DETERMINATION OF THE OPTIMUM PATH ON THE EARTH'S SURFACE

Marinos Kavouras

Department of Rural and Surveying Engineering
National Technical University of Athens
Zographou, Athens, GREECE 15780
e-mail: mkav@theseas.ntua.gr

Emmanuel Stefanakis

Department of Electrical and Computer Engineering
National Technical University of Athens
Zographou, Athens, GREECE 15780
e-mail: stefanak@theseas.ntua.gr

Abstract

Various algorithms have been proposed for the determination of the optimum paths in linear networks. Moving on a surface is a far more complex problem, where research has been scarce. An example would be the determination of the shortest sea course between two given ports. This paper presents a solution to the problem, which can be easily applied to a variety of surfaces, while considering different travel cost models. The applicability of the solution to representative surfaces is examined.

1 Introduction

The determination of the *optimum path* between two physical locations is a very common problem in applications such as Cartography, Robotics and Geographic Information Systems (GISs). Optimum in this context refers to a minimal accumulation of what amount to incremental travel costs associated with different media. It may be the shortest, fastest, least-expensive, or least-risky path.

When movement is restricted to the chains of a linear network, like road or railway networks, a weighted graph can be used as a model and associated algorithms [1,2] may be applied for the determination of the optimum path.

Moving on a surface is a far more complex problem, where research has been scarce. Examples would be the determination of the fastest path between two villages; the shortest sea course between two ports; the most regular gradient on ground path over a mountainous terrain; the least-risky path in a hostile environment (e.g., the path with the maximum concealment time vis-à-vis an enemy).

Clearly, the surface under study has its own peculiarities. For instance, it may be composed of a number of regions with different travel cost values assigned to them (e.g., walking on grass or sand); it may involve various means of travel (e.g., walking or driving); the direction of movement may introduce a variable cost value (e.g., *moving up-hill* or *down-hill*).

The problem of the optimum path finding has been examined in the past [3,4,5,6]. The proposed solutions suffer from several weaknesses [7], which have a negative impact to the determination of the optimum path. The objectives of this paper are twofold: first, to introduce a new approach to the optimum path finding problem (Section 2); and second, to examine its implementation to representative surfaces (Section 3).

2 A New Approach to the Optimum Path Finding Problem

This section introduces a new approach to the determination of the optimum path on a surface. The concept behind this approach is to establish a network connecting a finite number of locations (including departure and destination points) on the surface, so that effective algorithms coming from the weighted graph theory [1,2] can be adopted to indicate the optimum path for the desired trip.

The new approach consists of the following five steps:

2.1 Determination of a Finite Number of Spots on the Surface

The inconvenience that characterizes the movement on a surface is the infinite number of *spots* (i.e., locations) involved in the determination of the optimum path. The proposed solution to overcome this problem is based on the technique of *discretization of surface* under study. Discretization is the process of partitioning the continuous surface into a finite number of disjoint areas (cells), whose union results the surface. By representing each of these cells with one spot (e.g., its center point), a finite set of spots is generated to be involved in the process of the determination of the optimum path(s).

A wide variety of *tessellations* (also termed meshes) are available for obtaining the desired partitioning of the surface under study [8]. Tessellations may consist of regular or irregular cells. In the former case, the surface is partitioned by a repeatable pattern of regular polyhedra (regular polygons on plane surface); while in the latter case, the surface is partitioned by an extending configuration of polyhedra with variable shape and size. Some available tessellations are the regular, hexagonal or triangular grid for the plane surface; the geographic grid or the polyhedral tessellations for the spheroid. In order to achieve a uniform distribution of spots over the surface under study a regular tessellation should be adopted.

2.2 Establishment of a Network

After the determination of the spots over the surface under study, a *network* should be established to connect adjacent or non-adjacent spots and indicate the possible paths of movement.

The proposed scheme for the establishment of the network is based on the tessellation used for the generation of spots involved in the determination of the optimum path. Specifically, each spot is connected through network edges with the spots assigned to the neighbouring cells of its representative cell. Independently on the tessellation adopted to partition the surface under study, each cell has three types of neighbour cells: a) *direct,* i.e., neighbours with shared edges; b) *indirect,* i.e., neighbours with common vertices; and c) *remote neighbours.* Remote neighbours are characterized by the level of proximity to the cell of reference. For instance, level-one (level-two) remote neighbours are the cells which are direct or indirect neighbours of the direct or indirect neighbours of the cell of reference (of the level-one remote neighbours of the cell of reference). Note that, by increasing the number of neighbours considered, the directions of movements are augmented. For instance, in a regular tessellation on the plane surface, the direct neighbours introduce a set of four directions, the indirect neighbours another set of four directions, and the level-one remote neighbours another set of eight directions of movement. An exhaustive network would consider all direct, indirect and remote (of any level) neighbours for each spot.

2.3 Formation of the Travel Cost Model

The *travel cost model* assigns weights to the edges of the network established in the previous step. Its form depends on both the surface under study and the application needs. Examples of travel cost models are: the shortest path (minimum distance); the fastest path (minimum time); the least expensive path (minimum expenses); the least risky path (minimum risk).

The *model of the shortest path* is examined next in more detail. This is the simplest model to understand and it is used in the examples of Section 3. The surface under study is represented by a set of spots which may be *accessible* or *non-accessible* (i.e., lying on obstacles; usually not considered). For instance, spots lying on the sea are accessible, while those lying on the continents are non-accessible for a ship. The travel cost between two accessible spots is equal to the length of the shortest line connecting them, if they are *intervisible* [6]. Two spots are intervisible, if the shortest line connecting them passes through no obstacle. If this is not the case, the travel cost between the two spots is computed implicitly passing through intermediate spots.

2.4 Assignment of Accumulated Travel Cost Values

After the establishment of the network and the formation of the travel cost model, *accumulated travel cost values* from a spot of reference (i.e., departure or destination spot) to all spots of the surface under study (as chosen in step 1) can be assigned. This task is performed by adopting techniques from graph theory [1,2], concerning the determination of the optimum path on weighted graphs. Of the many classical algorithms, Dijkstra's algorithm is the most famous one.

2.5 Determination of the Optimum Path(s)

After the assignment of the accumulated travel cost values to the spots (nodes) of the network, the *determination of the optimum path(s)* from the spot of reference to a spot of interest (or the reverse) follows. Starting from the spot of interest, all its neighbours are examined and the spot(s) with the smallest cost value assigned to it (them) is (are) considered as the previous spot(s) of the optimum path(s). Starting from this (these) spot(s), the operation is repeated, until the spot of reference is reached.

3 Implementation

The scope of this section is to examine the applicability of the new approach in two representative surfaces: the *plane surface* and the *spherical surface* as an approximation of the earth's surface.

3.1 The Plane Surface

The plane surface is the simplest surface to study. The finite number of spots can be easily determined by using one of the available two-dimensional tessellations and locating one spot to the center point of each cell. The network is then established by connecting the spots through straight line segments, which are assigned weights depending on the structure of the surface and the travel cost model considered.

Figure 1 shows an example on the determination of the shortest path on a plane surface with obstacles (shaded areas). The tessellation used is the regular grid with a resolution of 10 units in both X and Y dimensions, while the established network connects each spot to its direct and indirect neighbour spots (Figure 1b) or direct, indirect and level-one remote neighbour spots (Figure 1c) by straight line segments (i.e., eight or sixteen direction of movements are considered). From the accumulated travel cost values assigned to the spots, the shortest paths are derived (solid lines in Figures 1b,c). Note that, by increasing the number of neighbours considered, a more accurate optimum path is derived, for the same set of spots (i.e., resolution of the tessellation) on the plane.

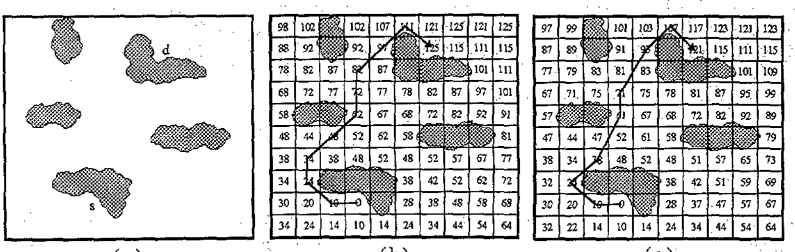

Figure 1: The shortest path on a plane surface (values assigned to the grid cells refer to the accumulated travel costs of the corresponding spots).

3.2 The Spherical Surface

The movement on the surface of a sphere is examined separately due to the peculiarities of the underlying spherical geometry. The determination of the spots is based on the available tessellations for the sphere. The tessellation which is closer to cartographers and geo-scientists, in general, is the geographic grid. Contrary to the plane surface, the path connecting two spots is not a straight line segment, but the shortest arc of the *great circle* passing trough them.

Figure 2 illustrates an example on the determination of the shortest sea course on the earth surface approximated by a sphere. For visualization purposes a Plate Carrée projection has been used to represent the continents and oceans of the earth. The solid line shows the shortest path from Adelaide, Australia to New Orleans, USA. The line does not approximate sufficiently the curve of the great circle on the projection plane, due to the low resolution of the grid (10 degrees) and the limited number of neighbours considered on the established network (direct and indirect).

The major problem faced by this tessellation is that the set of spots generated is non-uniformly distributed over the surface, because of the highly variable shape and size of the geographic grid cells; and as a consequence a variable accuracy on the determination of the optimum path is introduced, which depends on the region of the sphere studied. Clearly, there is a large concentration of spots around the poles which decreases by approaching the equator.

A more uniform distribution of the spots over the spherical surface can be obtained by adopting tessellations that are based on the recursive decomposition of regular polyhedra (e.g., the hierarchical model for the representation of the earth surface on a Global Geographic Information System introduced in [9]).

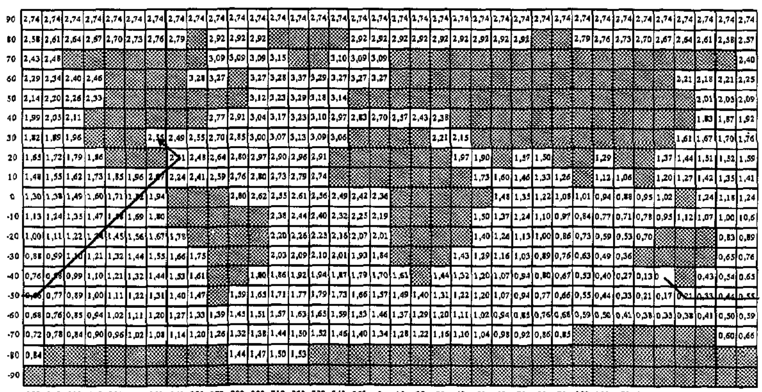

Figure 2: The shortest sea course (distances expressed in rad; shaded cells correspond to non-accessible spots, i.e., land; continents are simplified).

4 Conclusion

A new approach to the determination of the optimum path for free movement on a surface has been introduced. The general concept is based on the degeneration of the surface under study into a network, which can be simulated by a weighted graph, so that algorithms of graph theory can be easily adopted to indicate the optimum path for the desired trip. The new approach overcomes several problems faced by existing solutions and can be easily implemented to a wide variety of spaces with different travel cost models associated to them. The applicability of the new approach to surfaces used to represent the Earth has been examined.

References

[1] Gibbons, A., 1985. *Algorithmic Graph Theory*, Cambridge University Press Ltd.

[2] Sedgewick, R., 1990. *Algorithms*, Addison-Wesley Publishing Company, Inc.

[3] Warntz, W., 1957. "Transportation, Social Physics, and the Law of Refraction," *The Professional Geographer*, Vol.9, No.4, pp. 2-7.

[4] Lindgren, E.S., 1969. "A Study of the Movement of a Point on a Plane and in Space," *Harvard Papers in Theoretical Geography*, Cambridge, No.36, 16 pp.

[5] Douglas, D.H., 1994. "The Parsimonious Path Based on the Implicit Geometry in Gridded Data and on a Proper Slope Line Generated from It," *Proceedings of the International Symposium on Spatial Data Handling*, Edinburgh, Scotland, pp. 1133-1140.

[6] Aronoff, S., 1989. *Geographic Information Systems: A Management Perspective*, WDL Publications, Ottawa, Canada.

[7] Stefanakis, E., Kavouras, M., 1995. "On the Determination of the Optimum Path in Space," *Technical Report, KDBSL-95-06*, NTUA, 17 pp.

[8] Laurini, R., Thompson, D., 1992. *Fundamentals of Spatial Information Systems*, Academic Press Ltd.

[9] Goodchild, M.F., Shiren, Y., 1992. "A Hierarchical Spatial Data Structure for Global Geographic Information Systems," *Graphical Models and Image Processing*, Vol.54, No.1, pp. 31-44.