ALGORITHMIC DEVELOPMENT OF AN OPTIMAL PATH COMPUTATION MODEL BASED ON TOPOGRAPHIC MAP FEATURES

Mandy, Y.F. Tang[a], Lilian, S.C. Pun-Cheng[b]

Department of Land Surveying and Geo-Informatics
The Hong Kong Polytechnic University
Hunghom, Kowloon, Hong Kong
[a]lsgi.mandy@polyu.edu.hk , [b]lspun@polyu.edu.hk

KEYWORDS: GIS, Mapping, Modelling, Algorithms, Data Structures

Commission PS, Working Group IV/1

ABSTRACT:

Path finding problems have attracted widespread research interests with different GIS-T applications such as Logistics applications, Infrastructure Planning and Travel Demand analysis. Previous researches have largely been conducted on developing shortest path algorithms in GIS. The conventional approach is to adopt the arc-node network model. For example, a road is represented by a centerline that is formed between the two road margins. The interconnection of these lines and their intersecting nodes will form a network ready for path finding and computation of relevant parameters. However, the generation and maintenance of a centerline network is difficult and tedious because these are not natural but imaginary features. Human judgment and manual digitization are essential and will differ between individual operators in the creation of these lines. To remedy the situation, an alternative approach is suggested that path finding method is independent of any arc-node data structure. The network model for such computation is solely based only on feature outlines as appeared on the topographic maps. In other words, outlines or symbology of relevant features for a certain path finding application like road margins, building outlines, subways will directly be used to model the path finding network. This paper will investigate the rationale and logistics to develop such a model first for pedestrian walking. A small set of digital map data from a very congested Hong Kong urban area will be used to evaluate the model reliability and efficiency.

1. INTRODUCTION

1.1 General Instructions

With the development of geographical information systems (GIS) technology, network and transportation analyses within a GIS environment have become a common practice in many application areas. Many previous research studies focus on Geographical Information Systems for Transportation (GIS-T) have been published, including Vuren et al. (1987), Horowitz (1996), Sutton (1996), Zhan and Noon (1998), Kirkby et al. (1999), Chan and Yang (1999), Han et al. (2001) and Goodchild (2000).

The most common and convenient way to represent a road network is the node-arc representation. Generally, nodes correspond to street intersections while arcs correspond to street segments between intersections. In many applications, network arcs so called road centerlines are digitized to lie between the centers of physical road margins. In order to support routing application, all road centerlines should be snapped together to form a connected road network. This connectivity property of node-arc representation of road network allows system to traverse through the network efficiently with the support of network data storage structure.

However, road centerlines are not natural that can be found on ground. Data capturing techniques like surveying, photogrammetry and remote sensing are unlikely to capture this kind of data. It seems that manual digitization is one of the most common ways to generate the road centerlines to lie between the road margins. Although many GIS software are intended to automate the process of creating the centerline, the result could be unsatisfactory when some sophisticated cases are met. Therefore, the generation of centerlines could be cumbersome and time consuming.

The use of map features directly to perform path finding analysis provides an alternative method to replace the tedious generation of arc-node representations. Although there is no centerline shown on paper map, users can determine the accessible path based on both spatial features (e.g. vehicular road, walking path, buildings) and their descriptions (e.g. slope, stairs, turning directions) without the extra effort of generating another set of imaginable lines. Therefore, it should be possible to find a path with the use of base map features such as footbridge/elevated walkway, footpath, kerb line, pavement line, steps, subway and road/rail tunnel etc. provided in a digital map. This paper proposed an algorithm which is applied for a specific path-finding application – pedestrian walking.

In this paper, we developed a prototype that can be used to suggest walking paths for pedestrians with different scenarios of land feature configurations. Following the discussion of the current approach for finding path for drivers in section two, the idea of using map features to find walking paths is introduced in section three, and the initial results are presented there. The discussion of these preliminary results, suggestions for further work and conclusion are contained in the last section, section four.

2. NETWORK DATA MODEL

GIS provides users powerful tool to store, present and analyze geographical data digitally. In tradition, spatial data are presented on a paper map with the use of lines, symbols and color. But in GIS, spatial information have to be geometrically and semantically stored in a spatial database. The vector format is by far considered as an easily understood way of storing a feature semantically as it is.

Traditionally, vector representation has also been the domain of the network analysis in GIS. The linear network model is normally defined as a graph G, where G = (N, A) consists of (i) a finite set $N = \{n_1, n_2, \ldots, n_n\}$, whose elements are called nodes and (ii) a subset A of the Cartesian product A x A, the elements of which are called arcs. These two primitives: nodes and arcs represent the intersections and segments respectively in a continuous and connected linear network. Due to the connectivity property of the vector network, the complexities such as costs, distance, and time can be incorporated in the model easily. In the following discussions, road network will be specifically used for explanation because of its popularity in network application.

The conventional arc-node model is common and widely utilized to model the complexity of network elements in a logical way. The reason is clear. In a road network, arcs correspond to road segments that are the conduits for transportation and nodes correspond to road intersection connecting arcs together. This means, a connected network is consequently resulted to depict the complexities (such as turns, restrictions and lane information) of transport system no matter how sophisticated it is. Since the connectivity relationship of arcs and nodes and turn restrictions are implemented in attributes tables and turn table respectively, the optimum path between a source and destination(s) can be derived easily by traversing the topology of road network. It is just like performing a "search" operation in a branching tree with a parent root (the source) and corresponding child nodes (possibilities of destinations).

2.1 Problems Raised by Conventional Arc-Node Data Model

As stated in the beginning of this paper, the road centerline is an "abstract feature" to represent a road network. Therefore, generating road centerlines or developing the node-arc data model for network analysis are both important and difficult. It is important because road centerlines seem to be an essential element of a transportation network as discussed above; it is difficult because it involves tedious digitizing work. Although centerline mapping technologies are evolving rapidly, from traditional map digitizing to GPS, photogrammetry and remote sensing, each technology is associated with a performance range in terms of accuracy or resolution. There are many ways to model the transport system. Different definitions of roads, quality requirements or criteria make it difficult to have a consistent representation of transportation system. According to Dueker and Butler (2000), there are several GIS data models used for transportation applications. Prominent examples include Geographic Data File Standard (GDF), National Cooperative Highway Research Program 20-27 (NCHRP) and Dueker and Butler's enterprise GIS-T data model have been developed for transportation. However, these kinds of specifications do not lead to consistency due to their definitions and criteria differ.

Besides, the maintenance of a road centerline is hard. It is because the geometry and position of centerline are determined by the physical geometry of a road shown on a base map. The changes of physical geometry imply the changes of road centerline geometry and its related non-spatial attribute data. These may include node and arc identifiers, impedances and turns etc. Also, the existing arc-node representation of the road network results in a huge volume of data model. Especially in a planar network, the enforcement of a node at every intersection not only generates more turn possibilities at each junction, but also creates more arcs and nodes in the network. Take Hong Kong such a small territory as an example, its planar network consists of about 40,000 arcs and 30,000 nodes, with as many as 200,000 turn possibilities. Clearly, substantial amount of time are required for data input, preparation and validation in order to maintain a good quality road network and its associated attribute data.

In addition, the planarity of network does not represent well the real world properties of transportation networks that contain features such as "underpass" and "overpass" (Miller and Shaw 2001). Placing an additional node as under- or over-pass introduces a possibility of turning off the highway which does not make any sense in the real situation. To restrict those unrealistic turns, it can be done by assigning infinity impedance in turntable. However, this approach is an inefficient workaround.

With no doubt, data collection and maintenance are always the most expensive parts of a fully operational GIS. They could account for 60%-80% of the total cost in terms of time and money of a GIS project (Longley et al., 1999). Hence, the most convenient and efficient way to perform GIS analysis is to use the existing and already well-defined base map features directly, without attempting to generate a new supplementary data set for particular purpose. To accomplish this, it is assumed that digital map features are organized systematically with clear definition. Associated with user-specified application requirements, these map features are supposed to be sufficient enough to support several kinds of application. The following section provides an illustrative example of finding paths for pedestrians.

3. FINDING WALKING PATHS FOR PEDESTRIANS

As mentioned previously, path finding might alternatively be computed based on both user-defined relevat spatial features (e.g. vehicular road, walking path, buildings) of any geometry and their descriptions (e.g. stairs, turning directions) without the concept of using the road centerline on either a paper or digital map. Figure 1 illustrates an example of a digital base map and its organization/modeling of features in an associated database. The data structure is simply a product of any national or regional mapping agency, for no peculiar application. However, from users' understanding of a map, they may define the type(s) of features and attributes relevant to path finding. These will then form the basis for further computation as described next.

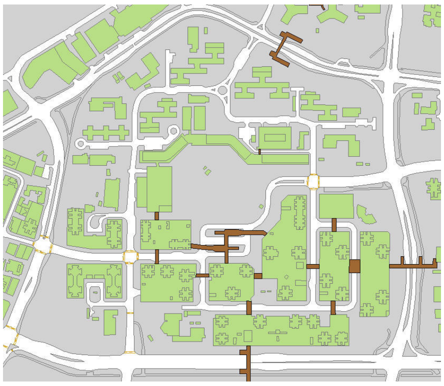

Figure 1(a): Example of digital data (Hong Kong, 1:5000)

Features	***Attributes***	***Descriptions***	***Path***
Street Blocks	ID	Unique Identifier	Y
	Type	Building blocks, barriers or traffic islands etc.	
Building Blocks	ID	Unique Identifier	N
Links (Bridges, zebra-crossings etc.)	ID	Unique Identifier	N
	Length	Length	
	Poly1	The first polygon ID	
	Poly2	The second polygon ID	
Roads	ID	Unique Identifier	N
	Type	Footbridges, steps, subways etc.	

Figure 1(b): Database descriptions

It is clear that the movement of cars is limited to one particular linear feature – road only. But for pedestrians, they are supposed to walk freely on all walking features (like pavements, stairs, elevated walkways) and on restricted places of roads (e.g. zebra crossing, traffic light areas). To start with, street block layer is chosen to be the key layer that presents the basic movement of pedestrians. Since other walking features such as steps, buildings, and pavements are all inside street blocks, the movement within street block is not our main concern here. The second important features to be considered are the "links" between street blocks. These "links" could be zebra-crossings, bridges or subways that enable pedestrians to move from one street block to another.

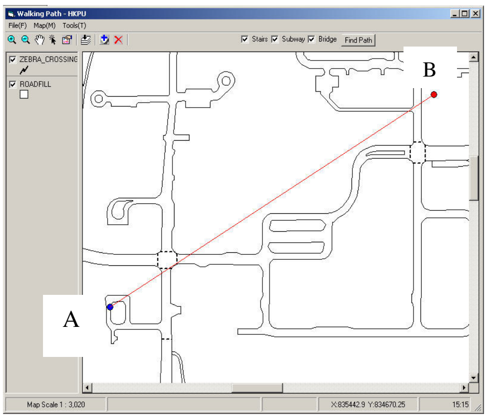

Figure 2: Example of computing a walking path between A and B

Suppose a walking path has to be computed between source A and destination B (Figure 2), the proposed algorithm is to construct a theoretical shortest line connecting these two places.

All calculations and and analyses are performed with reference to this theoretical line. This line is a vector which contains not only magnitude but also direction. The source is always considered as a from-node whereas the destination is considered as a to-node. After the direction is determined, the line is then broken down into several segments by extracting intersection points between the street block and "link" features. In the same way, each segment is considered as a directional line where the direction is the same as the drawn line. By overlaying the start and end point of each line segment with the polygonal street block layer, a list of polygons ID is resulted for further calculations (Figure 3).

Segment	Polygon ID	
	From	To
1	3	3
2	3	1
3	1	1
4	1	2
5	2	2

➡

Polygon ID
3
1
2

Figure 3: Polygon List

The resulting list indicates that to walk from A to B, the pedestrian is suggested to start at polygon 3, polygon 1 and then polygon 2. Then the next question is: how to walk across polygon 3 and 1, then polygon 1 and 2 on ground? To cross the road we need to find some features that connect the two street blocks, such as subways, bridges or zebra-crossings. Hence, the next step is to obtain the connectivity spatial relationship between the 'link' and street block features. This can be done by querying the attribute table of the "link" (Figure 4).

Attributes of links

FID	Shape*	ID	POLY1	POLY2
0	Polyline	1	88	143
1	Polyline	2	88	231
2	Polyline	3	203	231
3	Polyline	4	143	203
4	Polyline	5	112	143
5	Polyline	6	143	200
6	Polyline	7	200	203
7	Polyline	8	112	203
8	Polyline	9	221	230
9	Polyline	10	161	221

Figure 4: Attribute table of links

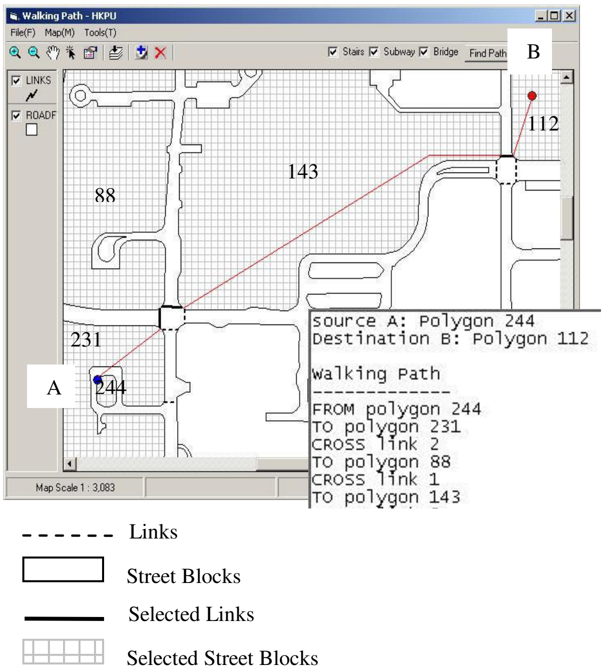

Figure 5: path between point A and point B

To know if there is any connectivity between any two polygons, say polygon 88 and polygon 231, just simply by making a query of "poly1=88 and poly2=231" or "poly1=231 and poly2=88". In this case, the record of link 2 is retrieved. Similarly, with the previous selected polygons, the result is obtained. Figure 5 shows the result of using base map features to find a walking path between two points. The pedestrian is suggested to start at polygon 244, polygon 231 and then cross the link 2 to reach polygon 88, cross the link 1 to reach polygon 143 followed by crossing the link 5 to arrive polygon 112, the destination. However, it is important to realize that land feature configuration in most cases is not that 'simple' as illustrated in the above example, but with a lot more complexities.

In Figure 6, there is a barrier between polygon 161 and 86 and there are two links connecting polygon 231 and polygon 203. In this application, barriers or traffic islands are not considered as parts of the solution because they are inaccessible for pedestrians. Therefore, these kinds of non-passable features are checked and removed from the polygon list (Figure 3) during the implementation of the algorithm. Furthermore, there are the possibilities of more than one existing link connecting two polygons (e.g. the two links between polygons 231 and 203).

With all these in mind, computation of the optimal walking path follows the algorithm as summarized below:

Step 1: Construct a line between the source and destination points
Step 2: Get all intersection points on the line
Step 3: Overlay the start and end point of each line segment with the street and construct a list with polygon ID
Step 4: Check inadequate polygons: to see if any physical dividers and traffic islands included in the above list. If yes, remove them.
Step 5: Check if there are more than one links connecting two polygons, if yes, compare the distance of each line with the constructed line. Choose the closest one.
Step 6: Query the link with provided polygon ID in the list to obtain the connectivity relationships between polygons
Step 7: Display the result

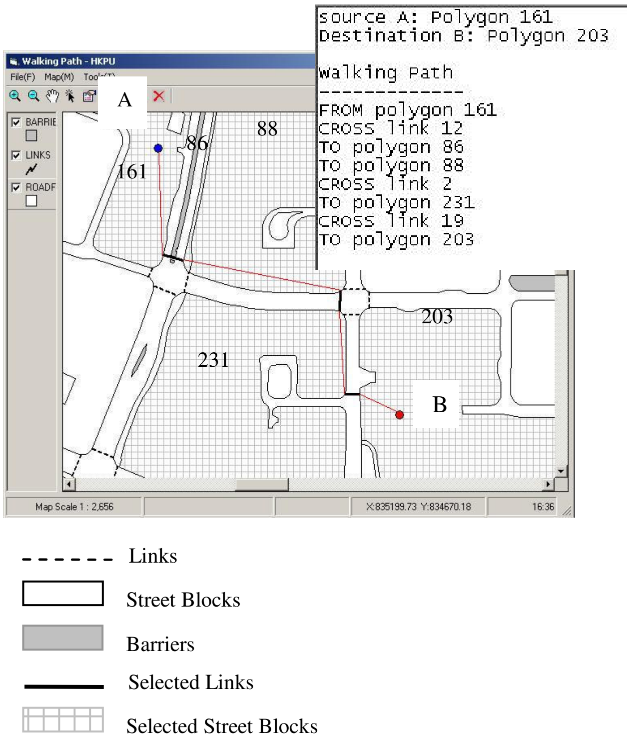

Figure 6: Computation with the existence of barrier and more than one link between polygons

4. DISCUSSION AND CONCLUSION

The study indicates that there is a potential to compute walking path without the use of road centerline. Although the work presented here lacks evaluation or examination, the results should be considered as encouraging. It is understandable that the accuracy required for walking path is relatively lower, when compared with accuracy required for driving path. But one of the important components of the future work should be included the evaluation part of the proposed algorithm to test the reliability of the model. Besides, the algorithm should be further developed and investigated with more walking path features such as building blocks, stairs, subways, bridges. As can be seen in Figure 5, pedestrians are allowed to move from one polygon to another polygon (from polygon 244 to polygon 231 in this case) without crossing any links in between. Therefore, some sophisticated assumptions or considerations are needed to take into account in the algorithm. For this particular case, the algorithm should consider the road width and the existence of barriers etc. in order to determine if it is suitable for pedestrians to cross. As a result, further analytical and theoretical works and studies are required to explore the feasibility of the proposed algorithm.

This paper introduces a new idea of using an independent of any arc-node data model to find path for pedestrians. In fact, the objective of transportation system is to improve individual accessibility. However, most GIS-T focus on vehicle navigation, individual accessibility is less common. This research could be important and useful for those people who interested in this area.

5. REFERENCES

Chan Yen Liang, Yang Hsu Hao (1999). Shortest Path in Traffic-light Networks. *Transportation Research* Part B 34 (2000) 241-253.

Dueker Kenneth J. and Bulter J.Allison (2000). A Geographic Information System Framework for Transportation Data Sharing. *Transportation Research* Part C 8 (2000) 13-26.

Goodchild Michael F. (2000). GIS and Transportation: Status and Challenges. *GeoInformatica* 4:2, 127-139, 2000.

Han Gang, Jang Jie and Chen Jun (2001). Optimizing Path Finding in Vehicle Navigation Considering Turn Penalties and Prohibitions. *ISPRS*, Vol.34, Part 2W2, "Dynamic and Multi-Dimensional GIS", Bangkok, May 23-25, 2001. pp.118-122

Horowitz Alan.J (1996). Integrating GIS Concepts into Transportation Network Data Structures. *Transportation Planning and Technol.*, Vol. 21, pp. 139-153.

Kirkby S.D., Pollitt S.E.P., Eklund P.W (1996). Implementing a Shortest Path Algorithm in a 3D GIS Environment. 96' *Spatial Data Handling Conference*, Delpht, The Netherlands, August 12-16, 1996, pp.7b31-42

Longley Paul A., GoodChild Michael F., Maguire David J. and Rhind David W. (1999) *Geographical Information Systems Volume 2 Management Issues and Applications*. New York, NY : Wiley.

Miller Harvey J. and Shaw Shih Lung (2001). *Geographic Information Systems for Transportation: Principles and Applications.* New York : Oxford University Press, 2001

Sutton John (1996). Data Attribution and Network Representation Issues in GIS and Transportation. *Transportation Planning and Technol*., Vol.21, pp.25-44.

Vuren Van T. and Jansen G.R.M (1987). Recent Developments in Path Finding Algorithms: A Review. *Transportation Planning and Technology*, 1988, Vol.12, pp.57-71

Zhan F.Benjamin and Noon Charles E. (1998). Shortest Path Algorithms: An Evaluation using Real Road Networks. *Transportation Science* Vol.32, No.1, pp.65-73

5.1 Acknowledgements

The work described in this paper was substantially supported by an associated fund - RPg grant from the Hong Kong Polytechnic University (Project No. G.34.9051).