Finding a Reliable Route between Two Points in a Transportation Network Using GIS

A Project Report Submitted for partial fulfillment of course requirements for CVEN658

Civil Engineering Applications of GIS

by

Prem Chand Devarasetty

December 6th 2010

Instructor,

Dr. Francisco Olivera

Texas A&M University,

Zachry Department of Civil Engineering.
ABSTRACT

In this paper an effort was made to find an optimal route between any two given points subject to some constraints. The constraints used in this study were related to the reliability of travel time. The reliability of travel time was defined based on the standard deviation of speed on that road or link. This reliability was attributed as a cost of the link with other costs such as average travel time of the link and length of the link. The optimal route was found by minimizing the cost of travel and travel time. Network analyst extension in ArcGIS was used to find the optimal route.

KEY WORDS: Optimal Route, Reliable Route, ArcGIS, Network Analyst
INTRODUCTION

Traditional route choice involves finding the shortest or the fastest route between any two given points. The shortest route is obtained by minimizing the distance between the origin and the destination point. While the fastest route is obtained by minimizing the travel time between those points. The travel time between any two given points A & B is stochastic in nature. It may depend on many variables such as traffic on that route, congestion, traffic incidents, weather, etc. Hence, the fastest route during off-peak hours may not be the fastest route during the peak hours; in short, it may not be a reliable route for travel. For finding a reliable route (route with less travel time variation) it requires to have information on the travel times and their variation on all possible routes connecting the given points.

GIS have been increasingly used as an efficient tool for a wide range of applications over the past decade. It is used for applications such as emergency response and transportation planning and management. GIS has evolved beyond the initial stages of data management and mapping, and have advanced into spheres such as modeling and analysis thus facilitating spatial decision making. Several attempts have been made to integrate different models into GIS and combine GIS with the development of Intelligent Transportation System (ITS) (Huang and Pan, 2006). This project aims at finding a reliable route between any two given points in a transportation network. Houston road network was used for this purpose.

The report is organized as follows: in the next section literature reviewed on related topics is presented, it is followed by a methodology section. In the section later, the results are presented. Finally the report is concluded with a brief conclusion part and directions for future research.

BACKGROUND

Route planning has been an active research area in transportation and logistics. Huang and Raguraman (2006) used bi-level Genetic Algorithm (GA) and Geographic Information System (GIS) to derive a route in a given transportation network which satisfy multiple objectives such as, minimal cost, minimum travel time, etc. The authors applied the above methodology to a tour operator problem, where the tour planner has to plan a tour that has to cover a certain number of visiting points such that it has the minimum cost, minimum travel
time, have most scenic beauty along the route, have least vehicle operating costs and maximum safety compared to other possible routes. The authors identified different ways to quantify the above variable (travel time, scenic beauty, etc.) from previous research and calculated those values for each link in the transportation network. The calculated values for each link were then coded to a GIS map. With the coded information and using the spatial analytical capabilities of GIS the importance of a link with respect to any given criterion was found. This information is used to obtain a generalized cost of each link which is critical variable for finding the best possible route. For each link, each of the variables was given a weight then, GA was used to obtain the best possible combination of weights. Once weights are known GA was again used to obtain the best sequence possible to cover all the given visiting points. The bi-level GA was code in C++ and the programs were as Dynamic Linkage Library which were linked to the Avenue script in ArcView. Once the program is run the GA finds the best possible route and sends the data to GIS then the final route map is displayed in GIS. A tour planner problem with 19 visiting points was solved in the article, the above mentioned method was able to find the best possible sequence of visiting points simultaneously optimizing the given variables.

Jakimavicius and Burinskiene (2010) developed a route planning methodology of an advanced traveler information system (ATIS) developed for Vilnius city. ATIS is very important, it presents information on real time traffic which can be used to predict traffic congestion thereby allowing travelers to make better decisions for their trip based on their preference of shortest or fastest route. The methodology presented in this paper involves a search for an optimal route between any two given points within the transportation network. Di-jkstra algorithm was used to find the optimal (shortest or fastest) route. Di-jkstra algorithm requires weights for each link in the transportation network to find the optimal route. The weights for links were calculated as a function of its travel time and segment length. The travel time was calculated based on the information on both real time traffic data and historic data on each link. The weights were made available in the attribute table of GIS and were updated every five minutes. A Standard ESRI ArcGIS Server Network extension was setup which enables to perform optimal route calculation tasks. Data on real time traffic collected by sensors on the road were transmitted to the server and updated in the GIS database. A WEB based information systems application was created for use by travelers. The WEB application allows users to input their trip information and also allows them to include restrictions on their trip (e.g. travelling from A to B with a stopover at a point B)
when submitted it shows the optimal route (both fastest and shortest) on a GIS map and it also gives public transport as an alternative option for the trip. This application was implemented in Vilnius city and was proven to be successful. The development of this automatic traffic management system in Vilnius city allowed increasing the average speed of traffic during peak traffic flow by 6%.

Travel time reliability is an important research area which is gaining importance these days. According to (Barry et al. 2005), in presence of substantial road congestion, the travel time variability is valued more than travel time savings. VOR indicates the value travelers’ place on the reliability of travel time. There have been several studies in the past trying to estimate the VOR. VOR is the travelers’ willingness to pay to reduce the variability of travel time by one unit. This variability in travel time is defined differently by different researchers. Few researchers assumed that the variability to be the difference between the 90th percentile and 50th percentile of the travel time (e.g. (Ghosh, 2001), (Lam and Small, 2001)). Whereas, few assumed it to be difference between the 75th and the 25th percentile of travel time (Small et al. 2005). Some researchers defined un-reliability as the standard deviation of the vehicle speed. In this project, the un-reliability of a link was defined by the standard deviation of the speed on that link.

**METHODOLOGY**

Standard ESRI ArcGIS Server Network Analyst extension enables to perform optimal route calculation tasks. This extension was used to find the optimal route. The optimal route was defined as the route which minimizes the travel time between any two given point and maximizes the cost associated with the travel. The cost for travel was defined solely on the basis of reliability of the links. As mentioned before the standard deviation of the speed data was used to define the un-reliability of the link. A higher standard deviation in speed implies that the link is unreliable, hence, a higher cost for travel. Similarly, a lower standard deviation of speed data implies that the link is more reliable, or, the travel time is more predictable, hence, a lower cost for travel through the link. Figure 1 shows the methodology followed in this project.
Sensors are located on all major highways in Houston. Aggregate vehicle speed data is collected by the sensors for every 30 second period interval. Data from those sensors was obtained from Houston Transtar office. Statistical Analysis Software (SAS) was then used to get the maximum, minimum, average, and standard deviation of the speeds of each sensor. During off peak hours most of the roads have traffic flowing at respective speed limits so the travel time is predictable. Whereas, during peak hours (6:00 AM to 8:00 AM, and 4:00 PM to 6:00 PM) most of the Freeways and roads near the downtown area are mostly congested and the travel time on those roads varies considerably. Hence finding a reliable route which has a predictable travel time is very helpful during peak hours. So for this project only peak hour data was used to define the cost of travel.

The speed data from the sensors was joined to the sensor location data using the join technique. The sensor data was then used to assign average speed and cost to all the Freeways in the network using spatial join function of ArcGIS. For local roads near downtown area the average speed during peak hours was set to 40 mph and the standard deviation of the speed was assumed to be 4 mph. For local roads far away from downtown area the average speed during peak hours was set to 45 mph and the standard deviation of the speed was assumed to be 2 mph. The road network data was then used to create a new network dataset. Using the network analyst
extension the optimal route between two points was found. Both travel time and reliability of the roads were used to define their respective cost attributes.

APPLICATION, RESULTS, and DISCUSSION

The study area of this project is Houston. The road network shapefiles were obtained from Houston-Galveston Area Council’s website. The geographic locations of the sensors and speed data were obtained from Houston Transtar office. Figure 2 shows the project area and Houston road network and sensor locations. Roads were classified based on the location of the roads and their functional class and separate shapefiles were created for each classification. Roads were classified as highways or freeways, roads near downtown, and roads far away from downtown (see Figure 3). The attribute values of freeways/highways, average speed and standard deviation of speeds were obtained using the spatial join based on location function of ArcGIS (see Figure 4). For roads near downtown and far from downtown the cost values are assigned as described in methodology section (see Figure 5 for reliability of roads). The three shapefiles were again merged using merge function to create the complete road network shapefile with cost attributes.
Figure 2: Houston Road Network and Sensor Locations
Figure 3: Road Classification
Network analyst extension was enabled and network analyst toolbar was added to ArcGIS. From ArcCatalog a new network data set was created using the road network shapefile (see Figure 6). Global turns were specified for turns. For the cost of the links, standard deviation of the speed was defined as the default cost function. Other cost functions of each link were travel time and length in miles (see Figure 7). Directions were also enabled for the dataset. The street names were used for referencing the links in the final route.
Reliability of Roads

Legend

Houston_Roads_Cost
AM_STDDev

- 2.000000 - 3.000000
- 3.000001 - 5.000000
- 5.000001 - 10.000000
- 10.000001 - 15.000000
- 15.000001 - 20.000000

Figure 5: Standard Deviation of Speeds on Each link
Figure 6: Building Network Dataset from Shapefile

Figure 7: Cost Attributes for the Network Dataset
Two points were chosen in the Houston area using network location tool and using the solve function of the network analyst the optimal route was found. Figure 8 shows the optimal route from location 1 to location 2 found by the network analyst. The route shown is the most reliable route between the given two points. The travel time of the route is better predictable than any other route possible between the two locations and also the shortest travel time. Since, the directions were also enabled we can also get the driving directions. Figure 9 shows the driving directions for the route shown in Figure 8.
CONCLUSIONS

The optimal route calculation method used in this project is based on the reliability of travel time of the network links. The reliability of a link was defined based on the standard deviation of the speeds. The higher the standard deviation, the lower reliability and was given a higher cost for travel through the link for the network calculations. On the other hand lower standard deviation of speed implies that travel times are more predictable and so is more reliable and was given a lower cost for travel through the link for network calculations. Speed data was available for only the freeways and highways. For other roads, the standard deviation of the speeds was assigned.

Figure 9: Driving Direction for the Route
based on the distance of the link from the downtown. Links near downtown were given a higher value for standard deviation of speeds and links far away from downtown were given a lower standard deviation for speed. Network analyst extension was used to find the optimal route between any two points. Costs for the links were assigned based on the reliability of travel time, travel time, and length of the link.

**FUTURE WORK**

The speed data for local roads was not available, the optimal route calculations would have been more accurate if I had the speed data for those roads. While choosing the optimal route the cost function was assumed as a linear function. However in reality, the cost or the standard deviation of the speed is not a linear function and cannot be added linearly. More sophisticated cost constraints can be used in future to find a more accurate and reliable route.
REFERENCES


