World Applied Sciences Journal 3 (Supple 1): 25-33, 2008
ISSN 1818-4952

New Method for Finding Optimal Path in Dynamic Networks

M. Alivand, A.A. Alesheikh and M.R. Malek

Department of GIS, K.N.Toosi University of Technology, Tehran, Iran

Abstract: Nowadays, urban traffic congestion is a complicated and ubiquitous problem. Continuous changes of traffic congestion with respect to the time lead to change the travel times of transportation network. These changes show the importance of time in transportation analyses in addition to the location. So determining the optimal path in a time-dependent transportation network is a challenging task.This article proposes the spatial analysis of finding the optimal path between two specific locations in a network that its traffic congestion changes continuously. Conventional algorithms are analyzed, and their features and constraints are evaluated. Then, a new method based on the partitioning space-time is presented. In the proposed method, some heuristic functions which are extracted from graph features have been used to guide the solution in each partition. Finally, a business model for collecting traffic data is introduced. Using this model could help the traffic managers to have real-time traffic data in different times and locations. Also, it could help the users to obtain the best route in their urban trips using GIS analyses like finding the optimal path.

Key words: Optimal Path Finding · Transportation Networks · Partitioning space-time · Heuristic methods · Mobile GIS

INTRODUCTION

The main characteristics of GIS that differentiate it from other information systems are its spatial data and geo-statistical analyses [1]. Such analytical functions usually present the best solutions to the users of GIS. Integration of these analyses with high visualization capabilities lead GIS to be widely used in decision making process [2].

The last decade has witnessed the rapid emergence of Internet-enabled mobile terminals, mobile computing, location aware technologies like GPS, and GIS capabilities. As a result of this integration, a new generation of GIS named mobile GIS have been developed which are capable of delivering geographic information and analyses to mobile users via the Internet and wireless networks [3]. Generally, Mobile Geospatial Information Systems (MGIS) is a type of GIS which its main research is about non-geographic moving object in geographical space. This research works on relationship between moving object and geographic entity, or moving object between another moving object. The main difference between a mobile GIS and traditional (static) GIS is the existence of a mobile agent [1].

However, many of human daily activities contain time in addition to space [4]. One of these important activities is transportation. In activities related to transportation, information changes both in space and time. Webster's dictionary defines transportation as "a means of conveyance or travel from one place to another". This definition implies that location and dynamics (movement in time) are critical components in transportation [5]. So, for solving many of transportation problems, the time constraints must be observed, in addition to spatial features. Considering this issue is very obvious in mobile GIS because the user is moving and its location is a function of time in addition to space. Consequently, solving many of analyses in mobile GIS needs time to be accounted for calculations [4].

Finding the optimal path is one of the most important analyses which are used in mobile GIS [2]. It is also one of the widely used applications in transportation problems. Finding optimal path is defined as "How to determine the optimal path from one or more places in network to other locations."

In many researches for finding the optimal path, time is regarded as a static parameter that leads to obtain similar solutions for a shortest path problem in different

Corresponding Author: M. Alivand, GIS Department, Faculty of Geodesy and Geomatics Eng., K.N.Toosi University of Technology

times [6]. Since traffic congestion changes continuously in a transportation network, so travel time of the links changes continuously in all the network locations. If the user needs to find a path with minimum travel time based on cost of travel, the previous methods will lose their efficiency and new algorithms must be developed in dynamic networks.The current article examines the spatial analysis of finding optimal path between two specific locations in a network which its traffic congestion changes continuously and a new algorithm based on partitioning space-time is proposed.

Optimal Path Between Two Specific Locations: Network analysis for transportation purposes is a common GIS practice. One of the most famous network analyses is finding the shortest path which delivers optimal path to the users based on the conventional algorithms. These algorithms determine the shortest path from a specific point to the other locations of a network. Generally, finding the shortest path is a type of optimization problems.

Finding shortest path between two specific locations is performed for ages. The first efforts for presenting a mathematical model for solving this problem were done by Euler and graph theory was proposed for modeling network representation and analysis. In 1959, Dijkstra suggested an algorithm for finding the shortest path which were used as the basic algorithm for many other algorithms up to now [7].

Definition of Optimal Path: It can be said that the optimal path is variable for different people. For example, the best route for a person who is finding a specific place in an unknown city, maybe the one which the probability of losing will be the least [8]. Since the mobile users of transportation network want to find a path with minimum time as cost, the optimal path is considered as finding a path between two specific points in transportation network which needs minimum time to traverse.

Graph Theory and Optimal Path: A weighted directed graph (or network) is a directed graph G(N,A) with real valued weights or lengths assigned to each edge [9]. N is a set of elements called *nodes* or *vertices*, and A is a set of unordered pairs of members of N called *arcs* or *links* [9]. The nodes of the graph are shown as points, while the edges are shown as lines connecting pairs of points. Equivalently, a weighted graph is a triple (N, A, w) where w is the weight of network links and maps the elements of E into real. The *length* of a path in a weighted directed graph is the sum of the lengths of the edges on the path. An *optimal path* (or *shortest path*) between a pair of vertices O and D in a weighted directed graph is a path from O to D with the least length [9].

Conceptually, the optimal path finding could be differentiated in a static network (with putting constant costs for links for traversing the network), and optimal path finding in a dynamic network (with putting variable costs for links for traversing the network).

Optimal Path Finding Algorithms in Static Networks: The network which the weights of its links are constant is called static network. In this state, the weight of all links is considered as fixed and the problem is solved globally for finding the optimal path between two nodes. Optimal path finding in static networks have been studied a lot during the years [3,4,9]. Many algorithms have been proposed for solving this problem. These algorithms are divided to one-to-one, one-to-all and all-to-all [3]. Traditional softwares which solve shortest path in a static network, usually calculate all-to-all shortest paths for all nodes and answer all the shortest paths questions based on them. The main basis of these algorithms is similar, and is based on Dijkstra's method.

Transportation Networks: Nowadays GIS faces more and more complex spatial-temporal queries; so time is regarded as a crucial component of GIS data models for solving spatial-temporal problems [5]. Sinton [4] defined time as a necessary component for solving spatial problems. Nevertheless, after more than 25 years, time is not examined much as an integral part of GIS.

In all the solutions for static shortest path that described in previous section, time is removed from the process of solving the problem. If the goal is to find the optimal route between two nodes with cost function on distance, it can be said that the problem would always have only one solution, but people want to receive optimal route based on travel time of each link as cost function. Consequently, new problems will arise with attention to changing time continuously which may cause wrong answers using the aforementioned static solutions [3].

A network that its links' weights change with respect to time is called Dynamic Network. Transportation network is a type of networks that have dynamic features which need more complex methods for solving the optimal path problem than static one. The most important parameter in transportation networks is travel time of each link Travel time of each link is a function of traffic congestion. Since traffic conditions changes continuously with respect to time, travel times change continuously. So it is important to consider the time as an important parameter for finding the optimal path in dynamic networks.

Many researchers have worked in shortest path problem in dynamic networks [3,6,10]. Generally the optimal path algorithms in dynamic state can be categorized into two classes [11]:

1. Time-dependent shortest path problem which network characteristics change with time in a predictable fashion. In this type, each link has a predictable function of travel time with respect to time.
2. Recalculation of optimal path due to consecutive, instantaneous and unpredictable changes in network data. Current methods solve this issue with re-optimization of a set of closely-related static shortest path problems.

Both of these algorithms have some problems and constraints which are examined in the next section.

Problems of Optimal Path Finding Analysis in Dynamic Networks: The first method needs many traffic data from different time and locations of the network in addition to continuous calculation the prediction function. If exact traffic data were available an exact prediction function in different times could be integrated with conventional algorithms for obtaining the optimal solution [6]. Since these traffic data do not be existed in many networks this method is not applicable for most of the cities. Also existence of many instantaneous phenomenon like accidents, roads constructions and special events lead to some disturbances in traffic prediction function which lead to the optimal solution does not achieve. So this method could not be used in many transportation networks.

The second method recalculates a new optimal path from the current location when traffic conditions change in the previous path. This method have been examined a lot because of the problems of the first method [3,10]. Nevertheless, this method has some problems for finding a reliable path as well. Figure 1 shows the problems of using the path recalculation method. In this figure, the first path is the one that is provided to the user from the beginning. Then because of traffic changes, recalculating algorithm have been used, and the second path have been provided from the determined node (A). Now, suppose that the current traffic data are available from the beginning, so the third path is better than the total travel time of the first and second paths. The occurred conditions, shows the much deviation of the path in the node A. In other words, that node has a high risk of deviation when it faces severe traffic congestions.

These constraints in previous algorithms lead to some unreliability in dynamic optimal path analysis.

Constraints of Optimal Path Finding in Tehran: ehran is one of those cities that have a lot of traffic problems. Traffic congestion imposes lots of economical and mental damages to the people. If suitable services can be presented to the users for finding optimal path with minimum time as cost function, then, it would prevents wasting the people's time in addition to decreasing traffic congestions. Now, there are few data from the different places of transportation network of Tehran. These data are gathered from some resources like computer vision and inductive loops, and etc. They only install in some constant places of network and couldn't sense in all locations and times; so, they do not present full spatial-temporal coverage in different places of network and different times of days and weeks, in addition to their low accuracy [12].

As a result, using a traffic prediction function as stated before, adds high unreliability to the computations (that has uncertainties itself). In this case, it may be better to present another patterns for finding the optimal path in dynamic transportation network of Tehran. This new pattern is based on a new vision that adapts the situation of transportation network of Tehran. In this method, the goal is to find a path that have few variance rather than the optimal path with current data, but if the traffic conditions changed, this path could modify itself with minimum changes in total time with respect to the current path. In other words, a path is selected as optimum which if accidents occur in that path, more paths will be generated from it to the destination.

Local Optimization: It is widely believed that in order to solve large scale global optimization problems an appropriate mixture of local approximation and global exploration is necessary. Optimal path finding in dynamic transportation networks is on of those optimization problems that could be solved better using local characteristics [5].

Characteristics of dynamic networks and problems of shortest path analysis in Tehran have examined until now

[13]. Now an algorithm is proposed for finding the optimal path for the users which have navigational tools in their cars.

The following features can be occurred in finding the optimal path in cities like Tehran:

1. Users are moving
2. Users do the shortest path analysis with a mobile terminal attached to a GPS
3. Users need shortest path analysis in an unknown environment
4. Users do not have the data (map, traffic and so on)
5. Traffic data does not exist in some regions
6. Real-time data are updated continuously on-line.

Above situations almost occur in all optimal path analysis in Tehran. They are features of a Mobile Environment that is based on mobile computing concepts [14,15]. In these environments, some algorithms based on partitioning space-time in continuous spaces for solving spatial problems have examined [14,15]. In such system that designed by Malek [15], current location of the user is determined by GPS and analyses is done in partitions of spaces and times in continuous space. In this section the idea of partitioning space-time in network models has been introduced. Using this method needs to use the local characteristics in determining the solution. In this method, the space of problem is divided to partitions and then shortest path analysis is done in each partition step by step until reaching the destination. This method will be more effective when there are some paths between origin and destination that have few time variances [10].

Finding the Optimal Path with Partitioning: The firs problem for finding the optimal path with partitioning is to partition space and time. There are different methods for partitioning space-time that could be divided into methods based on systematic and non-systematic partitions [9,14] In systematic approaches, space of problem divides to regular shapes. Greedy, tree and quad-tree methods are placed into these group. Non-systematic procedures which are more realistic with respect to the irregularity of time changing, divide space to unordered shapes. Partitioning space with Separator theorem is contained in this method which is the basis for many algorithms like parallel algorithms for solving shortest path problems [9]. Also some new methods based on Accessibility theory have mentioned recently [16].

One of the methods for finding the optimal path in dynamic networks is based on discretization of continuous time to smaller time intervals which traffic condition is constant in those intervals. Now, the space must be partitioned according to time divisions that lead to a partitioned space-time environment.

Figure 2 shows the first partition and the shortest path from the specific origin to specific destination. Figure 3 shows whole partitions which could be obtained using the end node of each partition.

The second and the most important point is finding the appropriate path in each partition. This solving must be such that the total provided paths in partitions do not deviate from the global optimally. The following example illustrates this problem.

Figure 4 shows a symbolic representation of a city. In this city the shortest path with minimum cost should be computed from A to the central point of city K. It could be seen that the region is organized from small partitions. Now, if the problem solved by local optimization, the determined path in Fig. 5 will be provided because it seeks the minimum weight link to traverse (this approach is the basis of Greedy method); instead, the global solving of problem leads to determined path in Fig. 6.

Consequently, if each partition solved independently, the optimal solution would not be obtained. So, to reach optimal solution it needs to integrate this method with heuristic methods. With this integration, optimal solution in each partition could be guided to destination. The features that could be used as heuristic are extracted from graph theory and will be explained in the next section.

Solutions for Integration of Partitioning and Heuristic Methods: After partitioning space and time, shortest path analysis could not be used independently in each partition because it will make the answer far from being optimally(like the above example in greedy method). For solving this problem, in each step a final node should be selected for each partition which is the first node of the second partition. The final node of each partition must be selected in a way that it would be on the destination direction to ensure that the provided path in each partition is guided to the destination correctly. Thus, functions that hold the direction must be used as heuristics. To obtain this heuristic, each final node at each partition could be related with a parameter which equals to direct distance to the destination (like the A* method for finding ht shortest path). In this case, the nodes that have lower direct distance to the destination could have higher weight in the computations. Also, first shortest path could be obtained globally without partitioning, and then a path is calculated in each partition which has the

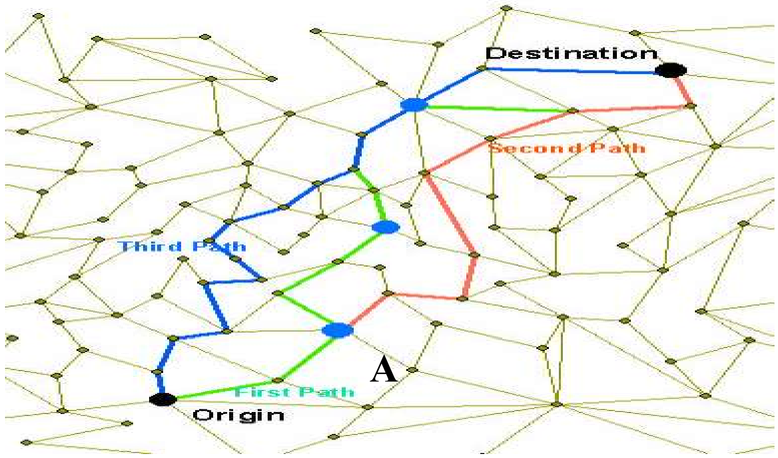

Fig. 1: The problems of path recalculation method for finding the optimal path

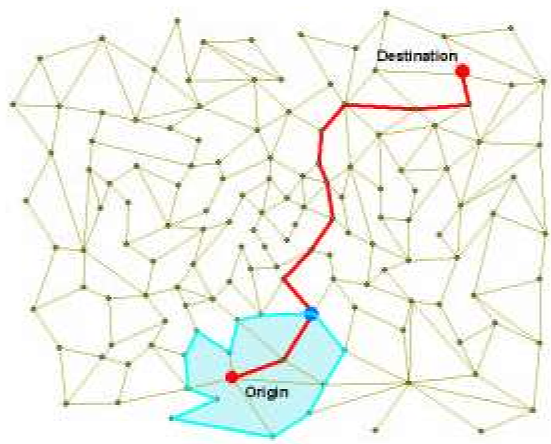

Fig. 2: Shortest path and first Partition

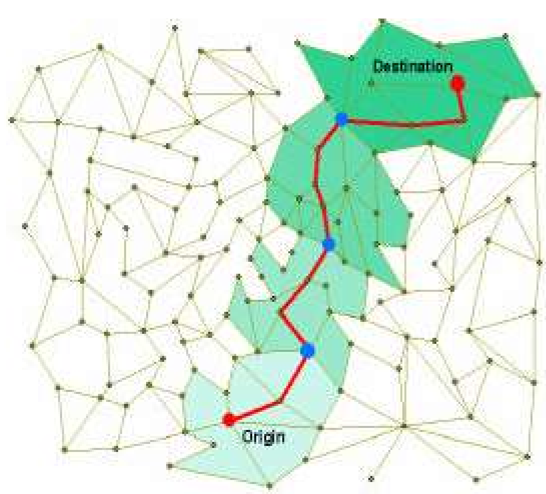

Fig. 3: Partitioning whole space-time

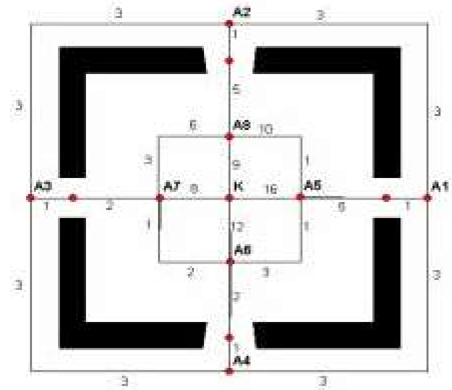

Fig. 4: The partitioned City

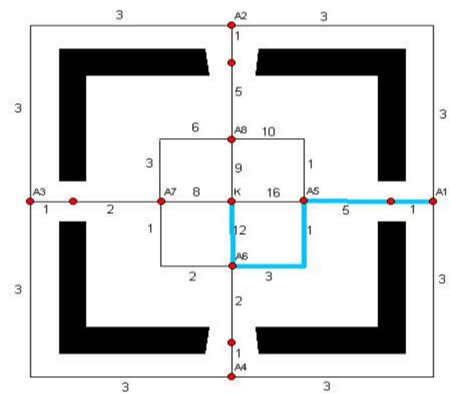

Fig. 5: Selected path with respect to local features

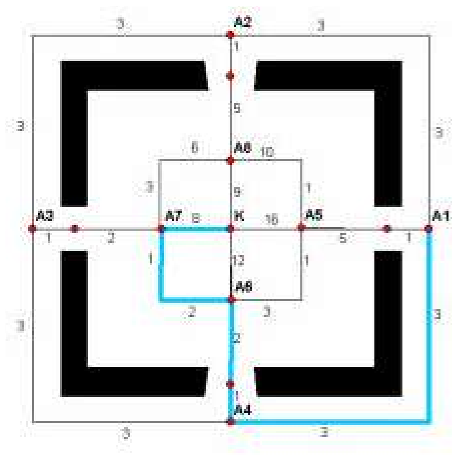

Fig. 6: Selected path with global features

minimum angle deviation from the global shortest path. After providing final node of the first partition with respect to the above criteria, start solving the problem in the second partition until reaching the last partition (Fig. 7).

As a result, first the necessary parameters for using in optimization problem will be obtained; then, the relevant cost function is determined with attention to these parameters that is a linear function. Now, the goal is to minimize the determined cost function.

The advantage of this method is that the space of problem is divided to smaller partitions, and only the traffic data of that partition is needed. So, at the time of solving the first partition, if an event occurs in other partitions, data of those partitions is updated before reaching the mobile user to them, and at the time of computations of that partition the updated traffic data will be examined.

A Business Scenario for Providing Traffic Data: It stated that in partitioning method traffic data of each partition are updated continuously before reaching the mobile agent to that partition. Now, resources of collecting traffic data are limited. Only a set of static sensors are placed into limited parts of network for estimating traffic volume which does not have high precision to estimate travel times at networks. Also, because of local limitations for installing these sensors, the collected data do not involve suitable spatial-temporal coverage. Using mobile cars as distributed mobile sensors is proposed by many researchers to solve this problem [12].

Collecting Traffic Data with Distributed Mobile Sensors: In this method, GPS is installed on some floating car data (FCD) in the network [12]. These GPS provide location data, instantaneous velocity and time, and store in a temporary memory. Then, these data will be sent to traffic management center via mobile communications for calculating travel time parameter in network (that is useful for shortest path analysis). Shortest path analysis could be run either in traffic center or in mobile client's computer. This is a type of smart client architecture that could be used both thin and thick client architecture [2].

Travel time of each link will gain from the time difference between first and final nodes of that link which GPS is registered [2]. These data will be sent to the center real-time and after on-line computation of these distributed mobile data together, travel times will send to the users. These data could be placed into Internet or could be sent to the users via mobile communication network [2]. Since installing and using these equipments have high costs, and also lack of an efficient system for exact management of these data, a business scenario for collecting traffic data with the use of spatial analysis has presented [12].

Introducing the Business Scenario: In this model, users that need on-line GIS's analyses in dynamic networks like finding optimal path, collect real-time traffic data themselves in the path. Users that need for such spatial analysis usually have equipments like GPS, pocket PC and etc in their own car; thus, they do not need to pay for installing these equipments, and only they pay for the communication to the center (with conventional approaches like Web) for sending the collected data will be paid [17]. Instead, they also can access to the suitable and efficient spatial analyses with attention to real-time collected traffic data. So, we send these analyses to people who gather traffic data, and it depends on interests of people for gaining online analysis [12]. To do this a client/server architecture has been chosen to comply with the users needs (Fig. 8).

CONCLUSION

Optimal path finding in transportation networks is one of the most important analyses in Mobile GIS and LBS. Information changes in transportation networks both in space and time; so, the solution is affected by time in addition to location. The cost of traveling is changing continuously due to traffic variants. So, it is inefficient to use static approaches for calculating the optimal path in dynamic networks.

Thus, dynamic methods and algorithms for computing the optimal path are presented that are based on discritiztion of time to smaller divisions and using static algorithms. This case leads to uncertainty on provided solution. In this paper a new vision based on partitioning space-time for solving shortest path problem is proposed. In this vision, heuristic methods in addition to optimization techniques are used. Implementation and test of this method shows that this pattern could be used as a new vision for calculating shortest path. However, to gain an exact model and algorithm needs more researches in this topic. Also, additional parameters must be extracted as new heuristics to choose the best end node at each partition.

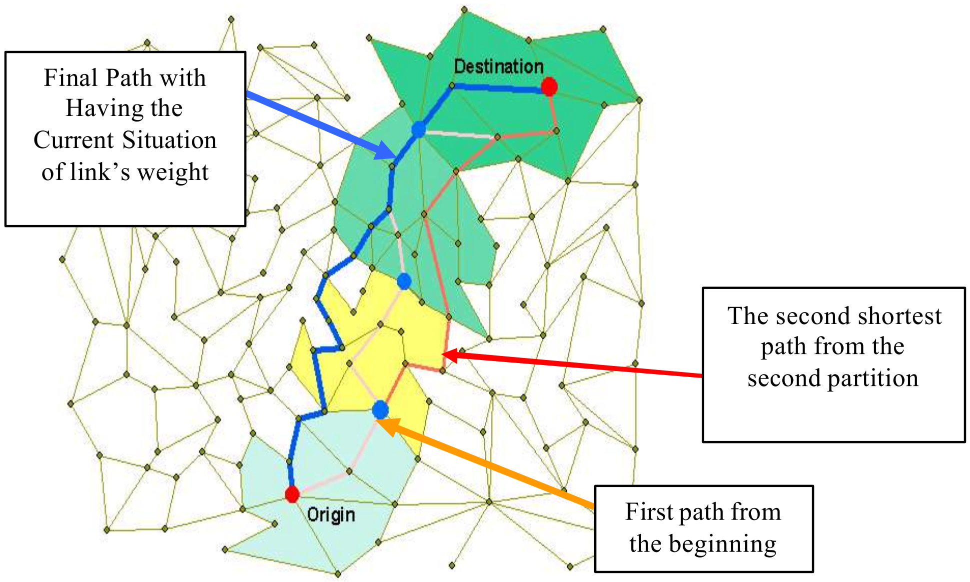

Fig. 7: Three paths can be generated in dynamic networks. Finding the third path from the beginning could be challenging

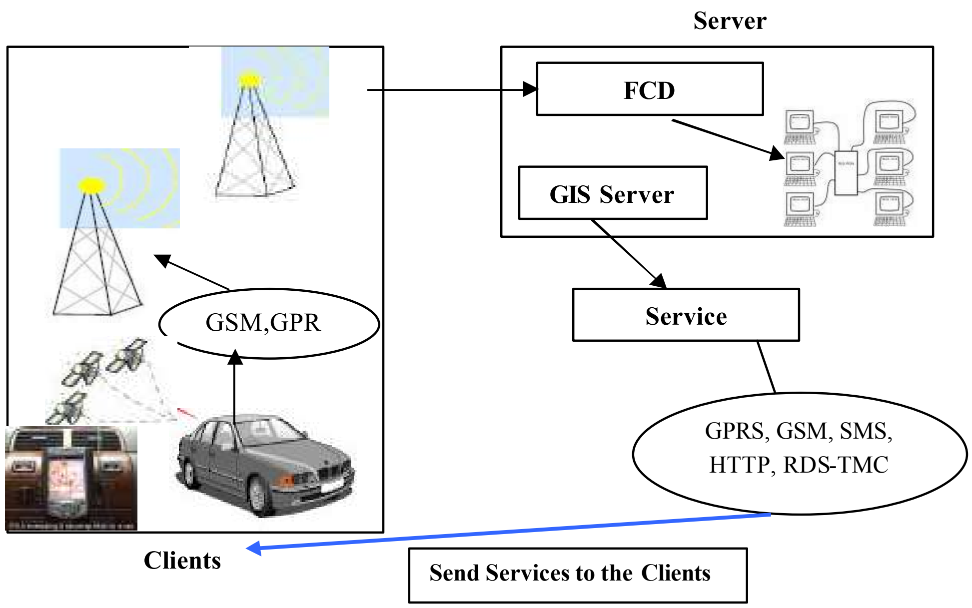

Fig. 8: Central Architecture for Implementing the System

Finally, due to the lack of suitable traffic data with appropriate spatial-temporal coverage, using a business scenario for collecting traffic data have been proposed and a confident architecture for implementation have been explained.

REFERENCES

1. Li, L., C. Li and Z. Lin, 2002. Investigation on the Concept Model of Mobile GIS, Symposium on Geospatial Theory, Processing and Applications, ISPRS, Ottawa, Canada.
2. Alivand, M., M.R. Malek and A. Alesheikh, 2007. Distributed Mobile GIS, Proc. of Map Asia Conference, Malaysia.
3. Wu, Q., 2006. Incremental Routing Algorithms for Dynamic Transportation Networks, M.Sc Thesis, Calgary University, Calgary, Alberta, Canada.
4. Neumann, A., 2005. Navigation in Space, Time and Topic, XXII International Cartographic Conference, A Coruna, Spain.
5. Koncz, N. and T.M. Adams, 2001.Temporal Data Constructs for Multidimensional Transportation GIS Applications, Department of Civil and Environmental Engineering, University of Wiscosin-Madison, USA.
6. Chabini, I., 1998. Discrete Dynamic Shortest Path Problems in Transportation Applications, Department of Civil and Environmental Engineering, MIT, USA.
7. Ahmadi, S., H. Ebadi, M.J. Valadan Zeoj, 2008. A New Method for Path Finding of Power Transmission Lines in Geospatial Information System Using Raster Networks and Minimum of Mean Algorithm, World Applied Sci. J., 3(2): 269-277.
8. Duckham, M., 2003. Simplest Path: Automated Selection for Navigation, Proc. of COSIT, Australia.
9. McHugh, J.A., 1990. Algorithmic Graph Theory, Prentice Hall, Englewood Cliffs, New Jersey 07632.
10. Dean, B.R., 2004. Shortest Paths in FIFO Time-Dependent Networks : Theory and Algorithms, Technical Report, Massachusetts Institute of Technology, USA.
11. Gentile, G., K. Noekel and L. Meschini, 2006.Time and Space Discretization in Dynamic Traffic Assignment Models, Department of Transportation, University of Rome, Italy.
12. Alivand, M., M.R. Malek and A. Alesheikh, 2007. Distributed Mobile Sensors for generating Traffic Information, Proc. of Location Asia Conference, Hong Kong.
13. Azizi, A., 2003. Network Analysis for Urban Traffic in GIS, M.Sc Thesis, Khajeh Nasir University, Tehran, Iran.
14. Malek, M.R., M.R. Delavar and A.U. Frank, 2006. A Logic-Based Foundation for Spatial Relationship in Mobile GIS Environment, Proc. of second International symposium on LBS & Telecartography, Vienna, Austria.
15. Malek, M.R., 2004. A Logic-based Framework for Qualitative Spatial Reasoning in Mobile GIS Environment, Lecture Note in Artificial Intelligence, 3066: 418-426.
16. Miller, H.J., 2005. Place-Based Versus People-Based Accessibility. In: Levinson, D. and K.J. Krizek (Eds.) Access to Destinations. Elsevier, London, pp: 63-89.
17. Amirian, P. and A. Alesheikh, 2008. A Service Oriented Framework for Disseminating Geospatial Data to Mobile, Desktop and Web Clients. World Applied Sci. J., 3(1): 140-153.