Analysis of Uncertainty Considerations in Path Finding Applications

Meysam Effati[1], Mohammad A. Rajabi[1], J. A. Rod Blais[2]
[1]Dept. of Geomatics Eng., University of Tehran, Tehran, 14395-515, Iran
Tel: +98 21 8833 4341, Fax: +98 21 8800 8837, Website: www.ut.ac.ir
[2]Dept. of Geomatics Eng., University of Calgary, Calgary, Alberta, T2N 1N4, Canada
Meysameffati@ut.ac.ir, Marajabi@ut.ac.ir, Blais@ucalgary.ca
Paper Reference Number: 0111-677
Name of the Presenter: Meysam Effati

Abstract

Geospatial Information Systems (GIS) has considerably advanced in recent years. However, the power of GIS specially in the context of analysis is limited by uncertainty. This uncertainty mainly comes from the data sets used and the methods employed.

Path finding is a term used mostly by GIS applications to plot the best route from one point to others. The basic idea behind path finding is searching a graph, starting at one point, and exploring adjacent nodes from there until the destination node is reached. Generally, the goal is of course to obtain the shortest route to the destination.

This article presents the experiments related to studying path finding under spatial uncertainties. The path finding is done on the web for part of road network of Tehran, Iran. More recently, fuzzy weighted graphs, along with generalizations of algorithms for finding optimal paths within them, have emerged as an adequate modeling tool for prohibitively complex and/or inherently imprecise systems. These algorithms are reviewed and formulized with uncertainty which comes from weights on edges according to actual situation on the road such as weather conditions, and road capacity at the specified time.

Keywords: GIS, Fuzzy Logic, Web Graphics User Interface (WGUI), Uncertainty, Path finding.

1. Introduction

In many applications such as transportation, routing, communications, economical, and so on, graphs emerge naturally as a mathematical model of the observed real world system [19]. Indeed, many problems can be reformulated as a quest for a path between two nodes in a graph which is optimal in the sense of a number of preset criteria [1]. Very often, these optimality criteria are evaluated in terms of weights, that is, vectors of real numbers, associated with the links of the graph. In the other words, a network takes the form of edges (or arcs) connecting pairs of nodes (or vertices) [2]. Nodes can be junctions and edges can be segments of a road. In a real-world model, an edge will have to be associated with a direction and with a measure of impedance, determining the resistance or travel cost along the network [14]. Graph theory has many different applications in a system analysis, economy and transportation. In some specific cases one must use data with uncertainty where the uncertainty can not be considered in the calculations over usually used graphs. Fuzzy Logic and Fuzzy Graph Theory provides proper tools to use in these cases [18]. The basic idea behind path finding is searching a graph, starting at one point, and exploring adjacent nodes from there until the destination node is reached. But this path finding maybe uncertain and this article introduces the experiments of studying path finding under spatial uncertainties for taxi network in Tehran and reformulating the shortest path problem as a fuzzy optimization problem: Indeed, by allowing that the crisp link weights expressing optimization criteria are replaced by fuzzy numbers (fuzzy weights), and by using the operations of fuzzy arithmetic, the problem becomes finding a fuzzy shortest path.

2. Literature review

All prior studies on path finding and path planning behavior assumed that all required spatial information was available but in real life navigators deal with incomplete or imprecise spatial knowledge resulting in spatial uncertainties [30]. Wiener et al. in 2008 presented experiments studying path planning under spatial uncertainties. In those researches a hierarchical planning scheme was used to navigate the shortest possible path to find an object hidden in one of four places and to bring it to the final destination [30]. Mao in 2008 designed and compared different path finding algorithms for a graph whose edge weights mutate randomly to a significant extent [21]. Nikolova et al. in 2006 presented new complexity results and efficient algorithms for optimal route planning in the presence of uncertainty. They employed a decision theoretic framework for defining the optimal route and identified a family of appropriate cost models and travel time distributions that were closed under convolution and physically valid [25]. Cornelis et al. in 2004 showed which criteria must be met for path finding algorithm correctness and explained an efficient method, based on defuzzification of fuzzy weights, for finding optimal paths [8].

3. Uncertainty

Geospatial data are often used under the assumption that they are free of errors [5]. Quality of geospatial data in GIS and the propagation of errors through GIS analyses have received much attention recently. The reliability of the resulting policy decisions very much depends on the quality of geospatial data used for reaching these decisions since the quality of the data affects the quality of the decisions and the evaluation of decision alternatives [33]. It is essential to acknowledge that uncertainty is inevitable [37]. Uncertainty is our imperfect and inexact knowledge of the world [17]. The term uncertainty is used to express the degree, not actual value, of discrepancy between geospatial data in GIS and the geospatial reality these data are intended to represent [5]. The uncertainty in the output of a GIS based analysis contains two parts [4]: First due to the propagation of uncertainty from the input data and second due to the error in the model on which the analysis is based. The goal of data quality analysis is to help understanding and managing the risk in making a particular decision involving the use of these data. In other words, one should perform risk assessment of a decision based on imperfect data. So it is necessary to perform operations in GIS with consideration of uncertain data. This paper tries to perform path finding under uncertainty in the input data.

3.1 Methods for uncertainty analysis

An error model refers to a stochastic process capable of simulating the range of possibilities known to exist for spatial data. These possibilities may exist because measuring instruments are known to be of limited accuracy, or because vital information, such as datum or map projection is missing [3].The methods by which geospatial data uncertainty can be modeled may be categorized to analytical equations for simple models and numerical approaches that involving the use of a computer for more complex models [15]. For quantitative uncertainty analysis, an iterative approach should be used.

When there is a possibility of obtaining an incorrect result even if exact values are available for all of the model parameters model uncertainty is used. The term model uncertainty is used to represent lack of confidence that the mathematical model is a "correct" formulation of the assessment problem [15]. The best method for assessing model uncertainties is through model validation [16]. Model validation is often limited because of lack of data, limited experimental opportunities, and inadequate financial resources.

3.1.1 Analytical methods for uncertainty analysis

The analytical approach most frequently used for uncertainty analysis of simple equations is variance propagation

n [20, 23]. A quantitative uncertainty analysis can be performed using analytical methods for statistical error propagation in relatively simple equations so for more complex calculations, variance propagation techniques are more difficult to apply analytically, and in some cases their use may not be practical or possible.

3.1.2 Numerical methods for uncertainty analysis

As such as analytical methods have some problems, to overcome these problems numerical methods are useful to performing an uncertainty analysis. Some of important numerical approaches include:

1) Monte Carlo simulation: An entirely different approach to the evaluation of the probabilistic structural demand is the Monte Carlo (MC) method [31]. Monte Carlo analysis is usually performed using two random sampling processes: Simple Random Sampling (SRS) and Latin Hypercube Sampling (LHS) [23].

2) Monte Carlo analysis of statistical simplifications of complex models [10, 22 and 24].

3) Differential uncertainty analysis [6, 34]: in this method the partial derivatives of the model response with respect to the parameters are used to estimate uncertainty.

4) Nonprobabilistic methods such as fuzzy sets, fuzzy arithmetic, and possibility theory [13].

5) First order analysis with Taylor expansions [28]: this approach is based the mathematical approaches on computer implementation for formulating the analytical solutions for error propagation.

6)

4. Path finding algorithms

A path finding algorithm for transit network is proposed to handle the special characteristics of transit networks such as city emergency handling and drive guiding system, in where the optimal paths have to be found. As the traffic condition among a city changes from time to time and there are usually a huge amounts of requests occur at any moment, it needs to quickly find the best path. Therefore, the efficiency of the algorithm is very important [7, 29 and 32]. The algorithm takes into account the overall level of services and service schedule on a route to determine the shortest path and transfer points. There are several methods for pathfinding:

In Dijkstra's algorithm the input of the algorithm consists of a weighted directed graph G and a source vertexs in Graph. Let's denote the set of all vertices in the graph G as V. Each edge of the graph is an ordered pair of vertices (u, v) representing a connection from vertex u to vertex v. The set of all edges is denoted E. Weights of edges are given by a weight function w: E → [0, ∞]; therefore w (u, v) is the non-negative cost of moving from vertex u to vertex v. The cost of an edge can be thought of as the distance between those two vertices. The cost of a path between two vertices is the sum of costs of the edges in that path. For a given pair of vertices s and t in V, the algorithm finds the path from s to t with lowest cost (i.e. the shortest path). It can also be used for finding costs of shortest paths from a single vertex s to all other vertices in the graph.

Best First Search (BFS) algorithm has some estimate (called a heuristic) of how far from the goal any vertex is, instead of selecting the vertex closest to the starting point, it selects the vertex closest to the goal. BFS is not guaranteed to find a shortest path and runs much quicker than Dijkstra's algorithm because it uses the heuristic function.

A* was developed in 1968 to combine heuristic approaches like BFS and formal approaches like Dijsktra's algorithm and can guarantee a shortest path. The A* algorithm integrates a heuristic into a search procedure [9]. A* is the most popular choice for pathfinding, because it's fairly flexible and can be used in a wide range of contexts [9]. Therefore, this research uses A* algorithm to perform path finding in the case study network.

5. Fuzzy Logic in GIS

Fuzzy sets are an extension of crisp (two valued) sets to handle the concept of partial truth, which enables the modeling of uncertainties of natural language [12]. Different to classical sets, elements of a fuzzy set have membership degrees to that set [12]. The degree of membership to a fuzzy set indicates the certainty (or uncertainty) [12]. With universe set X, Fuzzy set S is a subset of X that [36]:

$$S=\{(x,\mu_s(x)) \mid x \in X \wedge \mu_s \in [0,1]\} \quad (1)$$

Where, μ_s is the membership function.

The support of a fuzzy set S, written as supp(S), is the crisp subset of the referential set X defined by [36]:

$$\text{Supp(S)} = \{x \in X \mid \mu_S(x) > 0 \} \quad (2)$$

The support of a fuzzy set is the set of all elements of the universe that have a membership degree greater than zero [11].

The α-cut of a fuzzy set S, denoted by S_α, is the crisp subset of X that contains all of the elements of S with at least the given degree of membership α:

$$S_\alpha = \{x \in X \mid \mu_s(x) \geq \alpha\} \quad (3)$$

Any unary function or operation can be generalized to apply to fuzzy sets. Consider a fuzzy set S with measure μ_S membership function; the measure for f (S) is defined as Eq. 4:

$$\mu_{f(S)} = \sup_{y=f(x)} \{\mu_S(x)\} \quad (4)$$

Fuzzy Graph is an important element that it will be closely described in following equations [35]. A graph G consists of a set of vertices V and a set of edges E:

$$G = (V, E) \quad (5)$$

Where,

$$V = \{ v_1, v_2, ..., v_n \}, \quad (6)$$

$$E = \{ e_1, e_2, ..., e_m \}, \quad (7)$$

In the above function n is the number of vertices and m is the number of edges.

In this case study which is a weighted graph, each edge has a weight:

$$w_i = W(e_i) \quad (8)$$

A path Q is a sequence of edges:

$$Q = (e_{i1}, e_{i2}, ..., e_{im}) \quad (9)$$

If the graph is weighted, the path has a length given by the sum of the weights for the edges in the path:

$$l_q = \text{length (Q)} = w_1 + w_2 + ... + w_i \quad (10)$$

6. Shortest path algorithm with Fuzzy logic

In a fuzzy graph G, let λ be the set of all paths from vertex a to vertex b and the fuzzy length of a path be l_q that is introduced above.

$$l_q = \text{length (Q)} = w_1 + w_2 + ... + w_i = \sum_{e_k \in q} w_K \text{ , Q} \in \lambda \quad (11)$$

In the Eq. 11, e_k are edges of graph G. A fuzzy set S on λ with memberships σ_S is the fuzzy set of shortest paths:

$$\sigma_S \text{ (Q)} = \min \{ \mu_{l_Q \leq l_P} \}, \text{ where Q, P} \in \lambda \quad (12)$$

6.1 Methodology of implementation

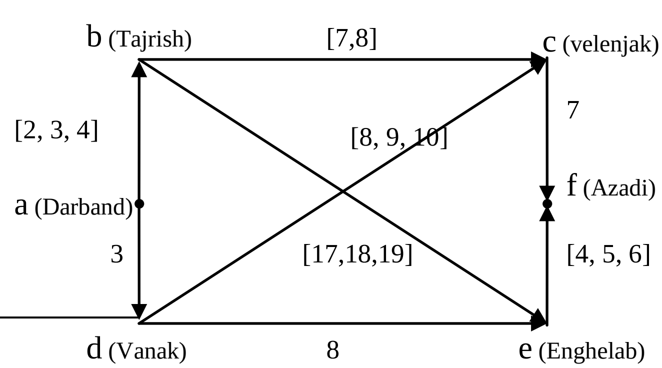
6.1.1 Data and study area

The data that have been used in this research are 1:50000 maps of study area that were obtained from National Cartographic Center [1] and taxi network paths of study area in Tehran that were obtained from Organization of Taxi Drivers[2] . Table 1 shows part of network data collected from OTD.

Table 1: Network data of taxi paths in part of Tehran

Id	Start	Destination	Line code	Going Path (km)	Return path (km)
1	a	b	1111	3.6	3.6
2	b	e	1112	18	18
3	p	s	1113	11	11.5
4	g	s	1114	14	14
5	b	d	1115	8	8.5
6	c	e	1116	7	8

Fig. 1 shows part of the taxi route that which fuzzy shortest path algorithm is implemented on it in three steps. There are uncertainties about weights on its edges according to actual situation on the road network (e.g. weather conditions, road capacity, and surface quality at the specified time.). These steps are:



[1] NCC
[2] OTD

Fig. 1: Taxis Fuzzy graph G

1- Construct graphs:
In this step G_1 and G_2, which are identical to G, are defined and the weights for the edges can be computed as Eq. 13 and Eq. 14:

$$\omega_{1a} = \sup\{\mathrm{supp}(w_1)\} \qquad \text{For } G_1 \qquad (13)$$
$$\omega_{2a} = \inf\{\mathrm{supp}(w_2)\} \qquad \text{For } G_2 \qquad (14)$$

2- Search for shortest path:
This step finds the shortest path Q from a to b in G_1. This is the shortest path problem and many well-known algorithms (which are explained above) can be used to solve it.

$$\mathrm{Supp(S)} = \{Q \in \lambda \,|\, \mu_{l_Q \le l_P} > 0, \forall P \in \lambda\} \qquad (15)$$

In above equation each edge e_i has a membership in the fuzzy set S':
Where

$$\mu_{S'}(i) = \max\{\sigma_{S'}(\mathrm{Q})\} \qquad \mathrm{i} = 1, \ldots, \mathrm{m} \qquad (16)$$

$$\mu_{S'}(i) = \min\{l_q\} \qquad q \in \lambda \qquad (17)$$

3- Calculate membership for each fuzzy path
If S_1 is the set of all paths from a to b in G_2 where their lengths are less than T (T is the length of path Q) and S is the set of all path in G, therefore, S is the set of all fuzzy shortest paths. Finally, membership for each fuzzy path from S with consideration of T is calculated.
In graph of taxi route shown in Figure 1 Vertex a is the source path (Darband) and f is the destination path (Azadi). The chart of fuzzy shortest paths for taxis fuzzy graph of Fig. 1 has been shown in Fig. 2.

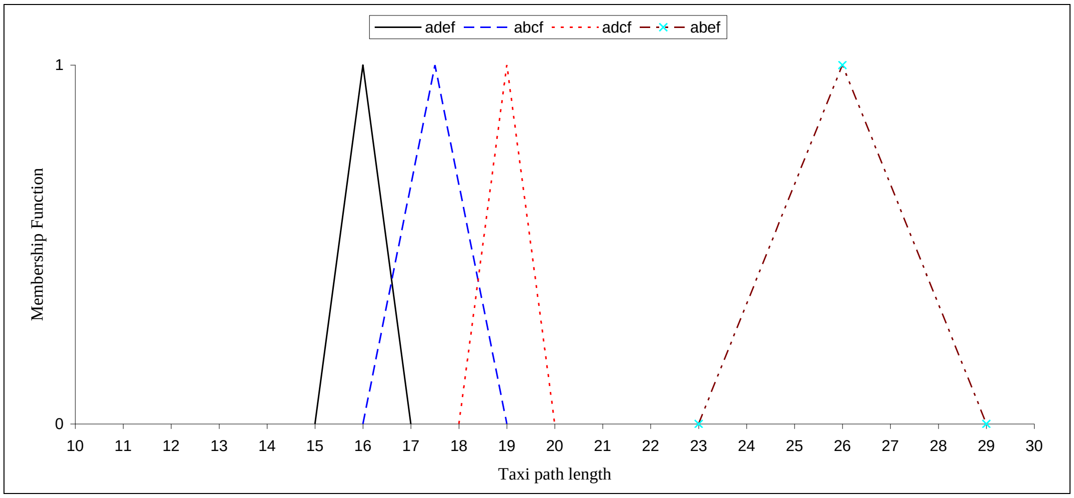


Fig. 2: Chart of fuzzy shortest paths for taxis fuzzy graph G

T=17 so the path adef has membership function $\sigma_S(adef)=1$ and the path abcf has membership function $\sigma_S(abcf)=0.5$ and other paths have the membership function zero $\sigma_S(adcf)=0$ and $\sigma_S(abef)=0$). An application of fuzzy shortest path algorithm is to plan the path of a vehicle which moves along the edges of a graph G [26], such as the taxi network which itself is an abstraction of the real road network. In this case we used data with uncertainty where the uncertainty could not be considered in the calculations over usually used graphs so fuzzy logic was used as a proper tool to handling uncertainty. By allowing that the crisp link weights expressing optimization criteria are replaced by fuzzy weights, and by using the operations of fuzzy arithmetic, the problem become finding a fuzzy shortest path. According to this result fuzzy model is a suitable model to generate the cost surface.

This optimum path is done on the server side with Access database. As such as Fig. 3 shows, users enters start and end points of the route through a Graphics User Interface (GUI) and passes them to the server through internet. The server performs optimum path finding based on the shortest path algorithms and web interface returns the best route as a picture to client.

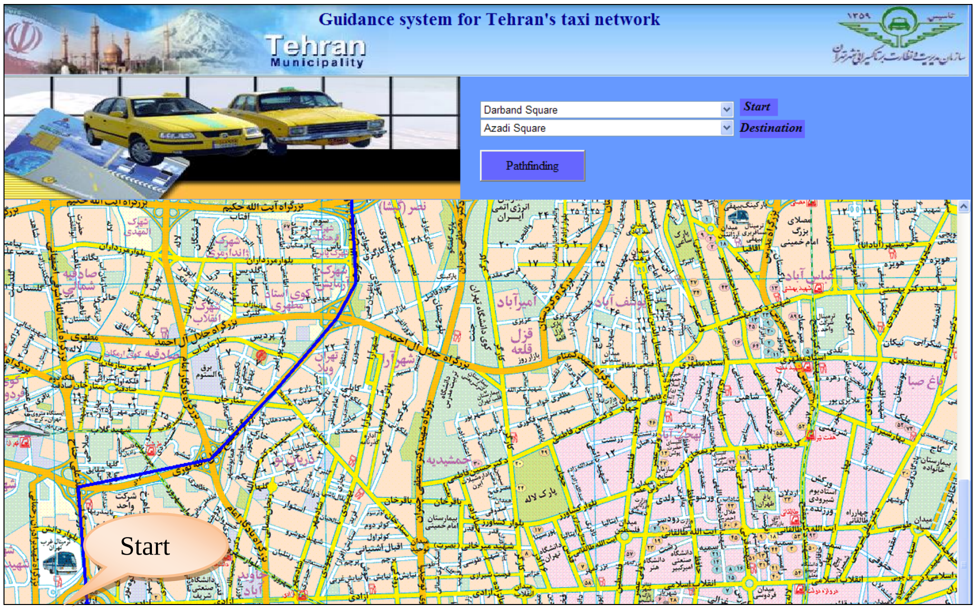


Fig. 3: Optimum taxi path on the web

7. Conclusions

The graph may be able to simulate a transportation network in order to find paths for pioneers even with uncertainty. The randomly mutating edge weights may represent an unknown cause of change in an environment, even if there is a systematic pattern to the change. In this paper we analyzed uncertainty in pathfinding using fuzzy logic and path finding is performed on a road network under uncertainties about weights on its edges according to actual situation on the road network. The optimum path showing to the client is a picture that user can not change and do any analysis on it. The future goal in this research is using neuro-fuzzy approach for adaptation of the fuzzy membership function of each fuzzy path according to the constructive learning concepts of neural networks. Another idea is performing fuzzy based path finding system on taxi network that gives more options to user for querying and analysis.

Reference

1. Aboelela, E. and Douligeris, C., (2000). *Fuzzy reasoning approaches for routing in B-ISDN.* Journal of Intell Fuzzy Syst, 9, pp. 11-27.

2. An, P. and Moon, W.M. (1991). *Application of fuzzy set theory for integration of geological, geophysical and remote sensing data.* Canadian Journal of Exploration Geophysics, Chapter 27, pp. 1-11.

3. Blais, J.A.R. and Boulianne, M. (1998). *Comparative analysis of information measures for digital image processing*. Int. Archive of Photogarmmetry and Remote Sensing, Vol. 27, Part B8, Commission 3, Kyoto, pp. 34-44.

4. Burrough, P.A. and McDonnell, R.A., (1998). *Principles of Geographical Information Systems*. Oxford: Oxford University Press.

5. Burrough, P. and Frank, A., (1996). *Geographic objects with indeterminate boundaries*. Taylor and Francis, London.

6. Cacuci, D. G. (1981). *Sensitivity Theory for Nonlinear Systems*. Part I and II. Journal of Math. Physics.

7. Chouy, Y.L., Edwin Romeijnz Robert, H. and Smithx, L. (1998). *Approximating Shortest Paths in Large-scale Networks with an Application to Intelligent Transportation Systems.*

8. Cornelis, C., Kesel, P. D. and Kerre E. E. (2004). *Shortest Paths in Fuzzy Weighted Graphs*. International journal of inelegance system, 19, pp. 1051-1068.

9. Dai, L. (2005). *Fast Shortest Path Algorithm for Road Network and Implementation*. Carleton University School of Computer Science.

10. Downing, D. J., Gardner, R. H., and Hoffman, F. O. (1985). *An Examination of Response-Surface Methodologies for Uncertainty Analysis in Assessment Models.* Technimetrics 27, pp.151-163.

11. Dubois, D. and Prade, H. (1980). *Fuzzy sets and systems: Theory and applications.* New York, Academic Press.

12. Engelbrecht, A.P. (2007). *Computational Intelligence: An Introduction*. Second Edition, John Wiley & Sons, Ltd.

13. Ferson, S. and Kuhn, R. (1992). *Propagating uncertainty in ecological risk analysis using interval and fuzzy arithmetic*. In, P. Zannetti (ed.). Computer Techniques in environmental Studies IV, Elsevier Applied Science, London, pp. 387-401.

14. Goodchild, M.F. (1977). *An evaluation of lattice solutions to the problem of corridor location.* Journal of Environment and Planning, pp. 727-738.

15. Hammonds, J. S., Hoffman, F. O. and Bartell, S .M. (1994). *An Introductory Guide to Uncertainty Analysis in Environmental and Health Risk Assessment*. SENES Oak Ridge, Inc. Oak Ridge, Tennessee under direction from the Environmental Restoration Risk Assessment Council, pp.3-19

16. Hoffman, F. O., Miller, C. W., and Ng, Y. C. (1984). *Uncertainties in Radioecological Assessment Models.* The environmental transfer to man of radionuclides released from nuclear installations Conference, Brussels.

17. Klir, G.J. and Folger, T.A. (1988). *Fuzzy sets, uncertainty and information.* Prentice-Hall Int. Editions.

18. Koczy, L.T. (1992). *Fuzzy graphs in the evaluation and optimization of networks, Fuzzy Sets and Systems.* 46, pp. 310-319.

19. Lin, K. and Chen, M. (1993). *The fuzzy shortest path problem and its most vital arcs.* Journal of Fuzzy Set Syst, 58, pp. 343-353.

20. Martz, H. F. and Waller, R. A. (1982). *Bayesian Reliability Analysis.* John Wiley & Sons, New York.

21. Mao, H. (2008). *Path finding Algorithms for Mutating Weight Graphs:* Project Proposal.

22. Mead, R. and Pike, D. J. (1975). *A Review of Response Surface Methodology from a Biometric Viewpoint.* Biometrics 31, pp. 803-851.

23. Morgan, M. G. and Henrion, M. (1990). *Uncertainty, A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis.* Cambridge University Press, New York.

24. Morton, R. H. (1983). *Response Surface Methodology.* Mathematical Scientist 8, pp. 31-52.

25. Nikolova, E., Brand, M. and Karger, D.R. (2006). *Optimal Route Planning under Uncertainty*. American Association for Artificial Intelligence.

26. Okada, S. and Soper T. (2000). *A shortest path problem on a network with fuzzy arc lengths.* journal of Fuzzy Sets and Systems, 109(12),pp. 129-140

27. Petrík, S., Madarász, L, Ádám, N. and Vokorokos, L. (2002). *Application of Shortest Path Algorithm to GIS using Fuzzy Logic*. Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovakia ,pp.1-8.

28. Scavia, D., Powers, W. F., Canale, R. P. and Moody, J. L. (1981). *Comparison of First-Order Error Analysis and Monte Carlo Simulation in Time-Dependent Lake Eutrophication Models.* Water Resources Research 17, pp.1051-1059.

29. Shad, R., Ebadi, H. and Ghods , M. (2003). *Evaluation of Route Finding Methods in GIS Application.* Dept of Geodesy and Geomatics Eng. K.N.Toosi University of Technology, IRAN.

30. Wiener, J. M., Lafon, M. and Berthoz, B. (2008). *Path planning under spatial uncertainty*. Memory & Cognition, 36 (3), pp. 495-504.

31. William, P. and Steven, M. (1997). *Guiding Principles for Monte Carlo Analysis.* Technical Panel, Risk Assessment Forum U.S. Environmental Protection Agency Washington, DC 20460.

32. Willhalm, T. (2005). *Speed Up Shortest-Path Computations*. http://www.mpi-sb.mpg.de/~sanders/courses/algen04/willhalm.pdf

33. Worboys, M. (2003). *GIS a computing perspective.* Second Edition, chapter 9.

34. Worley, B. A. (1987). *Deterministic Uncertainty Analysis*. ORNL-6428. Oak Ridge National Laboratory, Oak Ridge, Tennessee.

35. Zadeh, L.A., (1983). *A Fuzzy set theoretic approach to the compositionality of meaning: propositions, dispositions and canonical forms*. journal of semantics, 2(3).

36. Zadeh, L.A. (1965). *Fuzzy Sets, Information and control 8, 338-353.*California.

37. Zhang, J. and Goodchild, M.F. (2002). *Uncertainty in Geographical Information."* first published 2002 by Taylor & Francis.Chapter1, 2, 4.