A GIS - based Pipelining Using Fuzzy Logic and Statistical Models

***Hamid Kiavarz Moghaddam*†, *Mahmoud Reza Delavar*††**

† MSc.Graduate, Dept. of Surveying and Geomatics, Eng. Faculty, University of Tehran, Tehran. Iran
†† Assistant Professor, Center of Excellence in Geomatics Eng. and Disaster Management, Dept. of Surveying and Geomatics Eng., Eng.Facility University of Tehran ,Tehran, Iran

Summary

Oil and gas industry is an important pat of Iran's national economy. In this paper, geospatial information system (GIS) application in optimum route finding in raster model has been considered.
Effective criteria have been weighted using knowledge driven and data driven methods. Index overlay, fuzzy logic and statistical methods have been used for generating cost surface.
For creating accumulated cost surface, isotropic, partially anisotropic and full anisotropic models have been investigated. Since oil and gas pipeline route finding depends on location and direction of slope which varies in every direction, full anisotropic model has been implemented.
In case study, for calculating the weight of factors, a pipeline between Bidboland (in Chahar Mahale Bakhtiari Province) in central part of Iran and Farah (in Kohkoloyeh and Boyer Ahmad Province) with 25 kilometers length has been selected. Cartographic and topographic maps at a scale of 1:25000 and geology map at a scale of 1:100000 were collected. Factor layers have been weighted using knowledge and data driven methods and then integrated using a number of integrated methods. A number of evolution models including distance to existing route, calculated cost and the number of effective criteria have been tested. Using the three evaluation models, it is found that the path determined from fuzzy gamma cost surface is the best route. By error propagation method, the spatial error with 5.065 Meters has been calculated. The value of fuzzy gamma has been determined in Bidboland-Farah route. These values have been implemented in Ahvaz-Maron gas pipeline and the analysis indicated that it would be 31% cheaper than the existing pipeline route.

Key words:
GIS, Cost Surface, Fuzzy, Isotropic, Overlay, Weight of evidence

1. Introduction

A network model can be defined as a line graph, which is composed of links representing linear channels of flow and nodes representing their connections [10]. In the other words, a network takes the form of edges (or arcs) connecting pairs of nodes (or vertices) [1]. Nodes can be junctions and edges can be segments of a road or a pipeline. For a network to function as 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 [9].

Traditionally, a GIS represents the real world in either one of two spatial models, vector-based, i.e. points, lines and polygons, or raster-based, i.e. cells of a continuous grid surface [7]. Since the modeling of network structures intuitively refers to lines and points, vector GIS has dominated the realm of network analysis. This research will show that network analysis is equally feasible in raster GIS [3]. Because path finding is applicable to a number of networks such as pipelines, roads, utilities, water, electricity, telecommunications and computer networks, the total number of algorithms developed over the years is immense [6]. When finding least cost path in raster model the following points are essentially important [4]:

(i) Identify influencing factors for oil and gas pipelining. The results of investigation show influence factors are: elevation, accessibility roads, environment, ground cover, natural and geology critria.

(ii) Weighting influence factors based on the knowledge driven, data driven and integrated models.
(iii) Generate cost surface based on raster overlay models.
(iv) Generate accumulated cost surface from cost surface.
(v) Calculate least cost path.

2. Theoretical Consideration

This section describes the theorical models implemented for pipelining in raster GIS.

2.1. Weighting Theory

The weight of each factor map indicates the importance of its cost and value comparing with the other factor maps. Estimating correct weights can help finding cheaper cost surface. There are two knowledge driven and data driven weighting models [5].

2.1.1. Knowledge driven model

In knowledge driven model, influencing factors are weighted in a definite range using expert experience based on particular application. First, different ideas are collected and their dimensions are uniformed. Then, weights are normalized in the defined range using an appropriate scale [12].

2.1.2. Data driven model

In data driven model, influencing factors are weighted by calculating dependency value between factor maps and specified layer. One of the appropriate methods for determining dependency value is weight of evidence [12].

2.1.2.1. Weight of evidence model

Weight of evidence is a probabilistic model that can be used for integration of factor maps. This model is executed based on probability theory and rules similar to Bayesian model. This model weighted to influencing factors and create recent(a posteriori) probabilistic maps by probabilistic and statistic models [3].
To create recent probabilistic maps, two sufficiency and necessity ratios parameters calculated using Eq.1 and 2 [11]:

$$LN = \frac{P(\overline{B_i} | D)}{P(\overline{B_i} | \overline{D})} \qquad (1)$$

$$LS = \frac{P(B_i | D)}{P(B_i | \overline{D})} \qquad (2)$$

where B_i denotes the effective zone in each factor, $\overline{B_i}$ denotes the supplemental zone, D denotes the existing pipeline, LN denotes the necessity ratio and LS denotes sufficiency ratio [10].
$P(B_i | D)$ is conditional probability calculated by $\frac{P(B_i \cap D)}{P(D)}$. The value of LN and LS declare accordance effective zone and the supplemental zone of factors with existing pipeline.
The idea is to calculate the recent probability , $P(D \cap B_1 \cap B_2 \cap B_3 \cap ... \cap Bn)$ (Eq. 3, 4). The probability shows the occurrences. Occurrences are the ratio of the occurred probability as follows (Eq. 5) [3]:

$$O(D | B_1 \cap B_2 \cap B_3 ... \cap B_n) = O(D) \prod_{i=1}^{n} (LS_i or LN_i) \qquad (3)$$

$$P(D | B_1 \cap B_2 \cap B_{3...} \cap B_n) = \frac{O(D | B_1 \cap B_2 \cap B_{3...} \cap B_n)}{1 + O(D | B_1 \cap B_2 \cap B_{3...} \cap B_n)} \qquad (4)$$

$$O(D) = \frac{P(D)}{1 - P(D)} \qquad (5)$$

In Eq. 3-5, $O(D | B_1 \cap B_2 \cap B_3 ... \cap B_n)$ explains occurrences and $P(D | B_1 \cap B_2 \cap B_{3...} \cap B_n)$ explains the occurred probability of D .So the value of recent probability has been calculated for every pixel in the output map [6].

2.1.2.2. Weight calculation

By naperian logarithm computation, LN and LS , positive weight of evidence and negative weight of evidence for each factor map can be determined. Finally, by computing constant parameter called factor map, dependency value is determined (Eq. 6) [6].

$$C = | W_+ - W_- | = | Ln(LN) - Ln(LS) | \qquad (6)$$

where W_+ denotes positive weight of evidence, W_- denotes negative weight of evidence, Ln denotes naperian logarithm and C denotes factor map contrast calculated for every effective criteria .C shows the correlation between every effective criteria with exist pipeline. Data driven weighting has less errors than knowledge driven, however, its correct operation depends on existing pipeline. In our evaluation, because of inaccuracy in locations of existing pipeline layer, we use data and knowledge driven methods have been used for weighting the effective criteria [10].

2.2. Integration Models

Various models have been used for real world events simulation in GIS environment. Integration model is one of them that integrate related spatial data and effective criteria [7]. Some models have been tested in this paper including Boolean operation, indexing overlay and fuzzy logic.

2.2.1. Boolean Operation

In Boolean model, input maps are integrated using logical operators such as AND, OR, XOR and NOT. Although, Boolean operation is an easy and fast model to run, there are some problems in its execution routine. In this model, all input factor maps (processed data) have the same weights and appropriate sites can not be separated based on their priorities. This model has not been used because of the mentioned problems and nature of effective criteria and conditions considered [8].

2.2.2 Index Overlay Model

Factor maps are integrated using Eq. 7 [12]:

$$S = \frac{\sum_{1}^{n} \mathrm{S}_{\mathrm{ij}} w_i}{\sum_{1}^{n} w_i} \qquad (7)$$

where w_i denotes the weight of ith factor map, S_{ij} denotes the ith spatial class related to the weight of jth factor map and S denotes the spatial unit value in output map.

Comparing index overlay with boolean models routines, it is identified that index overlay model has more flexibility and ability for priority indication on spatial units of factor maps. With respect to the mentioned characteristics, this model is useful for comparing and evaluating integration models in generating cost surface [11].

2.2.3. Fuzzy Logic Model

For finding least cost path, every spatial effective criterion classified and defined their membership values between 0 and1.

There are some fuzzy operators such as fuzzy AND, fuzzy OR, fuzzy product, fuzzy sum and fuzzy γ used for factor maps integration.

2.2.3.1. Fuzzy AND

Fuzzy AND operates like common statements in classic sets theory as Eq. 8 [16]:

$$\mu_{Combinatio\ n} = Min(\mu_A, \mu_B, \mu_{C}, \ldots) \qquad (8)$$

Fuzzy AND operation extracts minimum value of the same location in output map and creating conservative estimate map from effective criteria. This is used when there have been two or more factors or evidences that can help solving the problem. Due to its weakness on participating all effective factors and lack of specific evidences for oil and gas route finding, fuzzy AND operator has not been used in this research [15].

2.2.3.2. Fuzzy OR

Fuzzy OR operates like common statements in classic sets theory as Eq. 9 [15]:

$$\mu_{Combinatio\ n} = Max(\mu_A, \mu_B, \mu_{C}, \ldots) \qquad (9)$$

Fuzzy OR operation extracts maximum value of the same location in output map creating optimistic map from effective criteria. Fuzzy OR operator is used when there are sufficient positive factors and evidences in study area. In our evaluation, because of lack of positive factors for pipelining, fuzzy OR operator has not been used [15].

2.2.3.3. Fuzzy product

Fuzzy product operator multiples input factor map membership values and presents the results on output map (Eq. 10). Therefore, it has decreasing affects on results and is used when input factor maps debilitate each others [16] [8]:

$$\mu_{Combination} = \prod_{i=1}^{n} \mu_i \qquad (10)$$

where $\mu_{Combination}$ denotes each unit value in output map and μ_i denotes the weight of ith factor map.

In oil and gas route finding, there were no negative factor maps, therefore, fuzzy product has not been used in our evaluation [8].

2.2.3.4. Fuzzy Sum

Fuzzy Sum operation is complementary of Fuzzy product operation (Eq. 11) [11]. Using Fuzzy Sum operation the values of membership tend to 1 so have increasing effects on each other. In our application, because of increasing effects of accessing factors, this operator was used and compared with other integrating operators [17].

$$\mu_{Combination} = 1 - \prod_{i=1}^{n} (1 - \mu_i) \qquad (11)$$

where $\mu_{Combination}$ denotes each unit value in output map and μ_i denotes the weight of ith factor map. This operator is used when input factor maps have increasing effects on each other. Because of increasing effects of accessing factors, this operator was used and compared with other integrating operations [17].

2.2.3.5. Fuzzy γ

Fuzzy γ is created from multiplication of fuzzy product and fuzzy sum operations (Eq. 12) [3]:

$$\mu_{\text{Combintion}}=(\text{Fuzzy Sum})^{\delta} * (Fuzzy\ Product)^{1-\delta} \qquad (12)$$

In Eq. 12, the value of gamma is determined between 0 and 1. Determining correct value of gamma generates output map showing adoption between decreasing and increasing trend in sum and product fuzzy operations [3].

2.3. Least cost path theory

Finding a minimum path over a surface partitioned into regions of different frictions to movement has two aspects: (i).Creation of an accumulated cost surface from a cost surface (where the frictions are stored); and (ii). Tracing a minimum cost on the accumulated cost surface from a source point to a destination [14].

2.3.1. Creation of accumulated cost surface

There are three main network problem models, related to the ways that weights or resistances are assigned to each of the network link including isotropic, partially anisotropic, and fully anisotropic [4,5].

2.3.1.1. Isotropic Model

In isotropic model costs are related to the node location, however, are the same for all directions (Fig. 1). Although many problems of route finding can be solved by the procedure described, it is limited to situations where the cost of passage or friction is the same for all directions of movement [4].

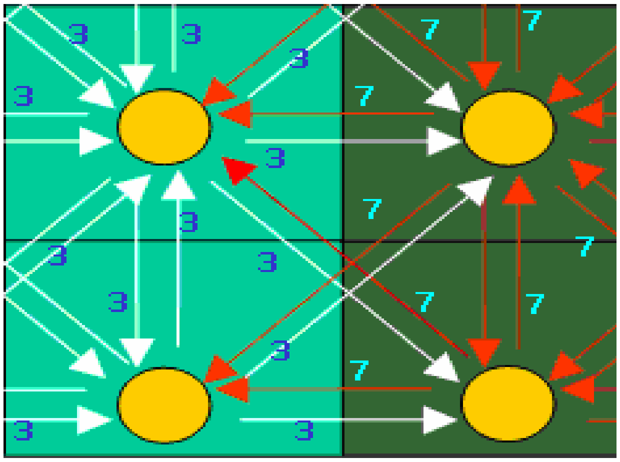

Fig. 1 Isotropic least cost model: the cost of passage depends on location only

2.3.1.2. Partially Anisotropic Model

In partially anisotropic model costs may be related to direction, however, this anisotropy has one prevailing direction and is valid for the whole surface (Fig. 2) [6].

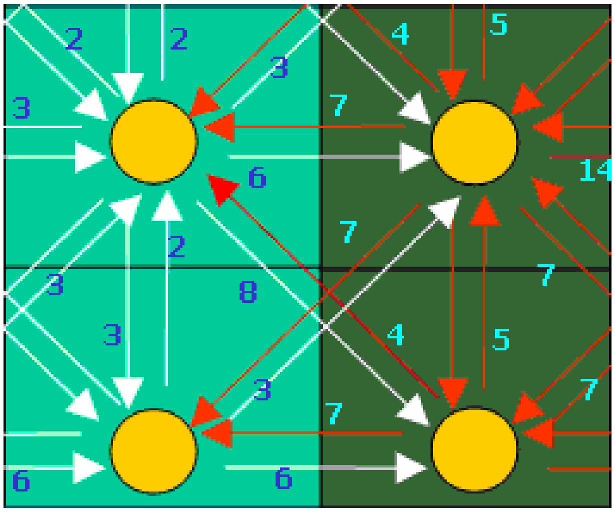

Fig. 2: Partially anisotropic least cost path model

Calculating partially anisotropic model divided to the two follow aspect [6]:

1. Calculate accumulated cost surface to move to the north, south, west and east directions by Eq. (13) and (14) [5,10]:

$$Accum-Cost = Link + (Cost2 + Cost3)/2 \qquad (13)$$

2. Calculate accumulated cost surface to moving to diagonal directions by Eq. 15 and 16 [14]:

$$Link = (Cost\,1 + Cost\,2)/2 \qquad (14)$$

$$Link = 1.414214 * (Cost\,1 + Cost\,2)/2 \qquad (15)$$

$$Accum-Cost = Link + 1.414214 * (Cost2 + Cost3)/2 \qquad (16)$$

where Cost 1 denotes the cost of cell 1, Cost 2 denotes the cost of cell 2, Cost 3 denotes the cost of cell 3, Link denoted moving cell 1 to cell 2 and Accum-Cost denotes accumulated cost.

2.3.1.3. Full Anisotropic Model

In full anisotropic model costs of passage are really anisotropic (Fig. 3). The cost of passage depends on location and direction, but there is one prevailing direction for the whole surface (for example it is easier to go left and up). In this type of problem, costs are variable over the surface and are direction dependent, however, the direction dependency is not in prevailing direction, and is also variable over the surface [4].

In planning routes for oil and gas pipeline design, the confirmatory pump position depends on slope, so topography acts as significant constraint. The case of pipeline, uphill reaches of the path should be restricted to a minimum and slope depends on the direction of the route. So oil and gas pipeline route planning depends on the position and direction (Fig. 3) [4].

Since the direction of travel for each cell is not defined at the beginning of the best route finding process, the cost-of-passage surface is also not defined, and the traditional cost-path finding procedure cannot be applied to solve problems with direction dependent costs. An alternative algorithm is presented in this paper, which can be used to draw the best path for a pipeline over a digital elevation model (DEM). The necessary costs to find the best path are related to the direction dependent slope through a pre-defined function.

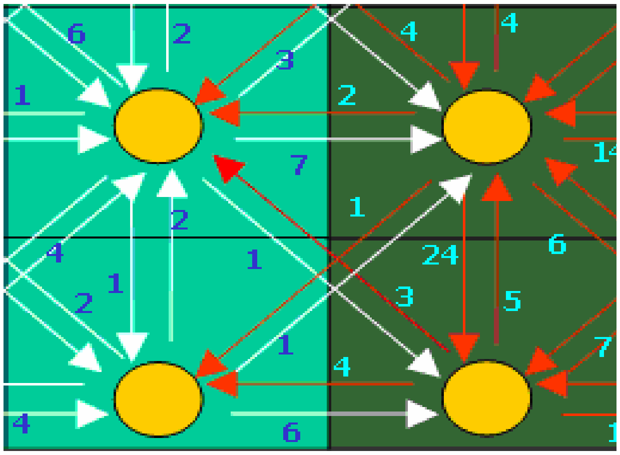

Fig. 3: Full anisotropic least cost path model

2.3.2. Path Finding

For finding least cost path between two points in raster environment, firstly source (starting point) and cost surface (created by gamma fuzzy overlay model) to full anisotropic model is introduced [13]. Full anisotropic model generated accumulative cost surface and direction surface that assigns number to each cell showing the direction of movement to calculate accumulated cost surface. Introducing destination point (ending point), accumulative cost surface and direction surface to path finding algorithm is generated. This algorithm follow minimum cost in every cell from destination point to reach the start point and finding least cost path between source point and destination point [11].

3. Experimental Consideration

In the case study, for calculating the weight of factors, the pipeline between Bidboland (in Chahar Mahale Bakhtiari Province) and Farah (in Kohkoloyeh and Boyer Ahmad Province) in the central part of Iran with 25 kilometer length has been selected. In this pipeline most of the influencing factors for pipelining have been considered, so it is a suitable route for estimating the weight of factors. In order to confirm the validity of the proposed weighting models, knowledge and data driven models have been jointly used to weight effective criteria. The estimated results by these models are shown in Table 1.

Table 1: The estimated weight for effective criteria

Criteria Name	Weight (%)
Elevation	25
Accessibility Roads	20
Environmental	10
Ground Cover	18
Natural Factors	17
Geology	10

Index overlay, fuzzy sum and fuzzy gamma models investigated as overlay models to create cost surface and the correlation and real cost of every model have been compared with the existing route (Table 2). The results show fuzzy gamma model is nearest model to reality for oil and gas pipelining, so fuzzy gamma model has been selected for overlay operation.

Table 2: Estimate of Cost

Model Cost	Index Overlay	Fuzzy Sum	Fuzzy Gamma	Exist Pipeline
Calculated cost (US $)	770000	751660	742551	744000

The value of gamma has been calculated and the results show gamma with 0.67 has the highest correlation with the existing pipeline shown in Fig. 4.

Full anisotropic models has been chosen as optimum model for generating accumulative cost surface

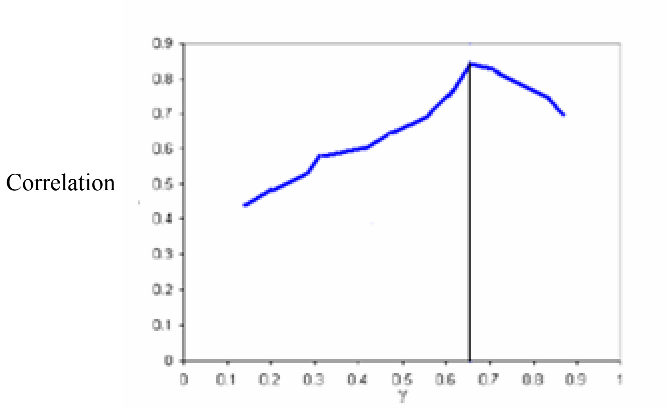

Fig. 4: The value of gamma correlation with the existing pipeline

Since the weighted parameters and the value of gamma are calculated based on the existing route pipeline (Bidboland-Farah), we tested the achieved value of gamma and weighted parameters on Maron-Ahvaz pipeline in the south of Iran and designed a pipeline route that is 31% cheaper than the existing pipeline route. This result shows, the calculated gamma value and weighted parameters are appropriate values for oil and gas pipeline design.

4. Conclusion

This paper outlines how to use statistical and fuzzy logic concept to achieve important criteria for pipelining and generating real cost surface.

The results show oil and gas pipelining is a nonlinear procedure, according to this result fuzzy model is a suitable model to generate the cost surface. Fuzzy gamma model is the best model to generated cost surface because oil and gas pipelining has increasing and decreasing trends.

There are some blind spots in knowledge and data driven models for weighing the effective criteria using combination of two models to estimate actual weights.

Pipelining in raster environment depends on pixel size.

In GIS the accuracy of models strongly depends on the quality of primary data, functions and type of analysis very selecting optimum methods for data generation is so important in analysis.

The achieved weights and parameters tested on Maron-Ahvaz gas route and the result shows the designed route pipeline is 31% cheaper than the existing route pipeline, indicating that the achieved parameters seems suitable for oil and gas route finding.

References

[1]. An, P., 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.

[2]. Aronof, S., 1989. Geographic Information Systems: A Management Perspective, Ottawa, Canada: WDL Publications.

[3]. Bonham Carter, G.F., 1991. Geographic Information System for Geoscientists: Modeling with GIS, Pergamon, Ontario.

[4]. Collischonn, W. and Pilar, J.V., 2000, A direction dependent least-cost-path algorithm for roads and canals. International Journal of Geographic Information Systems, 14, Vol.4, pp. 491-508.

[5].Delavar, M., H. Kiavarz, 2004," Network analysis in raster with emphasis on oil and gas pipeline", Proc. Map Asia 2004, Beijing, China, Aug, 26-29, 2004, 12 p

[6].Delavar, M., H. Kiavarz, 2005, "Route finding with least cost path algorithm for oil and gas pipeline, Proc. GIS Planet 2005, Estoril, Portugal , May 30 – June 2, 2005 , pp. 1-13.

[7]. Demers, M.N., 2005. GIS Modeling in Raster, Second Edition, Wiley.

[8]. Frank, A., 2005, "The Mathematics For Geographic Information System", Unpublished Manuscript.

[9]. Goodchild, M.F., 1977, An evaluation of lattice solutions to the problem of corridor location. Environment and Planning, 727-738.

[10] - Husdal, J., 1999, "Network analysis-raster Vs. Vector", A comparative study, University of Leicester, http//www.husdal.com

[11]. Kiavarz, H., 2005, "Optimum Gas Pipeline Routing Using GIS ", MSc. Thesis (In Persian with English abstract), Eng. Faculty, University of Tehran, Iran, 98p.

[12]. Kiavarz, H, M.R., Delavar, 2006, "Generation of cost surface for gas pipeline route finding ", accepted to be published in Proc. 25th. UDMS Conference, May 15-17, 2006, Alborg, Denmark, 9 p.

[13] Lee, J.; Stucky, D. 1998, On applying viewshed analysis for determining least-cost paths on Digital Elevation Models. International Journal of Geographical Information Science.Vol. 12, No. 8, pp.891-905.

[14]. Worboys M.F. and M.duckham, 2004, GIS: a Computing Perspective Second Edition. London, Taylor and Francis.

[15]. Zadeh, L.A., 1986. Probability measures of fuzzy events. Math. Anal. Applications, Springer Netherlands

[16]. Zadeh, L.A., 1988. Fuzzy logic, IEEE Computer, pp. 83-93.

[17]. Zimmermann, H.J. and Zayo, P., 1980. Latent Connectives in Human Decision Making: Fuzzy Sets and Systems, Chapter 4, pp. 37-51.

Hamid Kiavarz Moghaddam received the B.E from surveying and Geomathics Eng.Facalty, University of Iran University Science & Technology in 2002 and 2003 and M.E degree from Geo Spatial Information System Eng.Facuilty University of Tehran University in 2005 and 2006, next working as a research assistant in Tehran University and Tehran GIS Center depended on Municipality.

Mahmoud Reza Delavar received the BSc. degree in Civil Eng.- Surveying from K.N.Toosi Univ. Iran, in 1989, M.E. in Civil Eng. – Photogrammetry and Remote Sensing from Univ. of Roorkee, India, and a PhD in Geomatic Eng.-GIS from Univ. of New South Wales, Australia in 1997. Since 1998 he has been working in Dept. of Surveying Eng., Eng. Faculty, University of Tehran as an assistant professor. His research interest includes GIS-based artificial intelligence, agent-based spatial modeling, spatial data quality, spatio-temporal GIS, land administration, spatial data infrastructure, disaster management and cadastre. He is president of Iranian Society of Surveying and Geomatic Eng. He is also national representative of Iran in UDMS.