

MODELING TRANSPORTATION SUPPLY AND DEMAND FORECASTING USING ARTIFICIAL INTELLIGENCE PARAMETERS (BAYESIAN MODEL)

Mehrdad Arabi^{1*} Mohammad Ali Dehshir² Mahdi Shokrgozar³

¹Isfahan University of technology, Isfahan, Iran

²Amirkabir University of Technology, Tehran, Iran

³Iran University of Science and Technology, Tehran, Iran

The growth of population has led to an increase in the number of motor vehicles. This increase has caused traffic congestions in metropolitan cities and higher fuel consumption besides other problems. A solution for dealing with the increase in the number of vehicles is creating more capacity for roads, parking lots, etc. thus aiming at controlling the traffic problem. Providing proper approaches is on the condition that growth rate of vehicles in future years could be estimated by logical methods and so a relation between supply and demand at any time can be made. Considering this problem, in this paper one of the artificial intelligence parameters that is "Bayesian model" is introduced and by providing a model using this parameter, transportation situation (the ratio of capacity to demand) in future years can be forecasted. According to the results of this study, a model for supply and demand is suggested that can be used for traffic design of different regions and sections. Additionally the necessary capacity for roads and other important parameters in transportation is estimated according to the estimated capacity.

Key words: Artificial intelligence, Forecasting, Vehicles, Capacity, Transportation planning, Supply and demand, Bayesian model

INTRODUCTION

Trip generation and distribution are most important economic reactions that display the necessity of demand for transportation. Providing such demand creates transportation networks and also major transportation lines required between supply and demand locations. However demand for transportation is basically not separated from transportation supply, due to the fact that both generation and distribution and their interaction are highly related to each other through structure and efficiency of transportation networks. In fact, transportation networks are specific physical factors which are formed in and around major economic centers and so their main points of development are the points inside and around major economic centers [01].

TRANSPORTATION SUPPLY

Transportation supply means providing specific services which are made in response to transportation demand. The supply is provided by:

- Government organizations (at a local, regional and national level)
- Private organizations or companies (at a professional level)
- Businessmen and industry owners who have major benefits in providing facilities as well as owners of private cars (at an individual level) [02].

This supply includes: providing terminal buildings, installations, provision of vehicle, operations adjustment, and supervising the implementation of these services. In some industrial countries, construction and maintenance of roads is part of the services in this section. Preparing such operations and extra services and providing technical facilities is not possible to be supported by one supplier or individual organization unless supply facilities are managed according to technical and economical properties in the form of specific organized units.

Transportation supply may be offered in different levels of facilities and the difference in the quality and type of facilities increases the power of suppliers and causes a change in economic structure of transportation market [03].

Impossibility to preserve supply is the general characteristic in transportation economy and in many facilities, unused sections exist such as empty spaces in terminals, unnecessary staff, vehicles and etc.

Indivisibility of transportation supply is one of the factors in formation of oversupply because if we assume that vehicles have constant capacity and demand remains constant over time, it would be difficult to provide the actual combination of capacity on transportation facilities without errors.

Supply and demand in transportation services are often unbalanced and most of the times there is oversupply, thus, time is another cause of oversupply which requires the transportation suppliers to manage the operation based on a time table [04][06].

Technological developments may also lead to oversupply by causing difficulties in capacity fulfillment.

Different types of new transportation modes may create abstraction demand and the overcapacity can lead to an intense competition in transportation market in some cases. Experience has shown that by providing new facilities, demand is usually induced while it is possible that no traffic is attracted [05]

Traffic congestion, pollution, road development costs, lack of parking spaces, increasing dependency on cars and many other problems make most of the authorities to strive for a solution to overcome this situation. This paper aims at introducing the Bayesian model and investigating the problem of a specific type of management called transportation demand management by forecasting supply and demand for future years.

Before defining the Bayesian model, models of supply and demand are introduced.

Model of demand

The aim of this model is to find the optimal amount of traffic in a way that the total costs of transportation are minimized. In other words, the purpose is to obtain a demand situation in which the total costs of transportation are minimized. Therefore, the objective function aims at minimizing the costs and is defined as follows:

$$\text{MIN } Z = \sum_{i=1}^n C_i X_i \quad 1)$$

Where,

X_i is number of the vehicles entering the system;

and C_i is the costs of the vehicles entering the system.

Model of supply

The purpose of supply model is to obtain the optimal values of transportation supply, so that the benefits of vehicles' transportation as a whole is maximized. Therefore, the issue here is in contrary to the transportation demand model (i.e. the model is defined as a maximization model). However, both models follow the same general methodology and their number of activities and decision variables are the same. Supply demand is defined as the following equation

$$\text{MAX } H = \sum_{i=1}^n (C_i - P_i) X_i \quad 2)$$

Where,

X_i is number of the vehicles entering the system;

C_i is the transportation costs of the vehicles entering the system;

P_i is the transportation cost of the vehicles entering the system for users.

BAYESIAN MODEL

Bayesian models which are also known as belief networks, belong to the family of probabilistic graphical models. These graphical structures are used to show data in a field with uncertainty. More specifically each node in the graph shows a random variable and arcs show probabilistic dependency among variables. Such conditional dependencies are often evaluated by specific statistical and probabilistic methods. Bayesian models incorporate principles of graph theory, probability theory, computer and statistics sciences [07][08].

Bayesian models belong to a structure of graphical models named directed acyclic graphs which are popular in statistics, machine learning and artificial intelligence. Bayesian models provide effective representation and calculations of common probability distributions on a series of random variable. Moreover, Bayesian models quantify the rate of correlation between variables which allows the automatic update of convictional beliefs once new data are available [09].

Bayesian models have become popular models in the last decade. These models have many applications in various fields such as quality control of pavement, transportation situation control, error control codes, medical diagnosis, weather forecasting, and cellular networks.

In a Bayesian model, if there is a set of v variables including $\{x_1, \dots, x_v\}$, while each x_i is a random variable and parent nodes of this variable are shown by $\text{Parent}(x_i)$, the occurrence probability of $X=(x_1, \dots, x_v)$ is calculated through $P(X)=\prod_{i=1}^v P(x_i | \text{Parent}(x_i))$ [10][11].

PROBLEM STATEMENT

In this paper, in order to forecast the transportation situation (supply and demand) in 10 years, two methods are used to model transportation situation in consecutive years. One method is judgment data and the other is using the actual supply and demand data in 10 years (The data and experts are for 20 sections of Tehran city).

The number allocated to the transportation situation in the best case is 100 and in the worst case is 0, while it decreases from 100 to 0 over time under stable conditions.

In this study, 140 experts are asked about transportation situation and 5 cases which had the highest frequency are selected. These cases include the opinions that are repeated frequently among the experts. These data and the actual data are presented in Tables 1 and 2.

Table 1 shows the expert judgment about transportation situation in 10 years in the city of Tehran and Table 2 provides the actual data about transportation situation in this time period and in 20 sections of Tehran. The system ages in these tables are the years after which the system changes after designing (for example a system is able to support the traffic over the first years after the design and as time goes by, with the increase in traffic it is not able to provide an appropriate performance to the existing

traffic). The transportation situation in these tables presents the ratio of supply to demand.

Table 1: Expert judgment about transportation situation in 10 years

System Age	Transportation situation (supply and demand)				
	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5
1	90	91	89	89	89
2	88	88	85	86	85
3	85	88	90	92	95
4	80	81	81	79	77
5	73	76	75	78	76
6	62	59	59	60	61
7	51	54	55	56	59
8	46	48	47	45	44
9	37	37	37	35	36
10	30	30	27	24	23

Table 2: Actual data of transportation situation in 10 years

Section #	Age	Transportation situation (supply and demand)
1	2	87
2	9	28.5
3	6	51
4	8	45
5	6	52
6	2	83
7	6	56
8	1	88.5
9	6	58
10	10	29
11	7	51.5
12	10	24
13	7	47.5
14	3	75.5
15	3	81.5
16	10	22
17	1	97.5
18	5	62.5
19	8	40
20	10	24

The sections in Table 2 are the selected sections of the system for analyzing supply and demand, and each one had different ages at the time of measuring supply and demand (for example, the system age 2 means that 2 years has passed from the planning time in section 1 and the ratio of supply to demand is 87). The 20 selected sections are for 20 different points and are not located in connected paths.

According to the field data and expert judgment data, an equation must be provided to be able to anticipate the transportation situation in different years. Based on such information, the only independent parameter in this equation is Age. Transportation situation (supply and demand) is displayed by A.

$$A = b_0 + b_1 Age \tag{3}$$

So the target is to find b0 and b1 coefficients for Bayesian model (The final coefficients of bpos).

To this end, prior coefficients (bprior) and new coefficients (b) must be calculated at first.

Step 1: Calculation of prior coefficient (bprior):

1. In order to calculate the prior coefficient, expert judgment data are used. Thus, first the average value of experts' judgments is obtained and the average value of each group is considered as the representative of that group. Table 3 shows the average values of expert judgment data (the 5 cases which had the highest frequency).

Table 3: Average values of the expert judgment data

System Age	Y=A _{Average}
1	89.6
2	86.4
3	90
4	79.6
5	75.6
6	60.2
7	55
8	46
9	36.4
10	26.8

2. By using the following equation the bprior coefficient is calculated.

$$P(1) = P(0) \times TPM_1 \tag{4}$$

$$b_{pr} = A^{-1} \times (X_G^t \times Y)$$

Where,

P(1) is the probability in year 1,

P(0) is the probability in the beginning year,

TPM is total productive maintenance (the enhancer matrix of supply and demand situation in the n_{th} year), X_G is the independent variable that is age, Y is the dependent variable matrix (transportation situation), and b_{pr} is the prior coefficient.

It must be noted that in order to obtain X_G matrix, a column with elements of 1 is added to this index to calculate b_0 coefficient. Table 4 shows matrix of year and its transpose, Table 5 presents matrix A and its inverse and the value of the prior coefficient is shown in Table 6.

Table 4: Matrix of year and the transpose of matrix of year (for judgment data)

X_G											
1	1										
1	2										
1	3										
1	4										
1	5										
1	6										
1	7										
1	8										
1	9										
1	10										

Table 5: Matrix A and its inverse

$A = g(X_G^t * X_G)$		A^{-1}	
10	55	0.467	-0.067
55	385	-0.067	0.012

Table 6: Value of the prior coefficient

$b_{pr} = A^{-1} * (X_G^t * Y)$	
b_0	105.373
b_1	-7.421

The value of g used in table5 is G-prior factor and is generally considered 1.

Step 2: Calculation of new coefficient (b):

1. To calculate this coefficient, actual data are used.
2. The value of b is calculated using following equation.

$$H = X^t \times X \longrightarrow b = H^{-1} \times (X^t \times Y) \quad 5)$$

Where,

X is the matrix of independent variable data (time) and Y is the matrix of dependent variable (transportation situation).

Table 7 shows matrix of year and its transpose for actual data and Table 8 provides the values of Matrix H and its inverse.

Table 8: Matrix H and its inverse

$H = X^t * X$		H^{-1}	
20	120	0.246	-0.033
120	904	-0.033	0.005

3. Finally the values of coefficient b is shown in table9.

Table 9: Values of new coefficient (b)

$b = H^{-1} * (X^t * Y)$	
b_0	100.624
b_1	-7.571

Step 3: calculation of final coefficient (b_{pos})

1. In order to calculate the final coefficient of Bayesian model, the combination of prior and new coefficients is used.

$$M = H + A \quad b_{pos} = M^{-1} \times (A \times b_{pr} + H \times b) \quad 6)$$

Where A and H are matrixes obtained through multiplying the matrix of independent variable data by the transpose of this matrix. Table 10 shows the combination of prior and new coefficients and inverse of matrix M.

Table 10: Combination of prior and new coefficients and inverse of matrix M

$A * b_{pr} + H * b$	M^{-1}	
1749.6	0.160	-0.022
8169.6	-0.022	0.004

2. Finally the value of the model final coefficient can be obtained which is provided in Table 11.

Table 11: Value of the model ultimate coefficient

$b_{pos} = M^{-1} * (A * b_{pr} + H * b)$	
b_0	102.617
b_1	-7.594

Table 7: Matrix of year and transpose of matrix of year (for actual data)

X																			
1	2																		
1	9																		
1	6																		
1	8																		
1	6																		
1	2																		
1	6																		
1	1																		
1	6	X	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	10		2	9	6	8	6	2	6	1	6	10	7	10	7	3	3	10	1
1	7																		
1	10																		
1	7																		
1	3																		
1	3																		
1	10																		
1	1																		
1	5																		
1	8																		
1	10																		

Step 4: Modeling

Eventually, the Bayesian model which can be provided based on actual data is as following:

$$\begin{aligned}
 Y = b_0 + b_1X & \begin{cases} A = 105.373 - 7.421Age & \text{for } b_{nr} & 7) \\ A = 100.624 - 7.571Age & \text{for } b & 8) \\ A = 102.617 - 7.593Age & \text{for } b_{nos} & 9) \end{cases}
 \end{aligned}$$

By inputting the values of Age into each model, the values of the new A can be calculated and by computing the difference between the calculated A and the initial values, the value of error for each A is calculated. Finally, by squaring all error values and summing them, the sum of square error is obtained. Table 12 shows these results.

$$A_{new} = Age * b \quad e = A_{new} - A_{old} \quad 10)$$

RESULTS AND CONCLUSION

1. By considering the increasing growth of population and number of vehicles that lead to serious transportation problems and issues including environmental pollutions, reduced sources of energy, increased material and spiritual damages, increased delay and increasing growth of transportation demand particularly in the peak hour, paying special attention to adopting strategic policies by traffic authorities and urban planners against such important problem is very essential.
2. By adopting appropriate methods it is possible to address this problem, however, adopting such methods depends on planning and estimating the future situation and in this paper the Bayesian model was provided to estimate the future situation.
3. In this paper, the relation of supply and demand of transportation in different years is presented. Initially prior, new and final coefficients are calculated and a model for supply and demand for applying to the traffic design of different regions and sections is suggested. After using the Bayesian model to model the transportation situation in future years, various and numerous methods must be evaluated as travel demand reduction solutions to determine which method is applicable in the country.

4. Finally, by calculating the errors, the validation from the experts' judgment is performed. It was observed that the error due to bpr is more than the error due to b. This phenomenon is reasonable because the required data for calculation of bpr are based on the expert judgment and therefore they have more error compared to the actual data required for calculation of b. In other words, while there are error in both methods, the method in which the data are taken from the experts has more error because an expert gives the numbers based on his experience and has not tested and evaluated any data. But, the data provided for calculation of b are obtained based on the

experiments and previous data. However, the error due to calculation of bpos is higher because it is the cumulative errors from bpr and b, but its value is less than the sum of bpr and b.

5. According to the results of this study, for obtaining a more accurate and better estimation (situation of transportation or other related parameters in future years), it is possible to define an adjustment factor for the equation based on the experts judgment or a safety factor for the judgments of experts which can be investigated in future studies.

Table 12: Error values based on judgment data and actual data

Y_1	e_1	Y_2	e_2		
97.95	8.35	85.48	-1.52		
90.53	4.13	32.49	3.99		
83.11	-6.89	55.20	4.20		
75.69	-3.91	40.06	-4.94		
68.27	-7.33	55.20	3.20		
60.85	0.65	85.48	2.48		
53.43	-1.57	55.20	-0.80		
46.01	0.01	93.05	4.55		
38.59	2.19	55.20	-2.80		
31.17	4.37	24.92	-4.08		
Sum of square of Error = 230.05 Based on judgment data		47.63	-3.87		
		24.92	0.92		
		47.63	0.13		
		77.91	2.42		
		77.91	-3.59		
		24.92	2.92		
		93.05	-4.45		
		62.77	0.27		
		40.06	0.06		
		24.92	0.92		
		Sum of square of Error =186.28 Based on actual data			

ACKNOWLEDGEMENT

The authors contributed to this work equally.

REFERENCES

1. Albalate, D., Bel, G., (2009). Factors explaining urban transport supply and demand in large European cities. Working papers of the Research Institute on Applied Economics, Universitat de Barcelona. JEL codes: L91; L98; R41.
2. Allahverdizadeh, P., (2003). Methods of travel demand management and transportation system management, Secretariat of the Higher Council for Traffic Coordination in cities, Iran.
3. Allahverdizadeh, P., (2003). The regulations for travel demand reduction and transportation system management, Secretariat of the Higher Council for Traffic Coordination in cities, Iran.
4. Al-Saba, T., El-Amin, I., (1999). Artificial neural networks as applied to long-term demand forecasting. Artificial Intelligence in engineering, Volume 13, Issue 2, Pages 189-197.
5. Khadaroo, J., Seetenah, B., (2008). The role of transport infrastructure in international tourism development: a gravity model approach. Tourism Management, Volume 29, Issue 5, Pages 831-84.
6. Matas, A., (2007). Demand and revenue implications of an integrated public transport policy: the case of Madrid. Journal of Transport Reviews, Pages 195-217.
7. Heckerman, D., (1997), Bayesian Networks for Data Mining. Data Mining and Knowledge Discovery, Volume 1, Issue 1, pp 79–119, Kluwer Academic Publishers.
8. C-SHRP594 Canadian Strategic Highway Research Program, (1995). Bayesian Modeling: Joint C-SHRP/ Agency Application. Technical Brief No. 8, Ottawa. ISBN: 1-55187-024-X.
9. B. Korb, K., E. Nicholson, A., (2004), Bayesian Artificial Intelligence. Chapman & Hall/CRC, A CRC Press Company. ISBN: 1-58488-387-1.
10. Grossman, D., Domingos, P., (2004). Learning Bayesian Network Classifiers by Maximizing Conditional Likelihood. Proceedings of the Twenty-First International Conference on Machine Learning, ICML, Page 46. Doi:10.1145/1015330.1015339.
11. Neufville, R., Clark J., and Field F., (2013). Introduction to Linear Programming, Engineering Systems Analysis for Design. Massachusetts Institute of Technology.

Paper submitted: 12.07.2017.

Paper accepted: 15.11.2017.

This is an open access article distributed under the CC BY-NC-ND 4.0 terms and conditions.