JOURNAL OF TRANSPORTATION
SYSTEMS ENGINEERING AND INFORMATION TECHNOLOGY
Volume 11, Issue 2, April 2011
Online English edition of the Chinese language journal

Available online at www.sciencedirect.com
ScienceDirect

Cite this article as: J Transpn Sys Eng & IT, 2011, 11(2), 84−90.

RESEARCH PAPER

Node Importance Evaluation Method for Highway Network of Urban Agglomeration

SONG Xinsheng[1,2], WANG Xiaoxiao[1,2], LI aizeng[1,2] , ZHANG Lei[1,2]

1 Department of Traffic Engineering, Henan University of Urban Construction, Pingdingshan 467036, Henan, China
2 Urban Transportation Research Center, Henan University of Urban Construction, Pingdingshan 467036, Henan, China

Abstract: Urban agglomeration is a strong vitality subsystem of large regional city system with significant regional advantages in space and linked features in network. Objective and accurate assessment of each node importance of highway network is a vital part for regional road network layout planning. Problems in the existing node important evaluation methods are analyzed. And for the characteristics of urban agglomeration, urban flow intensity is included into node importance assessment indexes system. Factor analysis is used as an objective method to avoid random subjective values in node importance calculation. K-Means clustering method is used to distinguish degree level of each node importance for further analysis. Finally, central China city agglomeration is taken as an application example, and the result shows the method proposed in this paper has good practical value in application.

Key Words: traffic engineering; highway transportation; node Importance evaluation; factor analysis; urban agglomeration highway network; urban flow intensity; cluster

1 Introduction

Urban agglomeration consists of different levels of cities, generally with one or two economically developed central city of strong radiation drive function. In urban agglomeration, surrounding cities and hinterland form an urban system by their close economic contact, functional complementation and order grade[1]. Whether in regional level or on correlation space, urban agglomeration takes regional networked organization as the tie, and has basic characteristics of network[2]. Urban agglomeration node is the main source of traffic demand service, service object of traffic network, and also indispensable system unit of urban agglomeration traffic network. So node importance study can provide the basis for road network layout planning of urban agglomeration.

The existing studies on node importance evaluation mainly focus on three aspects. First, in assessment indexes selecting, Hu L G, et al.[2–5] have used population, GDP and other socio-economic assessment indexes; Zhou W, et al.[6-7] have used transportation assessment indexes such as passenger and freight volume, highway net density and mileage. However, as urban agglomeration has regional connectivity and open characteristics and cities within the urban agglomeration have mutual attraction concentration and diffusion radiation function, the assessment indexes reflecting urban agglomeration structural characteristics should be adopted. Second, in quantitative calculation process, the subjective method is usually used for index weights calculation, which can better reflect subjective judgments or experience of decision makers with simple implementation process. However, the subjective values determined with this method are usually at random and influenced by knowledge which lacks decision makers. So researchers have turned to use objective weighting method; for example, [3] and [5] used principal component analysis method to determine weights. Third, after node importance calculation, the results do not reflect the division of the node boundaries. In order to further distinguish the degree of each importance node level better, [2] and [7] have divided the results into several categories using system cluster. In view of the above analysis, urban flow intensity is included into node importance assessment indexes system, factor analysis is used as an objective method to determine weights value and K-Means clustering method is used to distinguish the node importance level in this paper.

2 Node importance

Received date: Nov 09, 2010; Revised date: Feb 21, 2011; Accepted date: Feb 25, 2011
*Corresponding author. E-mail: sxs@hncj.edu.cn

DOI: 10.1016/S1570-6672(10)60115-8

2.1 Concept of node importance

In planning study, for the ease of performance, visual descriptions and computer storage, abstract network chart is often used to simulate the actual highway network. The region studied is divided into several sub-regions. And each sub-region is abstracted as a point which is called node. Index reflecting the size of node function and status is called node importance [3]. It is usually calculated as follows:

$$\begin{aligned} D_k &= (a_{k1}\frac{X_{k1}}{X_{10}} + a_{k2}\frac{X_{k2}}{X_{20}} + \ldots + a_{kn}\frac{X_{kn}}{X_{n0}}) \times 100\% \\ &= \sum_{i=1}^{n} a_{ki} \cdot \frac{X_{ki}}{X_{i0}} \times 100\% \end{aligned} \tag{1}$$

Where, D_k is the node importance of the k^{th} node; X_{ki} is the i^{th} evaluation index value of D_k; X_{i0} is the average value of X_{ki} of all nodes; α_{ki} is the weights value of X_{ki}.

The premise of node importance calculation is to select evaluation index X. As node importance is influenced by many factors such as policy, economic, culture and so on, indexes should be comprehensively considered and generally selected from the following:

① Population in region X_1;
② Region GDP X_2;
③ Average income of urban resident family X_3;
④ Total mileage of highway X_4;
⑤ Passenger volume X_5;
⑥ Freight volume X_6;
⑦ Tourism income X_7; et al.

2.2 Node importance evaluation index of urban agglomeration

Node important evaluation indexes of urban agglomeration not only reflect regional social economy and transportation development conditions, but also should reflect the degree of city external economic exchanges and agglomeration radiation effects. As one of the important basis of indexes reflecting the strength of urban agglomeration city function, planning and development, urban flow intensity should be included in the urban agglomeration node important indexes system.

The economic ties between cities are achieved in the form of flow by concentration and diffusion. Urban flow refers to two-way or multi-dimensional flow phenomenon of space flow between cities such as logistics flow, information flow, capital flow and technique flow in urban concentrated regions. Urban flow intensity means concentration and radiation energy of outward function and quantitative influence connection between city and city, and between city and countryside [8]. Urban flow is an intensity quantitative connection index of city with the outside, and its expression is $F = N \times E$. Where, F is the urban flow intensity; N is the urban function and benefit, which is the actual effect of urban outward function capacity; and E is the urban outward function capacity.

In practice, urban employee number is usually taken as a measure of urban function capacity. Whether a city has an outward function mainly depends on one of its departments staff location quotients Lq_{ij}.

$$Lq_{ij} = \frac{G_{ij}/G_i}{G_j/G} (i = 1,2,\cdots,n; j = 1,2,\cdots,m) \tag{2}$$

where G_{ij} is the urban employees number of j department in city i; G_i is the urban employees number in city i; G_j is the urban employees number of j department in China; G is the urban employees number in China. If $Lq_{ij} < 1$, $E_{ij} = 0$. If $Lq_{ij} > 1$, j department in city i has outward function. The outward function of j department in city i beyond national average is E.

$$E_{ij} = G_{ij} - G_i(G_j/G) = G_{ij} - G_{ij}/Lq_{ij} \tag{3}$$

The total outward function of m departments in city i is E_i.

$$E_i = \sum_{j=1}^{m} E_{ij} \quad (i = 1,2,\cdots n) \tag{4}$$

Function and benefit N_i in city i can be denoted by GDP of the average urban employees.

$$N_i = GDP_i / G_i \tag{5}$$

The urban flow intensity in city i is F_i.

$$\begin{aligned} F_i &= N_i \times E_i = (GDP_i/G_i) \cdot E_i \\ &= GDP_i \cdot (E_i/G_i) = GDP_i \cdot K_i \end{aligned} \tag{6}$$

where GDP_i is GDP in city i; K_i is the proportion of outward function capacity in city i to total outward function capacity, which is called urban flow tendency[8].

3 Node importance evaluation method

3.1 Factor analysis method model

Suppose there is a set of m observable random variables $x_1, x_2, \cdots, x_m$ related to P underlying common factors $F_1, F_2, \cdots F_p (m \ge p)$. Each variable x_i has specific factors $U_i (i = 1 \cdots m)$. The specific factors $U_i (i = 1 \cdots m)$ are independent and independent of $F_j (j = 1 \cdots p)$. Each variable x_i is a linear combination to P common factors and its specific factors U_i. The general expression of factor analysis mathematical model[9] is as follows:

$$\begin{cases} X_1 = a_{11}F_1 + a_{12}F_2 + \cdots a_{1p}F_p + c_1U_1 \\ X_2 = a_{21}F_1 + a_{22}F_2 + \cdots a_{2p}F_p + c_2U_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ X_{\mathrm{m}} = a_{m1}F_1 + a_{m2}F_2 + \cdots a_{mp}F_p + c_mU_m \end{cases} \tag{7}$$

The matrix expression instead of the above equations is $X_{m\times 1} = A_{m\times p}F_{p\times 1} + C_{m\times m}U_{m\times 1}$. There are several assumptions that are made with this model:

(1) $m \ge p$;

(2) $\mathrm{COV}(F,U) = 0$, which means F and U are

independent;

(3) $E(F)=0$, $\mathrm{COV}(F)=\begin{bmatrix}1 & & \\ & \ddots & \\ & & 1\end{bmatrix}_{P\times P}=I_P$,

which means $F_1,F_2,\cdots F_p$ are independent with variance 1 and mean 0;

(4) $E(U)=0$, $\mathrm{COV}(U)=I_m$, which means $U_1,U_2,\cdots U_m$ are independent and in standardized form. Suppose $X_1,X_2,\cdots X_m$ are in standardized form but not independent.

Where, coefficient matrix A is the loading matrix. Its element (a_{ij}) size reflects the load of common factors to the observation variable X_i.

3.2 Node importance calculation

(1) Indexes standardization: As each index has a different order of magnitude and there is no unified metrics, the original data matrix X should be dimensionless or normalized to get a new matrix M.

(2) Coefficient matrix R calculation: Calculate eigenvalues λ and eigenvectors L of matrix R and get factors loading matrix.

(3) Determine the number of factors: Select the top t common factors, whose total variance contribution rate satisfies

$$\sum_{j=1}^{n} l_j > 85\%$$

Where, variance contribution rate is

$$l_j = \lambda_j / \sum_{i=1}^{n} \lambda_i$$

(λ_i is the i^{th} eigenvalue of original data matrix X).

(4) Factor rotation: Get the rotated factor loading matrix.

(5) Calculate factors score coefficient matrix using regression, Bartlett or Thomson estimation method. Consider every factor normalized value of variance contribution rate as weights. Linear combinations of factors score coefficient matrix and every factor weights are index weights α_{ki} of the k^{th} node importance.

(6) Node importance: Consider value α_{ki} into (1).

3.3 Node importance cluster analysis

The purpose of node importance cluster analysis is to differentiate node type, to determine control points of different levels of routes and to layout highway network better. K-Means clustering is a technology of dividing all samples into c groups and determining cluster center positions[10].

By K-Means clustering, vectors $X(t)=[x_1(t),x_2(t),\cdots x_n(t)]^T$ are divided into c groups $G_i\,(i=1,2\cdots c)$, cluster centers of each group are obtained and objective function minimum of non-similarity indexes are made. When Euclidean distance is selected as an index of vector x_k in group G_i and corresponding cluster center c_i, objective function can be defined as follows:

$$J=\sum_{i=1}^{c} J_i=\sum_{i=1}^{c}\left(\sum_{k,x_k\in G_i}\left\|x_k-c_i\right\|^2\right) \tag{8}$$

Groups divided can be defined as a $c\times n$ two-dimensional membership matrix U. If the j^{th} data point belongs to group i, element u_{ij} of U is 1; otherwise u_{ij} is 0. Once the cluster center c_i is determined, u_{ij} can be derived as Eq. (9) making Eq. (8) minimum:

$$u_{ij}=\begin{cases}1 & for\ k\neq i,\ if\left\|x_j-c_i\right\|^2\leq\left\|x_j-c_k\right\|^2\\ 0 & else\end{cases} \tag{9}$$

If c_i is the nearest cluster center of x_j, x_j belongs to the group j. The K-Means algorithm is as follows:

Step 1: Initialize cluster center $c_i\,(i=1,2,\cdots,c)$;

Step 2: Determine membership matrixes U according to Eq. (9);

Step 3: Calculate value function according to Eq. (8). If $J^l<\varepsilon$ or $\left\|J^l-J^{l-1}\right\|<\varepsilon$, the algorithm stops, else the algorithm continues (where, ε is a tolerance);

Step 4：Modify cluster center according to

$$c_i=\left(\sum_{j=1}^{n} u_{ij}\right)^{-1}\cdot\sum_{k,x_k\in G_i} x_k$$

return to step 2.

4 Application example

4.1 Description of central China urban agglomeration

The central China urban agglomeration is a close contact circle with a total of nine provincial-level cities and more than 14 county-level cities. It takes Zhengzhou as center and Luoyang as sub-center. The other cities of the agglomeration are Kaifeng, Xinxiang, Jiaozuo, Xuchang, Pingdingshan, Luohe and Jiyuan. With Longhai and Beijing-Guangzhou railway as the central axis, it forms a centralized and decentralized urban dense area. There are freeways of Beijing-Zhuhai, Huoerguosi-Lianyungang, et al. and national highway of G106, G107, et al. By 2009, the total mileage of highway has reached 87308 km, accounting for 36.3% of Henan; the total mileage of freeway has reached 2114 km, accounting for 43.7% of Henan; the highway net density (in land area) has reached 141.54 km/100 km^2 and the freeway net density has reached 3.43 km/100 km^2 in the urban agglomeration.

According to Central China urban agglomeration planning, “a half hour traffic circle” taking Zhengzhou as central radiation will have been developed until 2020. And an urban space layout of “two laps, dual-core, four belt, a triangle” and overall pattern of functioning will also be formed[11] as shown in Fig. 1. Land transportation is principal means of transportation in the agglomeration, and highway transportation plays an important role. In highway layout planning, global and local spatial layout of urban agglomeration should be considered and development should also be emphasized diversely in different stages. So node importance is exactly a significant parameter for spatial layout optimizing. However, only nine provincial-level cities without county-level cities are selected as nodes in the following studies.

Fig. 1 Spatial layout shape of central China urban agglomeration

4.2 Node importance analysis of central China urban agglomeration

Indexes shall be selected from macroscopic and microscopic aspects to reflect regional comprehensive strength. Considering different regional and economic characteristics of cities in Henan, eight indexes are selected into node importance evaluation system of central China urban agglomeration from three aspects of social economy, transportation and urban agglomeration, as shown in Fig.2. In aspect of urban agglomeration, urban flow intensity F was selected as evaluation index X_8.

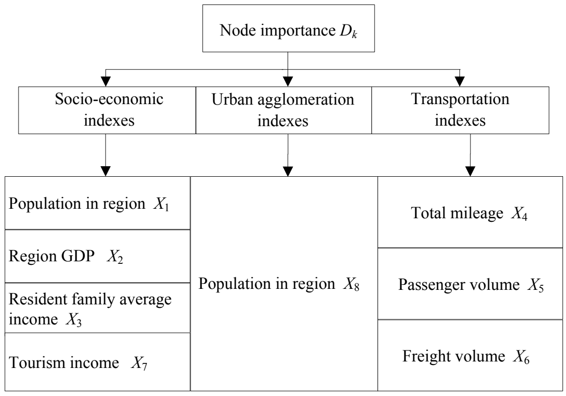

Fig. 2 Network of node importance evaluation indexes system

Table 1 Statistical data of indexes in central China urban agglomeration

Node Name	X_1	X_2	X_3	X_4	X_5	X_6	X_7
	/10^4 persons	/10^8 RMB	/RMB	/km	/10^4 persons	/10^8 tons	/10^4 dollars
Zheng Zhou	663	3 003.99	15 732	12 134	21 395	11 387	12 650
Kai Feng	484	689.37	11 342	8 574	4125	4 127	3 481
Luo Yang	654	1 919.64	14 672	17 636	9 009	8 372	11 285
Ping Ding Shan	501	1 067.60	13 531	13 089	6 536	6 501	276
Xin Xiang	561	949.49	13 000	12 599	5 068	5 184	254
Jiao Zuo	347	1 031.59	13 199	7 172	3 733	8 038	5 152
Xu Chang	456	1 062.05	12 448	8 925	5 354	9 740	171
Luo He	257	550.26	12 364	5 037	2 645	2 132	213
Ji Yuan	68	288.35	13 809	2 142	2 161	2 097	176

Note: The data is from "Henan Statistical Yearbook, 2009" (Bureau of Henan province).

Node importance evaluation indexes statistical data values of social economy and transportation are shown as Table 1. City employees index of second industry and tertiary industry is used in urban flow calculation such as mining, manufacturing, electricity, gas, wholesale and retail trade, real estate, hotel, restaurant and so on. Urban agglomeration city flow intensity can be calculated by (2) ~ (6), and the results are shown in Table 2.

Table 2 Outward function capacity and urban flow intensity in central China urban agglomeration

Node Name	Outward function capacity E_i	Function and benefit N_i	Urban flow tendency K_i	Urban flow intensity X_8
	/10^4persons	/10^8 RMB per 10^4 persons	%	/10^8 RMB
Zheng Zhou	16.805	30.950	17.315%	520.126
Kai Feng	5.016 4	26.876	19.557%	134.820
Luo Yang	6.169 8	38.454	12.359%	237.254
Ping Ding Shan	10.52 3	23.019	22.690%	242.242
Xin Xiang	5.647 8	22.907	13.626%	129.373
Jiao Zuo	4.944 6	33.342	15.981%	164.861
Xu Chang	4.183 0	38.775	15.272%	162.198
Luo He	3.791 9	26.506	18.265%	100.507
Ji Yuan	1.069 6	46.734	17.336%	49.987 5

Note: Data is from "China Statistical Yearbook, 2009" (Bureau of China) and "Henan Statistical Yearbook, 2009".

According to statistical data and city flow intensity calculation values in Tables 1 and 2, SPSS 18.0 is used for factor analysis. The results of main factors analysis matrix, indexes weights and main factors weights are shown in Table 3.

Consider data values of Tables 1, 2 and 3 into Eq. (1), node importance and sequences of nine cities in central China urban agglomeration are obtained as shown in Table 4.

Table 3 Calculating results by factor analysis method

Indexes Variables	Factor scores: The first main factor	Factor scores: The second main factor	Weights α_i
X_1	0.823	0.550	0.714 346
X_2	0.990	–0.103	0.554 986
X_3	0.768	–0.465	0.277 266
X_4	0.735	0.598	0.680 474
X_5	0.937	–0.229	0.472 932
X_6	0.844	0.101	0.548 286
X_7	0.864	–0.168	0.453 264
X_8	0.943	–0.168	0.500 822
Variance	4.228	2.790	The first main factor weights: 0.602 The second main factor weights: 0.398
Contribution rate (%)	52.848	34.877	
Total contribution rate (%)	52.848	87.725	

Table 4 City node importance and sort in central China urban agglomeration

Node Name	Zheng Z.	Kai F.	Luo Y.	Ping D. S.	Xin X.	Jiao Z.	Xu C.	Luo H.	Ji Y.
Importance	9.039	3.360	6.841	4.192	3.676	3.829	3.776	1.941	1.167
Sequence	1	7	2	3	6	4	5	8	9

4.3 Node importance levels division in central China urban agglomeration

By SPSS 18.0, node importance values are classified. The values of classification number K , maximum iterating times N and convergence standard are 4, 10 and 0, respectively. The clustering results and nodes classification are shown in Table5.

Table 5 Cluster and city node classification results

Final cluster centers			
1	2	3	4
9.04	6.84	3.77	1.55

City node classification in central China urban agglomeration			
Pole	Sub-pole	Important nodes	Common nodes
Zheng Zhou	Luo Yang	Pingdingshan Xinxiang Jiaozuo Xuchang Kaifeng	Luo He Ji Yuan

From Table 5, the node importance is divided into four levels by clustering: pole, sub-pole, important node and common node. Zheng Zhou and Luo Yang have the highest node importance and they are respectively the pole and sub-pole, which accord with “dual-core” characteristic of central China urban agglomeration layout planning. Luo He and Ji Yuan have the lowest node importance and they are common nodes. Other cities are important nodes.

5 Conclusions

Node importance evaluation method for highway network of urban agglomeration is examined in this paper. For the extroverted characteristics of urban agglomeration such as links between cities and outward function of cities, urban flow intensity is included into node importance assessment indexes system. Factor analysis avoids random subjective values and makes results objective in node importance calculation. The K-Means clustering method distinguishes degree level of each node importance. The result of central China city agglomeration application example shows the method proposed has good practical value. As node importance evaluation method in this paper is based on quantitative analysis, if there are individual nodes affected by factors such as politics, military, geography and so on, failing to reflect the real importance accurately, qualitative analysis should be made accordingly.

References

[1] Wang F Z, Guo Z F, Liu X L, et al. The structure optimization of the urban system in Zhongyuan region: Based on the integrated development of the urban agglomeration, Geographical Research,2007, 4(26): 637.

[2] Hu L G, Cheng L Q. Highway network layout of the city groups based on the layout method of node important degree with area location. Journal of Railway Science and Engineering,2008,5(1): 87–90.

[3] Miao J H, Yuan Z Z, SHA B, et al. Application of comprehensive evaluation method to deciding node importance of highway network. Transportation Science & Technology, 2005(5): 90–93.

[4] Jia Q, Yan K F, Zhang B Y, et al. Research on planning of highway transportation junction based on importance of nodes and traffic location. Shandong Jiaotong Keji, 2008(4): 4–9.

[5] Luo Z Z, Zhang F Y. Application of principal component analysis method in calculating the weights of the indexes of the highway nodes’ importance. Journal of Transportation Systems Engineering and Information Technology,2005, 5(6) :78–81.

[6] Zhou W, Yao Z G, Wang Y Q, et al. Optimizing items order of highway transport terminals based on analyzing importance of nodes. Journal of Chang’an University (Natural Science

Edition), 2006, 26(2): 70.

[7] Ma Y Q, Tan S P, Lan W L. The choice of logistics center location based on the regional highway network. Logistics Management, 2009(15): 11–13.

[8] Meng D Y, Fan K S. The intensity and rank of urban flow in Zhongyuan urban agglomeration. Journal of Shang Qiu Teachers College, 2010, 26(6): 116–118.

[9] Lu W D. SPSS for windows statistical analysis (The Third Edition), Publishing House of Electronics Industry, 2006.

[10] Li B C, Shao M Z, Huang J. Pattern recognition principles & applications. Publishing House of Xidian University,2008.

[11] Henan development and reform commission. Outline of Overall Development Plan for Central China Urban Agglomeration (2006–2020), 2006.