PLANNING OF DEVELOPMENT OF ROAD NETWORK BASED ON CLUSTER ANALYSIS

Komarov, Constantine L. *– D. Sc. (Tech), professor, head of the department of system analysis and project management of Siberian State University of Railway Engineering (STU), Novosibirsk, Russia.*

Zykova, Valeria Yu. *– Ph.D. student of Siberian State University of Railway Engineering (STU), Novosibirsk, Russia.*

Kuzmitskaya, Maria A. *– Ph.D. student of Siberian State University of Railway Engineering (STU), Novosibirsk, Russia.*

ABSTRACT

The object of research is existing and promising road network of areas that are parts of Novosibirsk agglomeration. Method of cluster analysis provides an objective assessment of road construction projects, their socio-economic relevance to local conditions and opportunities. A program of automated calculation of stable clusters is offered. Within the planning of transport systems development in the region approaches are developed to justify priority construction of roads and road structures of public service of regional and intermunicipal significance with account of socially significant economic effects.

ENGLISH SUMMARY

Background. *Weak development of the road network constrains economic development of Russian regions. In general, roads together with vehicles bring to their territories significant economic and social benefits. Experience of many countries shows in particular, how a developed network of rural roads acts as a catalyst for the development of agricultural regions. For instance Chinese practice perfectly confirms dependencies of economic results within the agglomeration on available transport capacity. Meanwhile still about thirty nine thousand settlements in Russia do not have reliable road links with regional centers. As a result, people do not receive full health care and other social services. In these circumstances underdevelopment of roads increases loss of economic benefits of agricultural producers and reduces investment activity in the regions.*

Objective. *The objective of the authors is to conduct a study of construction priority related to construction of roads and road facilities in Novosibirsk region with the help of cluster analysis method.*

Methods. *The authors used cluster analysis, mathematical calculations and comparative method.*

Results.

Approaches to the problem

Every year during the development and approval of the budget of Novosibirsk region, as well as of subjects of the federation, particular attention is paid to items «Road fund» or «Road construction». This is explained by the weak overall availability of roads in the country (in Novosibirsk region it is of 56,5 km per 1000 km ^{2}).

Currently not less than 30 billion rubles should be invested annually in the construction of new roads, their maintenance, and development of related advanced technologies to improve the situation in the region. In 2012 10, 3 billion rubles were spent on repair and construction of roads in the region, in 2013 – almost 12 billion. But answers to the questions: how effective is the return on investment and how to optimize efforts are not so clear and definite.

Plan for reconstruction and construction of roads is usually based on professional experience and political expediency. In this regard expected effects, including social ones, from construction of new road complexes are not achieved.

Since each district and each settlement has its own interests and proposals, which often do not fit into the approved budget items, it is relevant for maintenance of acceptable balance of local and region-wide needs to develop recommendations for planning and prioritization of construction of roads and road structures of regional and inter-municipal significance. This problem was solved in the course of studies, conducted by the authors, in the areas included in Novosibirsk agglomeration based on cluster analysis and objective assessment of social effectiveness of each project:

1) Existing traffic networks (hereinafter-TN) (Pic. 1) were analyzed.

2) Cluster approach to development planning of road network for different conditions and opportunities was used.

3) A program for automated analysis of stable clusters as areas with increased transport availability for residents living in the area allocated to the respective cluster was developed.

4) A method of calculating social effectiveness with regard to the number of residents, improving conditions of transport availability per unit of investment was proposed.

Cluster analysis is a set of mathematical methods designed to generate groups of remote objects which are relatively close to each other via criteria of distance or relations (proximity measures) between them [2]. In our case, it is a group of settlements (TN), most involved in transport interaction.

Cluster analysis method enables to classify objects using attributes that reflect essence, nature of the objects; to check presuppositions about the existence of some structure in the target group of objects; to build new classification for poorly studied phenomena when it is necessary to establish relationships, notably in that very case of transport links, within the set of settlements [1].

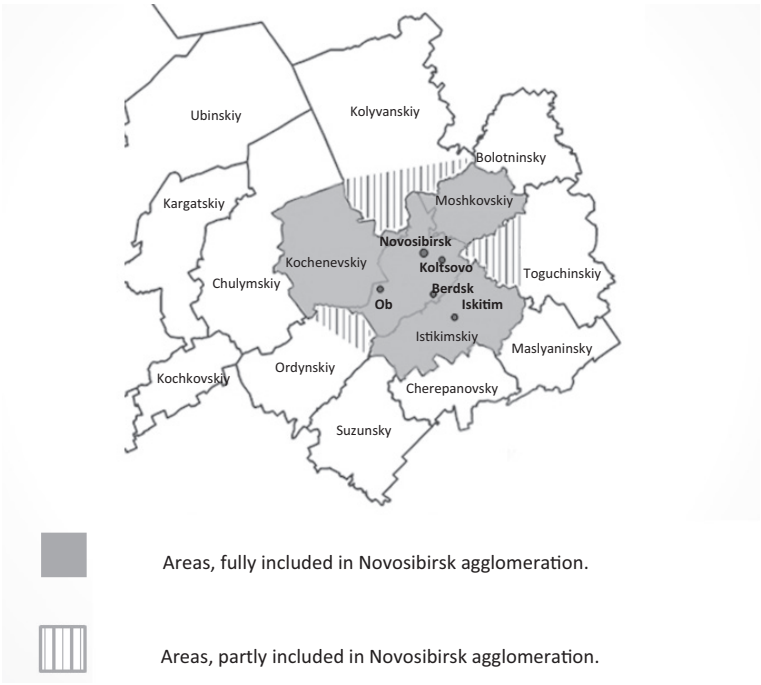
Formal representation of TN

As input data, taken as a basis, an approved list of construction of roads and road structures of regional and inter-municipal significance and administrative map of Novosibirsk region (scale 1:600000) were used.

Administrative map shows an overview of areas and road network, as well as graphical information on the problem. At the stage of formalization of the input data TN was shown as a simplified system of nodes (settlements) and the connections between them in the form of federal highways, roadways, highways, bridges and so on.

Traffic network of areas included in agglomeration, comprises 411 existing nodes.

At the next stage of formalizing of the input data the authors obtained binary matrices for each region, which are the interaction matrices of TN settlements. When drafting them the following principles were met:



Pic. 1. Novosibirsk agglomeration.

– The number of rows equals the number of columns and, accordingly, the number of nodes of TN;

– Intersection of node «with itself» is taken equal to «1";

– A direct link between nodes i and j implies «1» at the intersection of i-column with j-row;

– The lack of direct link between nodes implies «0» (Pic. 2).

Then, with transition to the situation related to commissioning of new roads or road structures, TN changes:

– Appropriate number of links that arise with the introduction of a particular road or structure is added (this is reflected graphically with change in location of the cluster);

– Existing links are neutralized between nodes, if following the emergence of a new road or a new structure a new node is formed between two already existed nodes.

Construction of matrices after insertion of changes in the adjacency table is conducted on the basis of the original matrix by:

– Changes in quantitative parameters (dimensions of the matrices);

– Changes in qualitative (substantial) component, reflecting structural properties of a new matrix with respect to initial content.

Thus, this process is a formal representation of the traffic network of areas included in the agglomeration and binary adjacency matrices of TN make it possible to obtain all necessary input data for problem analysis.

Formal technique of analysis

Cluster analysis method was selected as the formal technique for following reasons:

1) Ability to process large databases using software package.

2) The natural decomposition of TN aggregate on congestion areas of objects (clusters).

3) Submission of initial data in the form of adjacency matrix.

4) Ability to analyze data sets and make own conclusions, depending on the task.

	Sedovozaimka	Bibikha	Zeleniy mys	**Kubovaya**	Sosnovka	Stepnoy	Krasniy yar	Lomovskaya dacha	**Mochishche**	Vorobievskiy	Kudryashevskiy	Priobskiy
Sedovozaimka	1	1	0	0	0	0	0	0	0	0	0	0
Bibikha	1	1	1	0	0	0	0	0	0	0	0	0
Zeleniy mys	0	1	1	1	0	0	0	0	0	0	0	0
Kubovaya	0	0	1	1	1	0	0	0	0	0	0	0
Sosnovka	0	0	0	1	1	1	0	1	0	0	0	0
Stepnoy	0	0	0	0	1	1	0	1	0	0	0	0
Krasniy yar	0	0	0	0	0	0	1	0	1	0	0	0
Lomovskaya dacha	0	0	0	0	1	1	0	1	0	0	0	0
Mochishche	0	0	0	0	0	0	1	0	1	0	0	0
Vorobievskiy	0	0	0	0	0	0	0	0	0	1	1	0
Kudryashevskiy	0	0	0	0	0	0	0	0	0	1	1	0
Priobskiy	0	0	0	0	0	0	0	0	0	0	0	1

Pic. 2. Fragment of interaction matrix of Novosibirsk region settlements.

There are still no generally accepted methods of identification of clusters, however, recently a cluster analysis has received sufficient recognition [4], and there is no reason to doubt its validity.

Software-analytical complex

Statistical Package for the Social Sciences (SPSS) was selected as statistical package to conduct necessary analysis.

The main advantage of this software is the fact that it is successfully combined with a large number of visualization tools of processing results.

In order to highlight the true clusters, it is necessary to verify the sustainability of data, thereby identifying local transport clusters predisposed to unite in a solid cluster earlier than others.

Adjacency matrices of the graph of TN of areas that are included in the agglomeration, in the format of MS Excel, served as a source for the SPSS program.

Table 1

Example of stability diagram of clusters of TN of Novosibirsk region

10	9	8	7	6	5	4	3	2	Brought to the scheme	Number of nodes in the cluster (1)	Number of nodes in the cluster (2)	№№ of nodes in the cluster
1	1	*9								35		
2	2	2	2	2	2	*1			3	4	4	5,6,8,40
3	3	3	3	3	*1				4	4	4	7,9,41,42
4	*5									4		
5	*4									4		
6	6	6	*1							5		
7	7	7	7	*1					5	6	6	43,44,61,62,64,65
8	8	8	8	8	8	8	8	8	1	1	1	47
9	9	*1								3		
10	10	10	10	10	10	10	*1		2	5	5	58,59,60,63,66
								Total coverage		71	20	

As a result of the processing of each of the matrices by SPSS the researchers got the tables showing the appurtenance of separate settlements to different clusters (Cluster Membership).

Particularity of the study was that «behavior» of nodes was evaluated in the constantly diminishing number of clusters from ten to two. Thus, affiliation tables show that the same number of nodes initially divided into ten clusters, gradually combining, reduces to two large clusters. During the destruction of initial clusters along with the development of TN new enlarged clusters are partially formed.

Stability diagrams of clusters (Table 1) allow to determine «life cycles» of each of the original aggregate of nodes prior to their enlargement, highlighting in such a way the most stable ones. For clarity, each cluster gets a serial number from one to ten. «Life cycles», respectively range from one (least stable) to ten (most stable) phases. On the right side of the diagram, there are four additional columns:

- *«Brought to the graph»;*
- *«Number of nodes in the cluster (1)»;*
- *«Number of nodes in the cluster (2)»;*
- *«№ № of nodes in the cluster».*

The first column shows the most stable clusters (in order of increasing stability from five to one) and their serial numbers (the row number) to be further incorporated into TN scheme.

The column «Number of nodes in the cluster (1)» describes the total number of nodes in each cluster (when broken down into ten clusters). In particular this can be used for cheking operation (the number of nodes in the diagram must match the number of rows and columns in the matrix).

The column «Number of nodes in the cluster (2)» refers to the number of nodes in the clusters selected by the stability criterion and brought to the diagram. By summing the number of nodes in this column we obtain primary criterion for evaluating the priority of road construction in areas of Novosibirsk agglomeration. This refers to the coverage of certain localities with the most stable clusters. Using this criterion an initial evaluation and comparison of competing options are produced. In terms of this technique the most preferable highway is that, where five most stable clusters cover the largest number of nodes (settlements).

The fourth column «№ № of nodes in the cluster» helps to identify and add to the scheme clusters by numbers of nodes, included in it.

Thus, software and analytical complex SPSS enables to process input data by hierarchical cluster analysis method and interpret the results in the form of conclusions for initial evaluation of priority of road construction of regional and intermunicipal significance.

Justification of priority

On the basis of analysis of the stability diagrams five most stable clusters have been identified from the graph of each of the areas. After that, we obtain a quantitative initial assessment of priority of the construction of roads and road facilities following the principle of maximum system stability.

In Novosibirsk region according to long-term building plan the effectiveness of four objects was analyzed:

1. Approach roads to the industrial and logistics park (ILP).

2. Highway Baryshevo-Orlovka- Koltsovo, with its road tunnel, combined with the railroad.

3. Transport junction and approach roads to the exhibition complex «Expocentre of Siberia».

4. Road from Krivodanovka to Northern Bypass of Novosibirsk.

Based on the analysis primary effectiveness of roads was subject to a comparative assessment of the number of nodes (settlements) englobed by five most stable clusters.

To determine the optimal sequence of the implementation plan of road construction works, the authors introduced a social coefficient ($K_{соц}$).

With its use social efficiency of projects is estimated based on the number of residents, who will consequently improve their mobility, and per capita consumption of investments:

$$K_{соц} = K_{пр}/C \cdot 10^3 \qquad (1)$$

where $K_{пр}$ is a number of residents living in the cluster (person); C – cost of the project (mln rub).

According to the data obtained the following order of construction priority in Novosibirsk region was developed:

1) Approach roads to the industrial and logistics park.

2) Road from Krivodanovka to Northern Bypass of Novosibirsk.

3) Transport junction and approach roads to the exhibition complex «Expocentre of Siberia».

4) Project of a highway «Baryshevo-Orlovka-Koltsovo, with road tunnel with the railroad».

In Iskitimsky area under long-term plan for construction of roads and road facilities the effectiveness of four objects was analyzed:

1) Road Guselnikovo – Linevo.

2) Bridge crossing over the river Karacan at 55 km of road Zavyalovo – Fakel Revolutsii.

Table 2

Indicators of social efficiency of road construction in Novosibirsk agglomeration in 2014

Areas where the project is being implemented	Name of the road-building project	Cost of the project (million rubles).	Number of residents living in the cluster (million people).	$K_{соц}$ (million people/ million rubles) $\cdot 10^3$	Recommended sequence of the implementation plan of road construction works
Novosibirsk area	Highway Baryshevo-Orlovka- Koltsovo, with its road tunnel, combined with the railroad	775	1,550030	2,00	4
	Approach roads to the industrial and logistics park of Novosibirsk area	187	1,526959	8,17	1
	Transport junction and approach roads to the exhibition complex «Expocentre of Siberia»	400	1,538621	3,85	3
	Highway from Krivodanovka to Northern Bypass of Novosibirsk	285	1,530180	5,37	2
Iskitimskiy area	Highway Guselnikovo-Linevo	280	0,077552	0,28	4
	Bridge crossing over the river Karacan at 55 km of road Zavyalovo – Fakel Revolutsii	75	0,073705	0,98	2
	Highway and bridge crossing over the river Eek on the road Verkh-Eeki- Novososedovo	70,639	0,073705	1,04	1
	Highway N-0809-Kharino	95	0,074318	0,78	3
Ordynskiy area	Route 118 km of highway K-17p – Kamen-na-Obi at the site Kirza – the border of Altai Territory	1531,153	0,027762	0,018	2
	Bypass near regional village Ordynskoe	435	0,028725	0,066	1

3) Highway and bridge crossing over the river Eek on the road Verkh-Eeki- Novososedovo.

4) Highway N-0809-Kharino.

According to the results of analysis the first cluster covered 25 settlements, the second – 22, the third – 22, and the forth-19.

Based on the findings in this area the following sequence of construction is proposed:

1) Highway and bridge crossing over the river Eek.

2) Bridge crossing over the river Karacan.

3) Highway N-0809-Kharino.

4) Highway Guselnikovo-Linevo.

In Ordynskiy area the effectiveness of two construction projects was analyzed by the same method:

1) Route 118 km of highway K-17p – Kamen-na-Obi at the site Kirza – the border of Altai Territory;

2) Bypass near regional village Ordynskoe

According to the results of the analysis the first cluster comprised 21 localities, the second 18.

Assessment of obtained data set gave the priority for the bypass road of Ordynskoe settlement and the second object would have to wait.

Estimates of the social efficiency of road construction in 2014 are presented in Table 2.

Conclusions. *According to the authors, recommendations concerning choice of priority construction of highways, and results obtained in the course of cluster analysis may serve as confirmation of objectivity of method of comparing competing transport projects and may promote the formalization of procedures for the preparation of integrated management solutions.*

Input data for the study were provided by Ministry of Transport of Novosibirsk region and the study results were approved by this ministerial body and regional administration, members of the Legislative Assembly of Novosibirsk region. Therefore the proposed planning and analysis tools can be considered as initially adapted and validated.

Keywords: transport, region, cluster analysis, hierarchical cluster analysis, roads, road network, development planning.

REFERENCES

1. Labotskiy, V. V. Knowledge management [*Upravlenie znaniyami*]. Minsk, Sovremennaya shkola publ., 2006, 392 p.

2. Novorozhkina, L.I., Arzhenovskiy, S. V. Multivariate statistical methods in economics. Textbook [*Mnogomernye statisticheskie metody v ekonomike*]. Moscow, Dashkov i C° publ., Rostov-on-Don, Nauka-Spektr publ., 2007, 224 p.

3. Sbitnev, A. E. Conceptual problems of defining social efficiency of investment projects in road construction [*Kontseptual'nye problemy opredeleniya obschestvennoy effektivnosti investitsionnyh proektov v dorozhnom stroitel'stve*]. *Problemy sovremennoy ekonomiki*, 2011, No.2, pp.280–283.

4. Pyataev, M. V. Evaluating the effectiveness of the formation of regional transport and logistics clusters (at example of Novosibirsk Region). Ph.D. thesis [*Otsenka effektivnosti formirovaniya regional'nyh transportno-logisticheskih klasterov (na primere Novosibirskoy oblasti). Dis... kand. ekon. nauk.*]. Novosibirsk, 2010, 179 p.

5. Mastilin, A. E. Backbone transport system components of a metropolitan city [*Sistemoobrazuyuschie elementy transportnoy sistemy megapolisa*]. *Region: ekonomika i sotsiologiya*, 2008, No.4, pp. 265–272.

Координаты авторов (contact information): Комаров К. Л. (Komarov, C. L.) – komarovkl2015@gmail.com, Зыкова В. Ю. (Zykova, V. Yu.) – zikovavyu@mail.ru , Кузьмицкая М. А. (Kuzmitskaya, M. A.) – kuzmickayama@mail.ru.

Статья поступила в редакцию / article received 03.03.2014
Принята к публикации / article accepted 14.05.2014