

Linear structure taxonomy with the account of environmental polystructures impact

Tatyana Bobrova^{1,*}, and *Valeriy Vorobyev*²

¹Siberian State Automobile and Highway University (SIBADI), 644080 Omsk, Russia

²Siberian Transport University, 630049 Novosibirsk, Russia

Abstract. One of the specific features of transportation facilities design is the necessity to account changing environmental conditions in considerable distances along the track. The purpose of the study is theoretical substantiation of structural decomposition of linear structures based on the degree of uniformity of environmental factors effect on the geotechnical system of an object. A new approach to linear road zoning is suggested on the basis of comprehensive assessment of multidimensional spatial data on the environment. Theory of pattern recognition is used as a classification tool. Linear structure taxonomy, as a part of this theory, is based on hierarchical agglomerative clustering algorithms which provide consequential merge of operating territorial units into uniform linear complexes. Application of taxonomy methods is considered for automated tasks solving on various levels of roads information modeling: when designing structures, organizing construction and operating a facility.

Introduction

Various aspects of the impact of environmental polystructures on a linear structure are considered with respect to the stage of its life cycle. For the purpose of control of road structures with the account of environmental factors impact, regulatory documents and design guidelines adopted in different countries provide the use of systematized territorial zoning data [1-5]. Today, development of reliable road structures requires building regional climatic models and using special software to include the optimization techniques in the design process [6-8].

Various approaches exist for solving road climatic zoning issues, they are based on assessment of various conditions, including environmental and climatic, geological and geographical, physico-climatic parameters of the territory [5, 9]. Differentiation methods (division of territories) used for transport infrastructure design are developed on the basis of the general theory of physico-geographic zoning with the account of the industry peculiarities. Road climatic zoning of the type "zone-subzone-district" is carried out with the application of structured approach principles using correlation and regression analysis methods, as well as methods of principal components analysis, parameters gradation and cluster analysis [10-12]. Weather stations and zero-dimensional objects located on the road

* Corresponding author: bobrova.tv@gmail.com

network are taken as structured space objects. Territorial zoning enables to assess the chosen route direction on the pre-project phase and to substantiate the essential structural solutions. At the same time, design faults related to underestimating of environmental factors variability in a certain geological and geomorphological situation lead to intolerable deformations and structural failures on the road sections during operation [13]. It is especially important to prevent faults and adopt adequate design solutions in complex climate conditions [9].

Bigger number and higher accuracy of parametric characteristics of environmental polystructures are required on the stage of designing certain elements of a linear object. Solving this task becomes ever more relevant with the development of road information modeling (RIM) technology. A new technological framework is formed in economics today in Russia and other countries which is related to information and digital technologies. The RIM technology differs considerably from on-site structures modeling. This technology is primarily aimed at providing automatic calculations of parameters in relation with multidimensional spatial data on natural environment at different design stages [14, 15].

The purpose of the study is theoretical substantiation of structural decomposition of a linear structure by the level of uniformity of environmental factors impact, and development of mathematical model for its implementation in a CAD system. In contrast to the previously applied zoning principles, the approach used in the study is based on unstructured data representation in the form of cluster of linear elements. The main input data for modeling in the form of environmental factors characteristics on elements (stations) are formed on the stage of surveying along the object trace.

Research methods

Multiple theoretical substantiations of zoning are based on hierarchical division of the studied objects into classes according to similarities and differences of their essential attributes. This approach takes its basis in the theory of pattern recognition (PRT), in which a pattern denotes a combination of data on an information object. The data include quantitative parameters obtained with the use of measurement systems or mathematical models [16]. Any recognition algorithm can be presented as an abstract functional system R consisting of three components:

$$R = \{A, S, P\}, \quad (1)$$

where $A = \{A_k\}$, where $k = 1, \dots, K$ – alphabet of classes – aggregation of categories to which patterns are distributed, $S = \{S_j\}$, $j = 1, \dots, n$ – attributes dictionary – aggregation of characteristics which are used to describe the patterns, $P = \{P_l\}$, where $l = 1, \dots, L$ – number of rules for decision making.

Depending on the peculiarities of information components of system R three approaches to pattern recognition are distinguished: 1) template matching; 2) clusterization (taxonomy); 3) similarity of properties. In this system taxonomy is considered as a science dealing with classification and systematization of structures with complex organization. Choice of mathematical method depends on classification purposes applicable to components of "polystructural environment - linear transportation facility" system, as well as on the life cycle phase of the object: design, construction, operation.

Object of the study is a complex of a linear object and natural environment represented in CAD system using survey data. At this stage, calculations are highly detailed and are done with parametric modeling based on input data. For adopting reasonable structural solutions data of ground penetrating radar scanning along the axis of the road, right-of-way space and cross-section are used. Survey documentation includes detailed characteristics of

engineering geological units (EGU) and their state. Road can be represented as an aggregate of structurally similar sections adapted to the geocomplex conditions.

According to the selected PRT method - taxonomy - let us define the peculiarities of this method that are related to linear extended elements of transportation facilities, primarily to subgrade and road pavement for motorways. The input data is presented in GIS for transportation in the form of multidimensional spatial data. The term "multidimensional" is related to attributive indicators (numerical and order ones) of various environment characteristics, whereas the term "spatial" specifies location of these data with a station mark on the road axis. To reduce the scale, transition to operational territorial units (OTU) was performed due to integration of station sequences. Integration is performed until the value of at least one indicator exceeds the standard values of coefficient of variation (for physical parameters - 0.15; for mechanical parameters - 0.30). Thus, before taxonomy procedure starts, the road is divided into discrete OTUs with the average values calculated on the stations integrated in a single OTU. Then, the input data are presented in the form of "OTU - attribute" incidence matrix. As a result of applying standard classification algorithm [17] OTUs are grouped into geographically distributed classes (taxa) with relatively uniform environmental conditions. With the account of the industry specificity, these taxa are denoted as linear road complexes (LRC), and the process of taxonomy development on the level of a CAD model – as linear road zoning.

Euclidean distance is used to assess similarities (differences) between OTUs in N -dimensional fuzzy attribute space. Classification is aimed at obtaining a certain predetermined number of groups (classes) of OTUs S_1, \dots, S_K . Within one group, OTUs shall possess the maximum degree of "similarity", and OTUs that belong to different groups - the maximum degree of "difference". The paper [18] gives the following mathematical interpretation of OTUs classing. If a set of OTUs is written as $O = \{o_1, \dots, o_N\}$, where o_i - i^{th} OTU, N - number of OTUs, then the result of classing is represented by an N -dimensional vector $v = (v_1, \dots, v_N)$, $v_i \in \{1, \dots, K\}$ of OTU pertaining to one of the classes S_j :

$$v_i = j \Leftrightarrow o_i \in S_j, \quad i \in \{1, \dots, N\}, \quad j \in \{1, \dots, K\}. \quad (2)$$

Various standard methods are used to make the final decision on the number of classes [16]. The experience of using taxonomies for road objects showed that formal criteria of classification can lead to excessive division of territorial units into classes, and consequently, to bigger number of LRCs of smaller length. The final decision on structural decomposition of a linear object according to environmental conditions (in other words, on the number of classes) shall be made by a design engineer on the basis of analytical data obtained when applying taxonomy algorithm. To fulfill the set task it is convenient to apply the method of automatic visualization of classification using the ER-model or relational data model [18]. Simpler models can be used, such as bar graph, to demonstrate average values of environment attributes and coefficient of attributes variation by placing them next to the corresponding class (LRC).

The result of linear taxonomy implementation is decrease in dispersion of environmental factors within a LRC. Quality indicators of a road structure such as strength, robustness, reliability, etc. are functionally connected with qualitative parameters of environment, which are the basis of the most part of design calculations. Using the formula for error propagation known from the probability theory, variance of a function (result) is defined based on the variance of its elements:

$$\sigma_F^2 = \sum_{i=1}^n \left(\frac{\partial F}{\partial x_i} \right)^2 \cdot \sigma_{x_i}^2, \quad (3)$$

where F – the considered function, variance of which σ_F^2 is sought; $x_i, \sigma_{x_i}^2$ – element of the function (argument) and its variance.

This formula indirectly defines the necessity of forming such structural units based on area or length, within which uniform road structures will possess similar strength and robustness parameters. Further research suggests assessment of output response parameters of road structures caused by changes in variation of environmental factors parameters. This may serve as a basis for control and optimization of LRC length on a linear structure.

Results and application

The described method was used for developing design documentation for a road section with the length of 47.4 km located in the permafrost region. The average coefficient of variation of environmental factors parameters on the LRC did not exceed 20%, with 80% and higher on the road in general. Due to this, design solutions with higher degree of reliability could be implemented [19]. Table 1 presents the formulated suggestions on the use and development of taxonomy methods for linear objects at various design stage.

When performing information modeling, 6 levels of development of information model elements (LOD) are considered: pre-design, schematic design, design development (CAD), construction documentation, construction stage, as built [15]. Processing of information contained in geographical environment shall be performed at each level.

Table 1. Taxonomy methods in road information modeling system.

LOD	Spatial grading of natural environment	Group of classification factors for natural and man-made environment	Classification methods / criteria	Groups and elements of RIM for design	Decision area
Schematic design	Zone, subzone, road district	Climate, relief, geomorphology, hydrology, soil design paramters	Hierarchical division algorithm (territory division); post-hoc analysis / key factor	Standard structures of the object elements	Feasibility study, Environmental assessment
Design development	Station, OTU, taxon, LRC	Engineering survey data, special conditions (permafrost, seismic activity, etc.)	Hierarchical agglomerative mechanism (OTU clustering) / proximity measure: Euclidean distance	Detailed design of all elements in LRC section	Technical project, structural design, production drawings
Construction documentation;	LRC	Type and scope of works, technology, seasonality, resources	Methods of PRT: template matching, similarity properties of	Methods of work performance. Resource provision	Process design (construction management plan and project execution plan)

As built	Zone, subzone, maintenance section	Physico-climatic factors, azonal factors, man-made structures	Division of area by negative conditions parameters (snowfalls, blizzard, ice).	Parameters and state of road elements	Cost regulation, recruiting employees for work performance, maintenance routes
----------	------------------------------------	---	--	---------------------------------------	--

Implementation of taxonomy algorithm of a transportation facility was performed automatically using MS Excel. However, more advanced software is known [18]. When changing environmental conditions, the necessity to adjust project solutions may arise. Application of taxonomy allows making corrections in the process of road information modeling, i.e. performing variable design, adjusting certain previous solutions on taxonomy considering changes in factors of geocomplexes.

The final decision on formation of LRC is made by a design engineer based on visual presentation of automated zoning results. The following variants are possible: 1. Group sections with significant deviations on any regional factor in a separate linear complex (taxon); 2. Group adjacent sections of LRC or OTU if the suggested structural solutions of structural elements fulfill the standard requirements; 3. Adjust the LRC length by suggesting measures on changing natural environment conditions, for instance, due to engineering of drainage system, lowering of ground water level, etc.

Not only structural solutions are connected with spatial decomposition of a linear object, but also the peculiarities of process design and operation of an object (Table 1). On the level of a construction model, differences in structural solutions for LRC are accounted while designing construction management. Variations may concern methods of organizing flows, list of special squads, terms of works execution. Application of parallel flow method when executing works on LRC enables to reduce terms of construction.

When designing as-built level of a road model, most significant climate indicators are defined which affect the traffic conditions in winter season. This way, zoning of Western Siberia region was performed by three classification groups of the following environmental polystructures: snowfalls, blizzards and ice [20]. Linear division of roads during operation is connected with zoning and is performed with the account of regional and azonal factors, both natural and human-made: shelter belts, crash barriers, heavy cuts, etc.

Conclusion

The method of linear structure taxonomy provides scientifically proven approach to an object division into sections with more uniform environmental parameters. Such structural decomposition results in increase of reliability level and efficiency of design solutions. Inclusion of linear objects taxonomy method in the road information modeling system contributes towards solving the scientific problem of improving the level of automation while designing transportation facilities on various levels of project model development.

Information modeling of transportation facilities currently is at the early stage of development. The paper demonstrates the possibilities of taxonomy as an automation tool for solving tasks of natural environment zoning and decomposition of linear objects for designing reliable road structures. Improvement and development of the method is planned for various levels of road parametric modeling. Further research in the field are primarily related to substantiation of optimum LRC parameters while designing standard subgrade

structures, as well as to the issues of efficient computer implementation of taxonomy for the purpose of linear extended objects zoning.

References

1. SP 34.13330.2012. Avtomobil'nye dorogi [Automobile roads] (MinRegion, Moscow, 2012). (in Russian)
2. R. Kleiziene, A. Vaitkus, J. Zidanaviciute, E. Marcinkevicius, Proceedings of the 10th International Conference „Environmental Engineering“ (Vilnius Gediminas Technical University Lithuania, April 27–28, 2017).
3. L. Jukneviuciute-Zilinskiene, Baltic Journal of Road and Bridge Engineering, **Vol. 5** (1), 62-68 (September, 2010).
4. M.M. Shallal, S.E.M. Ahmed, IOSR Journal of Engineering, **Vol. 4**, Issue 4 (April), 13-26 (2014).
5. M.A. Elshaeb, S.M. El-Badawy, E.A. Shawaly, American Journal of Civil Engineering and Architecture, **Vol. 2**, Issue 3, 115-121 (2014).
6. American Association of State Highway and Transportation Officials (AASHTO), *Mechanistic-Empirical Pavement Design Guide: A Manual of Practice. Interim Edition* (Washington D.C., 2008).
7. Q. Li, L. Mills, S. McNeil, The Implications of Climate Change on Pavement Performance and Design. UD-UTC Final Report (2011).
8. J. Saha, *Evaluation of Climatic Effects on Pavement Performance using MEPDG* (LAP LAMBERT Academic Publishing, 2011)
9. P.N. Kachroo, N.G.K. Raju, Freeze-thaw Effects on Roadways (Approach to Pavement Design with Special Reference to Roads in Mongolia), International Seminar on the Appropriate Use of Natural Materials in Road, Ulaan-Baator 18-21 June 2002.
10. C.E. Zapata, W.N. Houston, *Calibration and Validation of the Enhanced Integrated Climatic for Pavement Design* (Transportation Research Board, Washington, D.C., 2008).
11. V. Efimenko, S. Efimenko, A. Sukhorukov, A. Yankovskaya, *IOP Conf. Series: Materials Science and Engineering*, **Vol. 71**(1), 012049 (2015).
12. V. Efimenko, S. Efimenko, A. Sukhorukov, *MATEC Web of Conferences*, **Vol. 143**, 01012 (2018).
13. E.A. Bedrin, A.M. Zavyalov, M.A. Zavyalov, *Obespechenie termicheskoy ustojchivosti osnovanija zemljanogo polotna avtomobil'nyh dorog* [Providing thermal stability of the road bed] (SibADI, Omsk, 2012). (in Russian)
14. D.S. Sarychev, A.V. Skvortsov, *CAD & GIS for roads*, **1**(8), 9-12 (2017). (in Russian)
15. STO AVTODOR 8.6-2016. Organizacionnaja i tehnologicheskaja podderzhka processov formirovanija informacionnyh modelej avtomobil'nyh dorog na vseh jetapah zhiznennogo cikla [Organizational and technological support of the generation of information models for highways at all life cycle stages] (Avtodor, Moscow, 2016). (in Russian)
16. L.N. Chaban, *Teorija i algoritmy raspoznavanija obrazov* [Theory and algorithms of pattern recognition] (Moscow, 2004). (in Russian)
17. T.V. Bobrova, A.A. Dubenkov, Vestnik of Tomsk State University of Architecture and Building. English version appendix, **1-4**, 134-141 (2013).
18. A.V. Zavarzin, Thesis of Candidate of technical sciences, Lomonosov Moscow State University, Moscow, 2003. (in Russian)
19. A.A. Dubenkov, Vestnik SibADI, **4**(32), 46-52 (2013). (in Russian)

20. Yu.V. Kodentseva, Thesis of Candidate of technical sciences, Siberian State Automobile and Highway University, Omsk, 2007. (in Russian)