Cluster analysis in the classification of transport infrastructure of large cities

Anna I. Terekhova*, and *Denis N.* Vlasov

National Research Moscow State University of Civil Engineering (NRU MSCU), Department "Urban Planning", 129337 Moscow, Russia

Abstract. Subject of the research: The subject of the research are transport and transfer hubs in the city of Voronezh (population 1,054,111 people). Transport and transfer hubs ensure the connectivity of the transport network, provide transport services to urban residents, and are an essential element of the system of settlement centers. Determination of classes of TIH in large cities. The role and significance of the TIH in the transport and planning structure of a city is determined by their classification. The node position in the general classification allows to determine its composition, to provide optimal and comfortable for passengers planning solutions, etc., etc. The selection of factors for the formation of TPU is carried out using a single-factor analysis of variance, which allows us to exclude less significant indicators more accurately. The selection left only 6 indicators, which were taken for further analysis. Cluster analysis was performed using a clustering tree. The results are five clusters, from which 4 classes of transport and transfer hubs in accordance with SP 395.1325800.2018 "Transport and transfer hubs. Design rules" and their characteristics are given. Conclusions on the compliance of the obtained clusters with SP 395 are made, the main indicators for the cluster analysis are outlined.

1 Introduction

The formation of a sustainable transport system is one of the main tasks of the development of settlements, in order to form a sustainable, comfortable and safe environment on the territory of cities. Sustainable transport provides the development of many other concepts of urban development: from "smart city" (SMART CITY)[1] to the concept of "mobility as a service" (MaaS)[2].

The formation of a sustainable transport system is a complex task that requires the implementation of infrastructure projects (construction of city streets and roads, subways, development of urban railroads, etc., etc.) and organizational measures (coordination between modes of transport, the formation of a single parking space, control over the observance of traffic rules, etc., etc.) [3].

* Corresponding author: Anna.terekhova.93@mail.ru

2 Literature review

In the current federal Code of Rules SP 42.13330.2016 "Urban Planning. Planning and Development of Urban and Rural Settlements" [4] the classification of transfer hubs is not presented. In 2019, SP 395.1325800.2018 "Transport interchange hubs. Design rules" [5], which presents the classification of transfer hubs. The classification is based on the functional (transport) purpose and interacting modes of transport. In addition, when determining the composition of the TI system of a settlement, its size is taken into account [6,7].

3 Materials and methods

The main purpose of the study is to clarify the existing or develop a new classification of interchange hubs. The basis for the study is the study of a set of urban indicators that assess the existing and planned development of the city, in order to determine the role and importance of transport hubs not only as an element of transport infrastructure, but also as an element of the planning structure of the city [8]. The system of transport hubs in Voronezh is considered as an object of research. Plans for future research include testing the results on the systems of transport hubs in other cities.

The system of transport and transfer hubs in Voronezh includes 23 hubs. In accordance with the purpose of the study there were defined the indicators, which were further used to develop the classification of the interchange hubs. Indicators: passenger flow size during "peak" hours, modes of transport interacting in the node and their characteristics, zone of influence of the TIH (served territory), TIH territory, location relative to the city center, distance to the TIH, transfer time, distance of the TIH from the existing buildings, parking space capacity, existing location relative to the city functional zones, public transport network density in the TIH influence zone, building height limits (according to the land use rules

After collecting baseline data, it is necessary to determine which of the factors should be taken into account when forming the classification of transport hubs. In the study, the method of factor selection is a single-factor analysis of variance.

A factor is considered strongly connected if two conditions are fulfilled: F>F critical and P-value <0.05 (Table 1), if one of the conditions is not met the factor is considered weakly connected (Table 2). Weakly coherent factors are not considered in the TPU classification.

Table 1. One-Factor Analysis of Variance (both conditions are met).

Analysis of variance						
Source of variation	**SS**	**df**	**MS**	**F**	**P-value**	**F critical**
Between groups	2,24E+09	15	1,5E+08	**1,909141753**	**0,021885447**	**1,699170372**
Within groups	2,4E+10	306	78362510			
Total	2,62E+10	321				

Table 2. One-factor analysis of variance (condition F>F critical is not satisfied)

Analysis of variance						
Source of variation	**SS**	**df**	**MS**	**F**	**P-value**	**F critical**
Between groups	2,2E+09	13	1,7E+08	**1,85234507**	**0,03589443**	**1,757652302**
Within groups	2,4E+10	262	1522626			
Total	2,62E+10	275				

4 Study results

Thus, the Strongly Connected Factors are:

1 The size of passenger traffic during "peak" frequencies;
2 The zone of influence of the TIU (the area served);
3 Territory of the NTH;
4 Interchange distance at the NTH;
5 Transfer time;
6 Remoteness of the TIU from existing buildings.

Classification of transport and transfer hubs in large cities will be conducted by cluster analysis, as the cluster analysis allows to form clusters (groups), united by common characteristics on the basis of given features [11].

In our study, the hierarchical cluster analysis - dendogram will be used [9,10].

The dendogram is constructed using a similarity matrix or a difference matrix. Such a matrix shows a measure of similarity of objects, measured on an ordinal scale. To build a matrix, a distance measure must be defined.

After the distance measures are defined, the grouping stage follows. Merging is done gradually, starting with the closest clusters, gradually increasing the distance (decreasing the strength of the cluster connection) until all clusters are merged into one.

After obtaining the results of the cluster analysis, it is necessary to determine how many clusters are allocated in a given tree, namely to choose the distance after which combining the features into clusters will not be appropriate, that is, the strength of connection will be insignificant. In the study, the method for determining the number of classes is to analyze the graph of the association process (Fig. 1). On the graph you need to find the break point and the number of step m, where the break occurred, then the number of clusters is equal to n-m, where n- the number of objects in the sample.

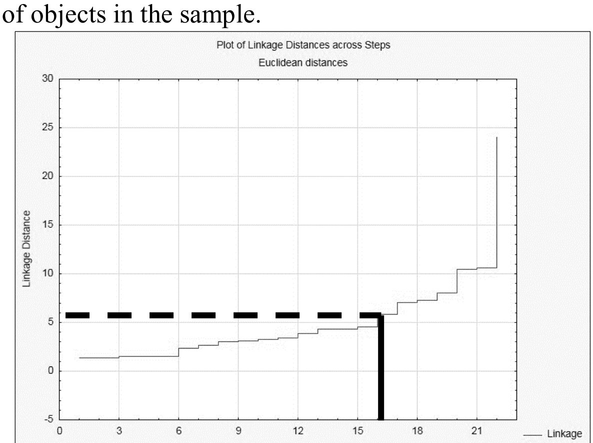

Fig. 1. Graph of the merging process.

Following the graph of the association process, the break point corresponds to step 16, and the number of objects considered equals 23, from this it follows that the number of clusters in the tree under consideration should be 7. Then it is necessary to mark on the hierarchical tree of cluster analysis with a perpendicular line at which step the number of clusters equals 5 (Fig. 2).

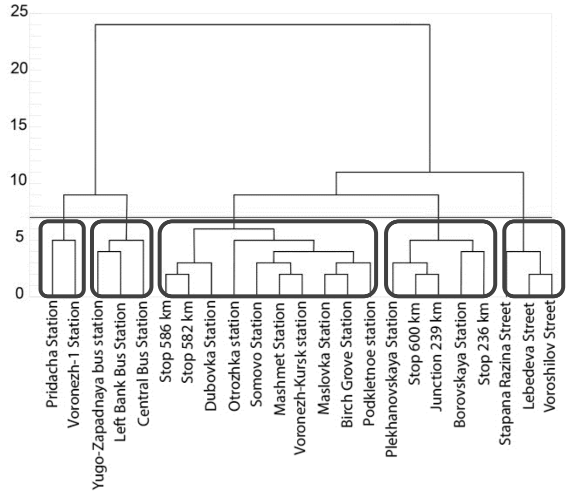

Fig. 2. Graph for determining the number of clusters.

In accordance with the results of the cluster analysis, we will classify the transport-transfer hubs of Voronezh in accordance with SP 395 [5,13].

5 Conclusions

1. As a result of the cluster analysis, 4 types (classes) of transport hubs can be distinguished in Voronezh, which correspond to SP 395:
- nodes of interregional significance;
- nodes of regional (agglomeration) significance;
- nodes of regional significance;
- nodes of local significance.

2. The analysis of variance identified significant factors for the classification of TPU:
1. The size of passenger traffic during "peak" periods;
2. The zone of influence of the TIU (served territory);
3. Territory of the NTH;
4. Transfer distance at the NTH;
5. Transfer time;
6. Remoteness of the NTH from existing buildings.

3. Thus, the TPU classification does not depend on the species interacting in the node, which is a major contradiction with SP 395.

4. To confirm the results obtained, it is necessary to conduct a cluster analysis of transport interchange nodes in at least one more large city to obtain a 15% sample.

References

1. D. N. Vlasov, Architecture and Construction of Moscow 2, 22-28 (2010)

2. D. N. Vlasov, N. V. Danilina, Industrial and Civil Engineering 9, 44-49 (2016)
3. I. A. Bakhirev, Urban Planning 2(42), 12-19 (2016)
4. D. N. Vlasov, Transport-transfer hubs: monograph (Publishing house of Moscow State University of Civil Engineering, Moscow, 2017)
5. "SP 42.13330.2011. Code of Practice. Urban Planning. Planning and development of urban and rural settlements. Revised edition of SNiP 2.07.01-89 *" (approved by Order of the Ministry of Regional Development of the Russian Federation of 28.12.2010 N 820)
6. SP 395.1325800.2018 Transport hubs. Design rules (approved by Order of the Ministry of Construction and Housing and Communal Services of the Russian Federation of 21.09.2018 № 609/pr), http://www.minstroyrf.ru/docs/17937/?sphrase_id=832396
7. N. V. Danilina, E3S Web of Conferences **11**, 00018 (2016)
8. Nikken Sekkei, Architecture and Urbanism **SI** (2013)
9. D. Vlasov, N. Danilina, Advanced Materials Research **869-870**, 201-204 (2014)
10. C. Ding, X. He, “K-means clustering using principal component analysis”. in *Proceedings of the Twenty-First International Conference on Machine Learning* (2004)
11. M. Jambu, *Hierarchical Cluster Analysis and Matching* (Finance and Statistics, Moscow, 1988)
12. D. N. Vlasov, Urban Planning 3(37), 11-19 (2015)