ATLANTIS PRESS

Advances in Social Science, Education and Humanities Research, volume 392

Ecological-Socio-Economic Systems: Models of Competition and Cooperation (ESES 2019)

Information and Communication Technologies in the Regions of Russia: Comparative Analysis

Elena Kochkina
Department of Chess Art and Computer Mathematics
Ural state University of Economics
Yekaterinburg, Russia
kem_d@mail.ru

Elena Radkovskaya
Department of Chess Art and Computer Mathematics
Ural state University of Economics
Yekaterinburg, Russia
rev_urgeu@mail.ru

Yurii Melnikov
Department of Chess Art and Computer Mathematics
Ural state University of Economics
Yekaterinburg, Russia
melniub@usue.ru

***Abstract*—The issue under study is pressing due to the high importance of information and communication technologies for the future of Russia that establishes new requirements for priorities in regional development strategy. The role and importance of information and communication technologies for the development of Russian economy are defined. The problem of imbalance in territorial economy structure which is typical for Russia as a multi-regional economic system leads to the need for a comparative analysis of regions in terms of development level and using information and communication technologies. The multivariance of studied process requires using special methods which include cluster analysis used by the authors; it is one of the methods of multivariate statistics and reflects the features of multivariance during the procedure of classifying objects. Clustering procedure was performed by the authors on the basis of hierarchical agglomerative methods using Euclidean distance between objects and farthest neighbor method in order to calculate the distance between clusters. In addition, the stability of obtained clustering results was checked using k-Means, Centroid, and Median methods. In order to get a dynamic clustering picture, we compared the results of dividing regions into clusters according to statistics from 2014 and 2017. Additionally, a comparative analysis of the regions was carried out according to the level of development of information and communication technologies based on the integrated comparison parameter developed by the authors.**

Keywords—information and communication technologies, region, cluster, hierarchical agglomerative methods, dendrogram, centroid, normalization.

I. Introduction

Analysis of global trends in social development shows that the basic capital in the 21st century will not be natural resources or even finances, but intellectual potential, therefore the future of Russia is in a great measure associated with the development of information and communication technologies.

Information and communication technologies are used in different fields, such as management, including public one, education, medicine, security, etc. As world experience shows, the increase in added value in the economy is largely due to intellectual activity and using communication and information technologies. Let us note that information and communication technologies quickly form a special economical segment which has an impact on the development of other industries [1-3].

The information sphere of Russia is becoming an integral part of global information space. Our country is developing a market for telecommunication equipment, information and telecommunication services, and a telecommunication infrastructure is being created [4]. The development of information sector which is focused on using local production base and scientific and technological potential will not only create new jobs and increase budget revenues but also satisfy the information needs of population, ensure international business and scientific contacts, and increase the investment attractiveness of our country. Implementation of new and more advanced information and communication technologies contribute to the liberalization of labor market and the organization of employee-employer relations via more flexible schemes [5]. The growth of labor productivity due to the wider using information and communication technologies will lead to an increase in the competitiveness of regions, and hence to higher profits and, consequently, to higher salary earned by employees, and, as a result, to the expansion of taxable base.

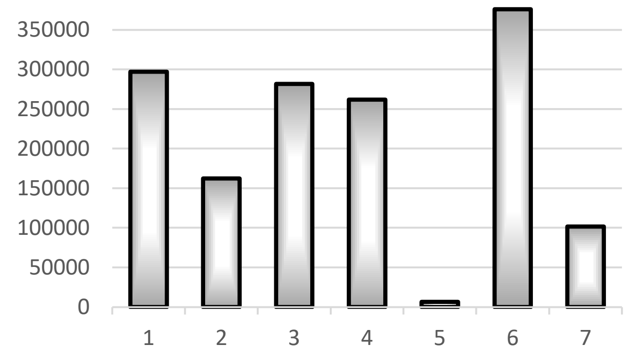
Unused scientific potential supported by modern systems of accelerated information exchange to the present moment serves as an actual basis for the country’s economic growth. For Russia, development and transportation of high technologies may be more effective than the transportation of many types of goods.

On June 28, 2017, the Government of the Russian Federation approved “Digital Economy of the Russian Federation” program [6] where the state policy for creating the necessary conditions for the development of digital economy was described. Implementation of this program will improve competitiveness in the global market of both individual sectors of the economy of the Russian Federation and the economy as a whole; but it requires significant financial, production and intellectual resources [7].

II. Materials and methods

Let us note that the regions of Russia differ significantly in their natural resource potential, structure and level of economic development, level of education, and other factors and, as a result, they differ in the effectiveness of using information and communication technologies. Uneven development of territories within a state is characteristic for all countries in the world; however, it is necessary to take this unevenness into account when making certain management and investment decisions [8-9].

Statistical data [10] show growing parameters characterizing the development of information and communication technologies in Russia. From 2005 to 2017, the number of organizations using Internet increased by 67%, the number of organizations having a website increased by 3.2 times, the number of personal computers per 100 employees increased by 2.17, the number of active subscribers for fixed and mobile broadband Internet access per 100 people increased by 72% and 67%, respectively. Positive changes are also characteristic for other parameters. Analysis of the structure of using special software tools shows that the most actively developing areas are the solution of organizational, managerial and economic problems and the making financial calculations in electronic form. Figure 1 shows the cost structure for ICT in Russia.



1. purchase of computer and office equipment;
2. purchase of telecommunication equipment;
3. purchase of software;
4. payments for telecommunication services
5. training of employees related to the development and use of ICT;
6. payment for the services of third-party organizations and ICT specialists (except for communication and training services);
7. other expenses.

Fig. 1. Cost structure for information and communication technologies in 2017.

Investigation of the level of development of information and communication technologies in the regions of the Russian Federation leads to the need to consider a large number of parameters. Development of theory, methodology and practice for statistical processing of analyzed data can move in two ways. One of them suggests the possibility of a probability interpretation of the processed data and the results obtained. The second way includes methods of cluster analysis, multivariate scaling, measurement theory, etc. which are not based on the probability nature of the processed data.

The authors have set the task of clustering the regions of Russia according to their level of development of information and communication technologies, i.e. to divide regions into homogeneous groups in selected attribute space [11-12].

III. Results and discussion

The term “cluster analysis” was first proposed by Trion. In the 1950s, the works of R. Lewis, E. Fix, and J. Hodges on hierarchical algorithms for cluster analysis were published. In 1963, R. Sokal and P. Sneath published their “Principles of Numerical Taxonomy”. One of the most popular classification tasks by combining a number of features is grouping of areas. Back in 1920, in his work “Connection between the Elements of Peasant Economy in 1917 and 1919”, B.S. Yastremsky investigated 34 characteristics of counties that influenced this connection [13]. Experiments of classifying cities by population are quite common. The famous Russian geographer and statistician V.P. Semenov-Tyan-Shansky was the first scientist who studied and classified the cities of our country from an economic point of view.

Cluster analysis allows investigating objects by combining them into groups with similar features. This method can be applied in a wide variety of fields [14-18]. The undoubted advantage of cluster analysis over other classification methods is that with its help it is possible to divide objects not by one parameter, but by a set of them. In addition, there are no significant restrictions on the type of objects under consideration; therefore, it is possible to study the initial data of an almost arbitrary nature.

In the general setting, the problem of clustering objects (regions of the Russian Federation) was to divide the analyzed set of objects R={Ri} which is statistically represented as a data matrix (parameters of using information and communication technologies) into a relatively small number of homogeneous in a certain sense groups, or classes. In order to formalize this problem, the regions of the Russian Federation were interpreted as points in the corresponding attribute space. Each region was characterized by a set (vector) of features showing the level of the development of information and communication technologies shown in official statistics. Parameters of thirteen types were chosen as coordinates of this vector. In preparing the statistical base for subsequent clustering, the authors excluded duplication of information associated with the use of correlated or non-informative features, i.e. parameters with little change during the transition from one object to another. Cluster analysis procedure involves calculation in some way of the distance between objects which in some cases cannot be correctly made due to the heterogeneity of units. This problem was solved by the authors using standardization of variables which allowed bringing variables to a single range of values.

Standardization of variables can optionally be supplemented by weighted coefficients; here it is reasonable to use expert estimations. The product of variable by the weight entered for it will allow estimating the distance between the objects taking into account the unequal weight of variables. As a result, expert opinion on the priority of the development of certain areas will be taken into account. However, this approach may be subjective.

Conventionally, distances from a given object to all other objects in the selected attribute space are used to describe objects in cluster analysis. The composition and number of clusters depends on the selected dividing criteria. For the clustering procedure, the authors used Euclidean distance which is the most popular metric for cluster analysis and emphasizes the contrast of defined clusters. Cluster analysis methods include many approaches and algorithms. For this purpose, the authors used hierarchical agglomerative methods on which basis a sequential combination of the initial elements and a corresponding decrease in the number of clusters were performed. As a measure of proximity characterizing the relative position of separate groups of objects, we used the distance calculated by farthest neighbor method, when an object is included into the cluster only if the similarity between the candidate for possible inclusion and any of the cluster elements is not less the some threshold value.

Next, the stability of the obtained clustering results was checked. For this purpose, some other clustering methods were used: k-Means, Centroid, and Median methods. Results obtained by different methods had a coincidence fraction exceeding 70% what allows using the results of clustering for further studies. It is natural to try to determine the comparative quality of different methods of dividing the studied set of objects into clusters. This procedure, as a rule, is carried out very arbitrarily and is based more on empirical and professionally intuitive considerations than on any strictly formalized system. The main criterion for the quality and validity of resulted dividing is a conceptual analysis of the results based on an understanding of the possible causal mechanisms for the implementation and isolation of the obtained groups of objects.

To get a dynamic picture of clustering, clustering of the regions of the Russian Federation was carried out according to the data of 2014 and 2017. Selection of groups of similar objects allows analyzing the characteristics of each group and making a behavior model for them; it is more rational than creating a common model for all objects. Dividing of regions into clusters was carried out sequentially, starting from the dividing into two classes using four abovementioned parameters of proximity between them. In our opinion, dividing into five clusters is optimal, since the subsequent dividing makes single objects as clusters, and it is not informative. Dividing the considered set of regions of the Russian Federation into groups of similar regions in the selected attribute space allows simplifying further data processing and making decisions using different analysis methods for each cluster.

In 2014, according to the results of clustering, three full clusters were defined that included 74 regions. Given the values of centroids for defined clusters, it can be argued that, taking into account the selected list of parameters, 12 regions were included in the best cluster, 30 regions were included in the worst cluster, and 32 regions form the middle cluster. The remaining two clusters include three Caucasian Republics and the Republic of Karelia. These regions can be considered as atypical objects that cannot be included in any of the defined clusters. The simplest explanation for this fact is that in these regions some parameters used for clustering have maximum values, while the others have minimum ones.

A convenient tool to quickly assess the success of dividing, the compactness of clusters, the presence of outlying cases, etc., are two-dimensional images of many points with indication of their group affiliation. Fig. 2 shows a two-dimensional image of defined clusters.

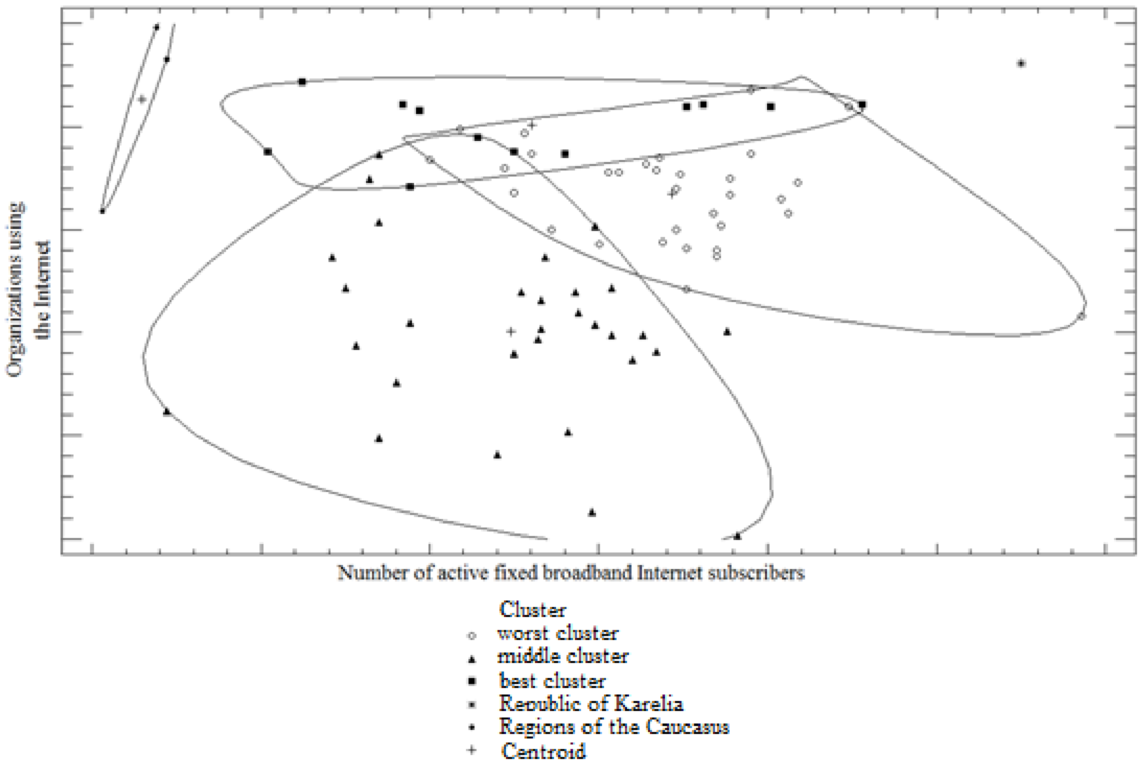


Fig. 2. Two-dimensional image of defined clusters

An important advantage of hierarchical clustering methods used by the authors is their visibility. The approach used allows building a dendrogram that is the result of hierarchical cluster analysis. This dendrogram is a nested grouping of objects which changes at different levels of hierarchy and describes the proximity of separate points and clusters to each other, and graphically represents the sequence of cluster unification (separation).

Clustering results according to 2017 data show the growth of most centroids and the alignment of regions in their using information and communication technologies.

When dividing regions into five clusters, as for 2014, two full clusters were defined that included 75 regions. Herewith, 37 regions were included in the best cluster, and 38 in the worst one. Regions of the Caucasus were included in the remaining three clusters, with each region being represented by a separate cluster. Analysis of the movement of regions between clusters shows that 23 regions moved from the middle cluster to the best, and 5 regions moved from the worst cluster to the best. Figure 3 shows the results of clustering of Russian regions in 2017 using selected attribute space.

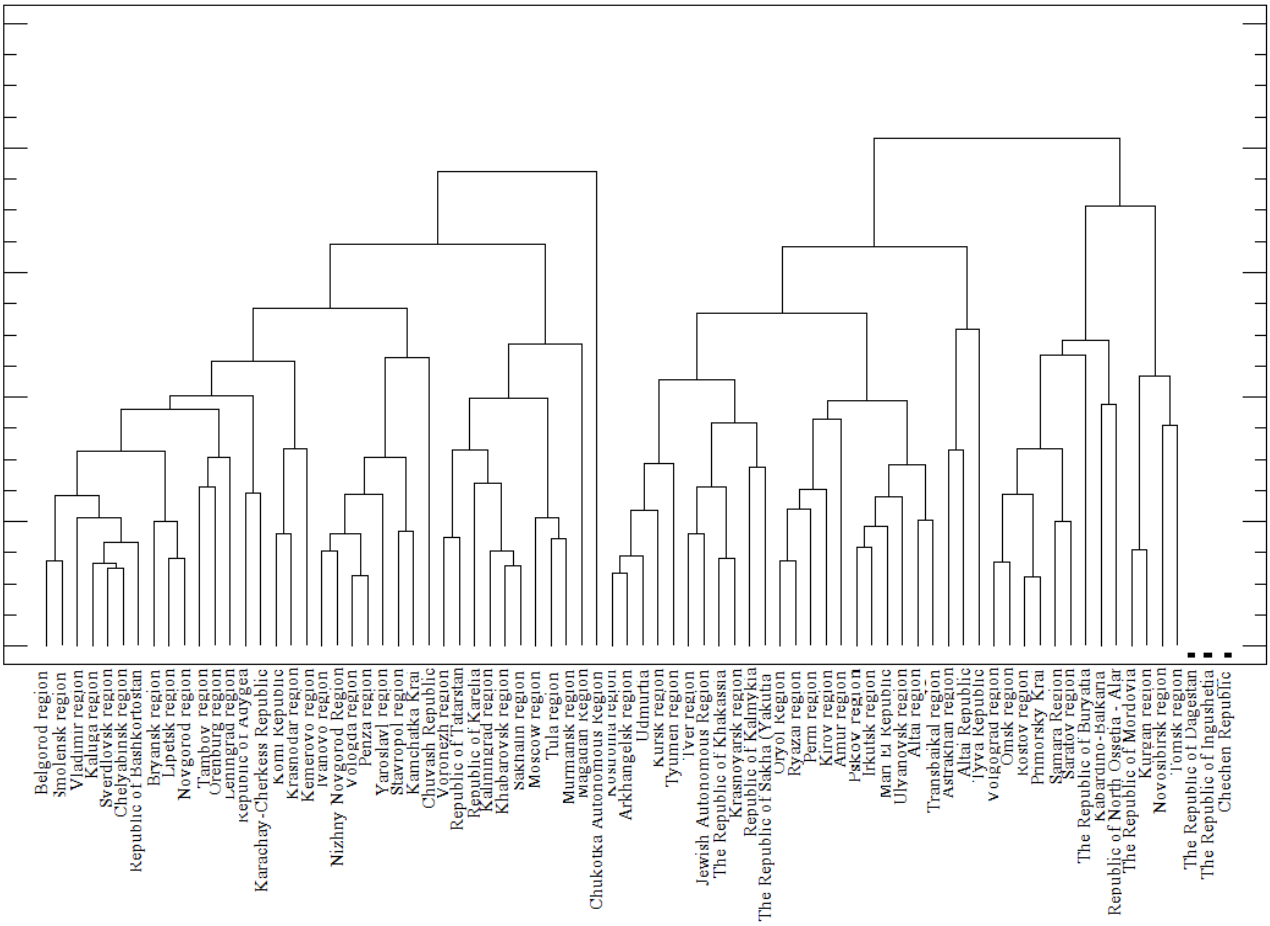


Fig. 3. Dendrogram of dividing the regions of the Russian Federation into clusters in 2017.

We should note that the values of centroids in the worst cluster in 2017 exceed the values of centroids even in the average cluster in 2014 what confirms the positive trends in the development of information and communication technologies in Russia.

Results of dividing the regions of Russia into clusters can be used for assessing the level of regional development in the field of development of information and communication technologies and redistribution of funds to support the regions taking into account reached levels.

Based on cluster analysis, the authors defined groups of regions that are similar in terms of the development of information and communication technologies in the selected parameter space. However, another comparative analysis is of additional interest – which defines the regions of Russia with fairly high level of the development of information and communication technologies, and the regions that lag behind in their development.

Based on the synthesis of the studied material, a methodological approach is proposed that allows assessing the regions depending on the set of parameters selected by the authors that reflect the development of information and communication technologies in the regions of Russia.

In order to assess the development of information and communication technologies, different parameters, both absolute and relative, can be used [19-21]. As a multivariate phenomenon, the development of information and communication technologies is characterized by a certain set of parameters published by official statistics. Combinations of different parameters characterizing the subject of analysis enable the analyst to obtain an individual assessment. Since there is almost always the possibility of different interpretation, obtaining a convincing assessment leads to the need to develop an integral parameter.

Integral parameter allows transforming the multivariance of studied phenomenon, simplifying and formalizing it, reducing many different criteria into a single assessment. The result allows placing, i.e. ranking the studied subjects (regions) on a given scale. Naturally, the more detailed is the set of used parameters, the more formally reliable is the calculating system and the results obtained have the greater confidence.

Iikt integrated comparison parameter was developed by the authors based on parameters used for clustering of the regions of Russia. In order to obtain comparable estimates, parameters of the development of information and communication technologies in the region (Pi values) were normalized to their maximum values for the regions selected for comparison (Si values), and the integral comparison parameter (Iikt) was calculated as the average value of normalized estimates.

Iikt integral comparison parameter is calculated by the formula (1):

$$I_{ikt} = \frac{\sum_{i=1}^{n} S_i}{n} \quad (1)$$

where n is the number of parameters used to characterize the development of information and communication technologies in the region.

To calculate Si values, the following formula (2) was used:

$$S_i = \frac{P_i}{\max\{P_i\}} \quad (2)$$

where Pi is the numerical values of the parameters used to characterize the development of information and communication technologies in the region; the maximum is taken for all analyzed regions. In general case, when it is necessary to take into account the negative development trends of studied process, the 1– Si difference is used as Si values.

For comparative analysis, unweighted estimates were used, i.e. it was assumed that all considered parameters are equally important for determining the value of integral parameter. A system of weighted estimates does not always provide a higher level of objectivity. For different regions, different parameters can have the largest share. In addition, a region is a complex dynamic system and the determination of the specific share of the studied parameters with an imperfect statistical base will be subjective, and, according to the authors, will give less effective results than the system of unweighted estimates.

Analysis of normalized estimates allows identifying regions leading in each of the selected areas. For example, the best electronic data exchange between own and external information systems is in the Astrakhan Region; the Novosibirsk Region leads in the number of personal computers per 100 employees; the best access to the Internet with a data transfer rate of 256 Kbps and higher is organized in the Republic of Tatarstan, etc.

IV. Conclusion

The approach used allowed defining the best and the worst regions of Russia not only for separate parameters, but also in combination for all selected parameters. The Republic of Tatarstan, the Republic of Karelia, Voronezh Region, Murmansk Region, Sakhalin Region, Stavropol Region, Kamchatka Region and Moscow Region are the regions with the higher level of development of information and communication technologies than in other regions of Russia taking into account the parameters selected for analysis. Regions with lower development of information and communication technologies according to the value of Iikt integral comparison parameter are the Kabardino-Balkarian Republic, the Chechen Republic, the Republic of Dagestan, the Republic of Buryatia, the Republic of Mordovia, and the Kurgan Region.

Obtained results make it possible to define regions with strong representation of any specific area of development and using information and communication technologies what makes it possible to find regions with experience that is worth to be shared to other regions, if it is possible.

Further improvement and expansion of the database of official statistical information will increase the number of parameters used for comparative analysis, and, as a result, more accurate results can be obtained.

References

[1] S.A. Rakhimova and R.F. Aminova, “Correlation of economy with information and communication technologies and its importance for today,” Science and Education, pp. 68-71, 2018 [V International Scientific and Practical Conference “Economics, Business, Innovations”, 2018]. (in russ.)

[2] Yu.S. Polozhentseva, “Management of differentiation of regional socio-economic systems on the basis of mobilization of internal development resources and attraction of external ones,” Izvestiya Yugo-Zapadnogo GU (Bulletin of Southwestern State University), No. 2(71), pp. 137-149, 2017. (in russ.)

[3] E.A. Stepanyan, “Digital economy as a factor of economic growth in Russia and its regions,” SNC, pp. 226-228, 2018 [International Research and Practice Conference “Theoretical, Methodological and Practical Problems of Innovative Ways to Improve the Energy Efficiency of Regional Industrial Complexes”, 2018]. (in russ.)

[4] G.N. Gasanov and D.A. Rabadanova, “Innovative technologies and problems of the digital economy of region,” Aktualnye voprosy sovremennoj ekonomiki v globalnom mire (Actual Issues of Present-day Economy in the Global World), No. 7, pp. 80-84, 2018. (in russ.)

[5] O.I. Shakhanova, “Influence of information and communication technologies on labor market,” Vestnik RGGU. Seriya: Ekonomika. Upravlenie. Pravo (Bulletin of the Russian State Humanitarian University. Series: Economics. Mangement. Law), No. 3, pp. 267-272, 2009. (in russ.)

[6] “Digital Economy of the Russian Federation” program approved by the Order of the Government of the Russian Federation No. 1632-rm as of July 28, 2017. (in russ.)

[7] O.A. Romanova, “Strategies for the socio-economic development of Russian regions in the context of the challenges of digital economy,” Izvestiya Uralskogo GEU (Bulletin of the Ural State Economic University), Vol. 19, No. 3, pp. 55-68, 2018. (in russ.)

[8] V.I. Tikhiy and S.S. Ivanov, “Socio-economic differentiation and its impact on the innovative development of regions during the transition to a digital economy,” Oryol State University, pp. 404-411, 2018 [International Research and Practice Conference “Formation of a New Paradigm for Economic Thinking of the 21st Century”, 2018]. (in russ.)

[9] R.V. Badylevich and E.A. Verbinenko, “Differentiation of the regions of the north in terms of their financial potential,” Intellekt. Innovatsii. Investitsii. (Intelligence. Innovations. Investments), No. 8, pp. 7-13, 2017. (in russ.)

[10] Regions of Russia. Socio-economic parameters, Collected papers Rosstat. Moscow, 2018. (in russ.)

[11] L.Kh. Gitis, Statistical classification and cluster analysis. Moscow: MSHU, 2003. (in russ.)

[12] I.M. Govorukhin, “Cluster analysis as a model of multivariant assessment of differentiation in regional parameters of innovative potential of national welfare,” Ekonomika i predprinimatelstvo (Economics and Entrepreneurship), No. 4-2(45), pp. 415-418, 2014. (in russ.)

[13] B.S. Yastremsky, “Connection between the elements of peasant farming in 1917 and 1919,” Vestnik statistiki (Bulletin of Statistics), pp. 19-21, 1920. (in russ.)

[14] V.N. Churilova, “Statistical approach to the analysis of business development based on cluster analysis methods,” Innovatsii i investitsii (Innovations and Investments), No. 1, pp. 39-44, 2016. (in russ.)

[15] A.A. Grigoriev, “Cluster analysis as a data processing tool in the analysis of information systems,” Izvestiya Rossijskogo ekonomicheskogo universiteta im. G.V. Plehanova (Bulletin of Plekhanov Russian University of Economics), No. 1(11), pp. 36-48, 2013. (in russ.)

[16] A.A. Khodkevich, “Analysis of the competitive environment in the market of communication services using cluster method,” Antrovita, pp. 148-158, 2018 [II International Research and Practice Conference “Social and Humanitarian Aspects of Human Well-Being”, 2018]. (in russ.)

[17] Yu.I. Treschevsky and V.M. Kruglyakova, “Cluster approach to the analysis of factors and conditions of investment activity in the regions of Russia,” Ekonomika i upravlenie (Economics and Management), No. 7(69), pp. 17-21, 2011. (in russ.)

[18] T.P. Kovshova, “On the cluster analysis of credit and investment policy of the second-tier banks of the Republic of Kazakhstan for 2013-2017,” Innovatsionnaya ekonomika: perspektivy razvitiya i

sovershenstvovaniya (Innovative Economics: Prospects for Development and Improvement), No. 3(29), pp. 65-73, 2018. (in russ.)

[19] S.G. Babich, “Index analysis of the differentiation of the regions of the Russian Federation according to basic innovation parameters,” Statistika i ekonomika (Statistics and Economics), No. 2, pp. 3-13, 2017. (in russ.)

[20] I.V. Vyakina, “Spatial differentiation of investment attractiveness of regions in the context of providing economic security,” Ekonomicheskij analiz: teoriya i praktika (Economic Analysis: Theory and Practice), Vol. 16, No. 8(467), pp. 1455-1472, 2017. (in russ.)

[21] R.V. Badylevich and E.A. Verbinenko, “Differentiation of the regions of the north in terms of their financial potential,” Intellekt. Innovatsii. Investitsii (Intelligence. Innovations. Investments), No. 8, pp. 7-13, 2017. (in russ.)