Mulesa O.,
Snytyuk V.,
Melnyk O.,
Nazarov V.

DEVELOPMENT OF ELEMENTS OF THE CONCEPT OF DETERMINING THE FUTURE DEMAND FOR MEDICAL SERVICES BASED ON THE RESULTS OF ANALYSIS OF DATA OF DIFFERENT NATURE

Planning and organizing the functioning of health care institutions is a priority area of activity of their founders. The purpose of such management activities is to ensure the timeliness, quality and completeness of medical services provided to the clients of the institution. At the same time, an important step is to predict the needs for medical services in future periods of time. Forecasting should be carried out taking into account the socio-demographic, medical and behavioral characteristics of persons – potential consumers of medical services and the characteristics of the population structure of the territory in which the medical institution operates. Thus, the object of research is the processes that arise during the analysis of operational and retrospective statistical, medical and social, expert and other data to determine the forecast values of the levels of demand for certain medical services. The results of the analysis should become the basis for making management decisions on planning and organizing the activities of health care institutions in future periods.

In the course of the research, a systematic approach, methods of mathematical modeling and other general scientific methods were used.

The research result is a developed method for forecasting the demand for medical services in future periods of time. The method consists in the implementation of four sequential stages of the analysis of the initial data. In this case, it becomes necessary to solve the problems of clustering, classification, identification and forecasting. The accuracy of the predicted values depends on the choice of methods and algorithms for solving the problems posed and on the completeness of the initial data. As a result of applying the method, it is possible to obtain:

– adivision of persons – potential consumers of medical services into groups in accordance with their socio-demographic portraits, medical data and behavioral characteristics;

– relationship between the number of educated groups and the demand for various medical services;

– predicted values of the number of groups, as well as the demand for medical services.

The results can serve as a basis for making managerial decisions on organizing the activities of medical institutions in future periods of time.

Keywords: *demand for medical services, clustering, structural and parametric identification, healthcare institutions.*

1. Introduction

Planning and organizing the work of healthcare institutions is one of the priority areas of activity of their founders, administrators and managers in the healthcare sector. The purpose of such management activities is to ensure the timeliness, quality and completeness of medical services provided by medical workers to clients of the institution. One of the key problems that arise in the process of planning the activities of medical institutions is forecasting the demand for medical services in future periods. Such forecasting should be performed taking into account the following factors:

– features of the activity and location of the healthcare institution;

– the staff and the competence of the employees of the institution;

– retrospective data characterizing the demand for medical services in past periods of time;

– medical and social factors affecting the emergence of the need for medical services in persons;

– social, economic, demographic and other features of the service area of the institution, and the like.

So, taking into account the need to analyze data of various nature, their coordination and the adoption of appropriate management decisions, it is relevant to develop a unified concept for determining the predicted values of demand for medical services in future periods.

2. The object of research and its technological audit

The object of research is the processes that arise during the analysis of operational and retrospective statistical,

medico-social, expert and other data to determine the forecast values of the levels of demand for individual medical services. The results of the analysis should become the basis for making management decisions on planning and organizing the activities of health care institutions in future periods.

3. The aim and objective of research

The aim of research is to analyze the stages of data processing to calculate the predicted values of the need for medical services in future periods of time. To achieve this aim, it is necessary to complete the following objectives:

1. To systematize the data on the basis of which it is necessary to analyze the needs for medical services.

2. To analyze the problems that arise in the process of forecasting the needs for medical services.

3. To develop a method for forecasting the demand for medical services in future periods of time.

4. Research of existing solution of the problem

The problem of implementing the function of managing a health care institution is covered in many scientific works. So, in [1] it is noted that a healthcare institution should be considered as a complex system. Despite the fact that the main goal of the activities of such institutions is to provide quality and timely medical services, it is important to improve the quality of such components of their management system as medical personnel, equipment, buildings, and the like. In [2], the authors note that the peculiarities of the medical services market include the presence of risks of diseases and the dependence of the demand for medical services on their supply. The study of the relationship between the socio-economic development of countries and trends in the demand for medical services in them is devoted to work [3]. It is shown that the range of medical services correlates with the level of socio-economic development of the country. The works [4, 5] are devoted to the analysis of the medical services market and the dependence of the demand for such services on the available supply. These studies argue that when analyzing the market for medical services and assessing the demand for them, it is necessary to take into account factors such as the structure of the region's population and the availability of medical services.

The work [6] is devoted to the problem of forecasting the demand for medical services, in particular, the need for emergency medical care. Forecasting is performed taking into account the geographic location of expected emergencies and is performed using artificial neural networks that are optimized using an evolutionary algorithm. Based on the results obtained, it is possible to determine the optimal locations of emergency medical vehicles. In [7], methods of clustering, forecasting, and regression analysis are demonstrated to assess the demand for medical services in different countries. The study contains examples of constructing regression models to determine the number of people who were discharged from hospitals in a calendar year based on historical data. Work [8] contains the results of predicting the need for vaccine for vaccination against tuberculosis based on forecasting fertility. Forecasting was carried out on the basis of the Box-Jenkins model based on time series analysis. The results obtained can be used to plan the procurement of the appropriate vaccine. The work [9] is devoted to the problem of assessing the demand for drugs. In the work, a model for predicting drug demand using an adaptive network based on a fuzzy inference system was created, based on the analysis of retrospective data on their sales volumes. In [10], the prediction of drug needs is carried out on the basis of two-stage clustering. This approach allows to achieve more accurate results, but it affects the training time of the neural network.

So, a large number of scientific studies are devoted to the analysis of the market for medical services and forecasting the demand for some of them. However, the question of developing a unified approach for predicting such demand for an individual healthcare institution based on both retrospective and medical and social data, and taking into account the population structure of the territory served by this institution, remains almost unsolved.

5. Methods of research

During the study, the following methods are used:
– systematic approach – when analyzing the problem of forecasting the demand for medical services in future periods of time;
– method of mathematical modeling – to formalize the tasks that arise in the process of analyzing retrospective and medico-social data and affect the forecast values of the corresponding demand indicators;
– method of analysis – in the study of the relationship between data on potential consumers of medical services and the demand for such services.

6. Research results

Assessment of the level of demand for medical services in future periods of time depends on the completeness and accuracy of retrospective data on potential and existing consumers of medical services, their socio-medical indicators. And also about the structure of the population of the territory served by this healthcare institution.

At the same time, data goes through several stages of analysis, among which are clustering, identification and forecasting. This process can be schematically represented as shown in Fig. 1.

So, in the process of analyzing data, in order to assess future needs for medical services, the following tasks arise:

1. The task of dividing into groups of persons – potential consumers of medical services – in accordance with the peculiarities of their medico-social and behavioral characteristics and the purpose of such a division.

Let a dataset be given in the form of a tuple:

$$<O, MD, ID, BD, SD>,$$

where O – set of persons; MD – set of values of medical indicators in given persons; ID – set of personal data such as age, gender, place of work, etc.; BD – behavioral characteristics of persons that may have an impact on the possibility of a need for medical services; SD – information about the services for which the demand is being investigated (for example, the demand for services in different periods of time, the individual characteristics of persons that affect the possibility of a need for a particular service, etc.).

It is necessary to split the set O into groups according to a given criterion.

The criterion for dividing persons – potential consumers – into groups may be the «similarity» of their socio-demographic or behavioral characteristics, which in turn affects their needs for medical services. In this case, the problem can be formulated as a clustering problem, the formulation of which is as follows:
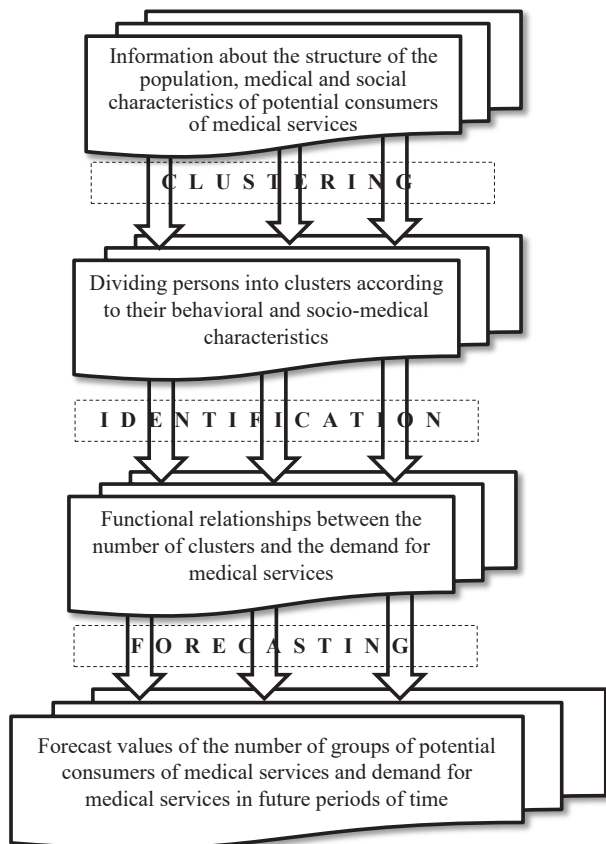

Fig. 1. Stages of data analysis in the process of assessing future demand for medical services

Let a set of persons be given:

$$O = \{O_1, O_2, ..., O_N\},$$

each of which is characterized by a variety of features:

$$X = \{X_1, X_2, ..., X_M\}.$$

Signs, in this case, can be elements of the socio-demographic portrait of a person, behavioral characteristics, components of the anamnesis, and the like.

Let a vector of feature values be given for each person $O_i \in O$, $i = \overline{1, N}$:

$$x^i = (x_{i1}, x_{i2}, ..., x_{iM}),$$

where x_{ij} – value of the feature $X_j \in X$, $j = \overline{1, M}$ for the corresponding person.

On the basis of the given data, it is necessary to construct a rule $s = D(x^{(i)})$ that maps a set X of possible values of features $\{X_1, X_2, ..., X_M\}$ to a set $\{1, 2, ..., K\}$ of cluster numbers, that is $s: X \to \{1, 2, ..., K\}$.

In other words, the task of clustering is to define the indicator variable s as follows:

$$s = \begin{cases} 1, \text{ if } x^i \in \Omega_1, \\ \dots \\ K, \text{ if } x^i \in \Omega_K, \end{cases}$$

where $\Omega_k, (k = \overline{1, K})$ – areas of space X corresponding to clusters.

There are frequent cases when the number of clusters K is not known in advance. Then K is also a parameter of the clustering model. The solution to the clustering problem in this case consists in varying the parameter K and searching for such a value that ensures the achievement of the specified optimality criterion.

To solve the formulated problem, it is possible to use clustering methods such as:
  – statistical [11, 12];
  – tree-like [13];
  – neural network [14, 15];
  – evolutionary [16];
  – expert [17];
  – methods of fuzzy clustering [18], etc.

Another criterion for dividing individuals into groups is their medical and social data. In this case, the partitioning problem can be formulated as a classification problem in the following setting [19].

Let's suppose a set of persons $O = \{O_1, O_2, ..., O_N\}$, for each of which the values of features from the set $X = \{X_1, X_2, ..., X_M\}$ are known.

That is, a given set of vectors:

$$x^i = (x_{i1}, x_{i2}, ..., x_{iM}),$$

where x_{ij} – value of the feature $X_j \in X$, $j = \overline{1, M}$ for the corresponding person. Each person belongs to one of the two specified classes A and B. Moreover, belonging to one of the classes means that the person either has a need for medical services (class A), or does not arise (class B).

It is necessary to set a rule according to which for some other person O', characterized by a vector $x' = (x'_1, x'_2, ..., x'_M)$ of the corresponding criteria from the set K, based on the data on the persons from the set O, it will be possible to make a decision about its assignment to one of the classes A or B.

To solve the formulated problem, it is possible to use methods for comparing feature values [20], neural network methods [21, 22], methods based on Wald's sequential analysis [19], etc.

The result of clustering will be the division of the set of persons O into clusters. Thus, even if as a result of the clustering done, K clusters were obtained: $G_1, G_2, ..., G_K$.

2. The task of establishing relationships between the capacity of formed clusters and the demand for medical services.

At the second stage of data analysis, it is necessary to estimate the size of each of the clusters in different periods of time. Let, as a result of the performed assessment, a table was formed (Table 1).

Based on the data obtained, it is necessary to perform the structural and parametric identification of the dependences of demand for services on the capacity of clusters. For this, it is possible to use fuzzy [23, 24], evolutionary [25] identification methods and other algorithms for estimating unknown parameters of systems [26].

Let's suppose, as a result of the identification done, a system of functions $P_j = \phi_j(Z_1, Z_2, ..., Z_K)$, $j = \overline{1, M}$ was built.

Table 1

Summary results for clusters and services at different time periods

Time period	Cluster capacity				Service demand			
	Z_1	Z_2	$\dots$	Z_K	P_1	P_2	$\dots$	P_M
T_1	z_{11}	z_{21}	$\dots$	z_{K1}	p_{11}	p_{21}	$\dots$	p_{M1}
T_2	z_{12}	z_{22}	$\dots$	z_{K2}	p_{12}	p_{22}	$\dots$	p_{M2}
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
T_t	z_{1t}	z_{2t}	$\dots$	z_{Kt}	p_{1t}	p_{2t}	$\dots$	p_{Mt}

Note: $T_1 < T_2 < \dots < T_t$ – specified time periods; z_{kl}, $k = \overline{1,K}$, $l = \overline{1,t}$ – results of assessing the number of the k-th cluster in the time period T_l; p_{jl}, $j = \overline{1,M}$, $l = \overline{1,t}$ – demand for the j-th service in the period of time T_l

3. Forecasting the values of the capacities of clusters in future periods of time.

The next step is to forecast the number of formed clusters in future periods of time. Forecasting should be carried out taking into account the retrospective data of Table 1. In this case, the setting of the forecasting problem is as follows:

Let $z_1, z_2, \dots, z_t$ be some discrete time series without length t gaps. The time series value is recorded at discrete points in time $i = \overline{1,t}$. The task of forecasting a time series is to build a decision rule F that allows to find an estimate of the value of a series in future periods of time, which is a forecast at a given point $t + \tau$ with a forecast step τ $(\tau \geq 1)$.

To solve the formulated problem, it is possible to use, for example, forecasting methods based on time series [27–29]. And also the forecasting method, taking into account the conclusions of experts [30].

Let's suppose that, as a result of the study, a forecast was made for the number of clusters $G_1, G_2, \dots, G_K$ in a given period of time T_{t+1}, $T_{t+1} > T_t$, as a result of which the predicted values $\tilde{Z}_{1,t+1}, \tilde{Z}_{2,t+1}, \dots, \tilde{Z}_{K,t+1}$ were obtained.

4. Determination of forecast values of demand for medical services in future periods of time.

Based on the results obtained at the previous stages, the forecast values of demand for medical services in future periods of time are calculated as follows:

$$p_{j,t+1} = \phi_j(\tilde{Z}_{1,t+1}, \tilde{Z}_{2,t+1}, \dots, \tilde{Z}_{K,t+1}).$$

The results obtained at the final stage can be used by health managers at the stages of planning and organizing the activities of the staff of medical institutions.

7. SWOT analysis of research results

Strengths. In the course of the study, a method was developed for assessing the demand for medical services in future periods of time, based on the available data on their potential consumers. The method consists in the implementation of four consecutive stages. As a result of the method, the division of potential consumers of medical services into groups will be obtained in accordance with the possibilities of their need for medical services in the future, as well as the calculated forecast values of demand for individual medical services.

Weaknesses. The accuracy of calculations in the course of applying the developed algorithm depends on the choice of methods and algorithms for solving problems arising at individual stages. Also, the completeness and reliability of data on potential consumers of medical services, socio-demographic and economic characteristics of the territory of service of a medical institution, and the like are important.

Opportunities. The results of the method, namely the estimated values of the demand for individual medical services in future periods of time, can be used by health managers, managers and founders of medical institutions in the process of organizing and planning the activities of the staff of healthcare institutions.

Threats. In the course of applying the developed method, it is important to choose methods for solving problems that arise at each separate stage. The accuracy of the obtained forecasts depends on the chosen methods of clustering, classification, identification and forecasting.

8. Conclusions

1. Data have been systematized, on the basis of which it is necessary to analyze the needs for medical services. It has been found that in the process of predicting future demand for medical services, it is necessary to analyze the following data sets:
   – information about the structure of the population of the territory served by the medical institution;
   – information about the socio-economic characteristics of the region or settlement;
   – set of values of medical indicators in persons – potential consumers of medical services;
   – socio-demographic portraits and information about the behavioral characteristics of persons – potential consumers of medical services;
   – information about the features of the provision of services, the demand for which is being investigated;
   – information about the demand for the investigated services in the past.

2. The analysis of the tasks that arise in the process of forecasting the needs for medical services has been carried out. It has been established that in the process of analyzing the input data, it is necessary to consistently solve the following tasks:
   – task of dividing individuals into groups in accordance with their socio-demographic and medical characteristics, as well as taking into account their behavioral characteristics that affect the emergence of needs for medical services;
   – task of identifying the dependences of the demand for medical services on the number of educated groups;
   – task of forecasting the number of groups in future periods of time based on retrospective data and taking into account the conclusions of experts.

3. A method for forecasting the demand for medical services in future periods has been developed. The method consists in performing four sequential stages of analysis of the input data. The result of applying the method is the forecast values of demand for medical services in future periods of time.

References

1. Hrabovskyi, V. A., Klymenko, P. M. (2014). Systemnyi pidkhid do upravlinnia zakladamy okhorony zdorovia. *Visnyk Natsionalnoi akademii derzhavnoho upravlinnia pry Prezydentovi Ukrainy, 3,* 136–142.

2. Danko, V. V. (2017). Upravlinnia zakladamy okhorony zdorovia v suchasnykh umovakh: teoretychnyi aspekt. *Visnyk KhNAU. Seriia: Ekonomichni nauky, 4,* 225–233.

3. Haponova, E. O. (2017). Suchasni tendentsii funktsionuvannia svitovoho rynku medychnykh posluh. *Visnyk Kharkivskoho natsionalnoho universytetu imeni VN Karazina. Seriia: Mizhnarodni vidnosyny. Ekonomika. Krainoznavstvo. Turyzm, 6,* 20–24.

4. Brownlee, S., Chalkidou, K., Doust, J., Elshaug, A. G., Glasziou, P., Heath, I. et. al. (2017). Evidence for overuse of medical services around the world. *The Lancet, 390 (10090),* 156–168. doi: http://doi.org/10.1016/s0140-6736(16)32585-5

5. Glasziou, P., Straus, S., Brownlee, S., Trevena, L., Dans, L., Guyatt, G. et. al. (2017). Evidence for underuse of effective medical services around the world. *The Lancet, 390 (10090),* 169–177. doi: http://doi.org/10.1016/s0140-6736(16)30946-1

6. Grekousis, G., Liu, Y. (2019). Where will the next emergency event occur? Predicting ambulance demand in emergency medical services using artificial intelligence. *Computers, Environment and Urban Systems, 76,* 110–122. doi: http://doi.org/10.1016/j.compenvurbsys.2019.04.006

7. Maltseva, S., Prokofyeva, E., Zaitsev, R. (2017). The Demand for the Healthcare Services: the Opportunities of Big Data in Predicting Patient Flow. *International Conference Information Systems 2017 Special Interest Group on Big Data Proceedings,* 5.

8. Kim, K.-W., Li, G., Park, S.-T., Ko, M.-H. (2016). A Study on Birth Prediction and BCG Vaccine Demand Prediction using ARIMA Analysis. *Indian Journal of Science and Technology, 9 (24).* doi: http://doi.org/10.17485/ijst/2016/v9i24/96151

9. Husein, A. M., Simarmata, A. M. (2019). Drug Demand Prediction Model Using Adaptive Neuro Fuzzy Inference System (ANFIS). *SinkrOn, 4 (1),* 136–142. doi: http://doi.org/10.33395/sinkron.v4i1.10238

10. Husein, A. M., Harahap, M., Aisyah, S., Purba, W., Muhazir, A. (2018). The implementation of two stages clustering (k-means clustering and adaptive neuro fuzzy inference system) for prediction of medicine need based on medical data. *Journal of Physics: Conference Series, 978,* 012019. doi: http://doi.org/10.1088/1742-6596/978/1/012019

11. Baturkin, S. A., Baturkina, E. Iu., Zimenko, V. V., Siginov, I. V. (2012) Statisticheskie algoritmy klasterizatsii dannykh v adaptivnykh obuchaiuschikh sistemakh. *Vestnik RGRTUb, 31 (1),* 82–85.

12. Rokach, L., Maimon, O. (2005). *Clustering methods. Data mining and knowledge discovery handbook.* Boston: Springer, 321–352. doi: http://doi.org/10.1007/0-387-25465-x_15

13. Amanuma, S., Kurematsu, M., Fujita, H. (2012). An Idea of Improvement Decision Tree Learning Using Cluster Analysis. *SoMeT,* 351–358. doi: http://doi.org/10.3233/978-1-61499-125-0-351

14. Kohonen, T. (1988). *Self-organization and associative memory.* New-York: Springer Verlag, 312. doi: http://doi.org/10.1007/978-3-662-00784-6

15. Škrjanc, I., Iglesias, J. A., Sanchis, A., Leite, D., Lughofer, E., Gomide, F. (2019). Evolving fuzzy and neuro-fuzzy approaches in clustering, regression, identification, and classification: A Survey. *Information Sciences, 490,* 344–368. doi: http://doi.org/10.1016/j.ins.2019.03.060

16. Mulesa, O., Snytyuk, V., Myronyuk, I. (2016). Forming the clusters of labour migrants by the degree of risk of hiv infection. *Eastern-European Journal of Enterprise Technologies, 3 (4 (81)),* 50–55. doi: http://doi.org/10.15587/1729-4061.2016.71203

17. De Koninck, P., Nelissen, K., Baesens, B., vanden Broucke, S., Snoeck, M., De Weerdt, J. (2017). An approach for incorporating expert knowledge in trace clustering. *International Conference on Advanced Information Systems Engineering.* Springer: Cham, 561–576. doi: http://doi.org/10.1007/978-3-319-59536-8_35

18. Mulesa, O. Yu. (2015). Adaptation of fuzzy c-means method for determination the structure of social groups. *Technology Audit and Production Reserves, 2 (2 (22)),* 73–76. doi: http://doi.org/10.15587/2312-8372.2015.41014

19. Mulesa, O. Yu., Snytiuk, V. Ye., Herzanych, S. O. (2019). Metod nechitkoi klasyfikatsii na osnovi poslidovnoho analizu valda. *Automation of technological and business processes, 11 (4),* 35–42. doi: http://doi.org/10.15673/atbp.v11i4.1597

20. Kornoushenko, E. K. (2017). Algoritm klassifikatsii putem parnogo sravneniia priznakov. *Avtomatika i telemekhanika, 11,* 151–166.

21. Ruff, L., Vandermeulen, R., Goernitz, N., Deecke, L., Siddiqui, S. A., Binder, A. et. al. (2018). Deep one-class classification. *International conference on machine learning,* 4393–4402.

22. Geifman, Y., El-Yaniv, R. (2017). Selective classification for deep neural networks. *Advances in neural information processing systems, 30,* 4878–4887.

23. Shtovba, S. D. (2003). Identifikatsiia nelineinykh zavisimostei s pomoschiu nechetkogo logicheskogo vyvoda v sisteme MATLAB. *Exponenta Pro: Matematika v prilozheniiakh, 2,* 9–15.

24. Zaichenko, Yu. P. (2007). Nechetkyi metod hrupovoho ucheta arhumentov pry neopredelennikh vkhodnikh dannikh. *Systemni doslidzhennia ta informatsiini tekhnolohii, 3,* 100–112.

25. Mulesa, O. Yu. (2016). Development of evolutionary methods of the structural and parametric identification for tabular dependencies. *Technology audit and production reserves, 4 (2 (30)),* 13–19. doi: http://doi.org/10.15587/2312-8372.2016.74482

26. Haber, R., Unbehauen, H. (1990). Structure identification of nonlinear dynamic systems – A survey on input/output approaches. *Automatica, 26 (4),* 651–677. doi: http://doi.org/10.1016/0005-1098(90)90044-i

27. Aizenberg, I., Sheremetov, L., Villa-Vargas, L., Martinez-Muñoz, J. (2016). Multilayer Neural Network with Multi-Valued Neurons in time series forecasting of oil production. *Neurocomputing, 175,* 980–989. doi: http://doi.org/10.1016/j.neucom.2015.06.092

28. Deb, C., Zhang, F., Yang, J., Lee, S. E., Shah, K. W. (2017). A review on time series forecasting techniques for building energy consumption. *Renewable and Sustainable Energy Reviews, 74,* 902–924. doi: http://doi.org/10.1016/j.rser.2017.02.085

29. Geche, F, Batyuk, A., Mulesa, O., Vashkeba, M. (2015). Development of effective time series forecasting model. *International Journal of Advanced Research in Computer Engineering & Technology, 4 (12),* 4377–4386.

30. Mulesa, O., Geche, F. (2016). Designing fuzzy expert methods of numeric evaluation of an object for the problems of forecasting. *Eastern-European Journal of Enterprise Technologies, 3 (4 (81)),* 37–43. doi: http://doi.org/10.15587/1729-4061.2016.70515

Mulesa Oksana, *PhD, Associate Professor, Department of Cybernetics and Applied Mathematics, State Higher Educational Institution «Uzhhorod National University», Uzhhorod, Ukraine, ORCID: http://orcid.org/0000-0002-6117-5846, e-mail: mulesa.oksana@gmail.com*

Snytyuk Vitaliy, *Doctor of Technical Sciences, Professor, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine, ORCID: http://orcid.org/0000-0002-9954-8767, e-mail: snytyuk@gmail.com*

Melnyk Olena, *PhD, Associate Professor, Department of Software Systems, State Higher Educational Institution «Uzhhorod National University», Uzhhorod, Ukraine, ORCID: http://orcid.org/0000-0001-7340-8451, e-mail: olena.melnyk@uzhnu.edu.ua*

Nazarov Volodymyr, *Postgraduate Student, Department of Cybernetics and Applied Mathematics, State Higher Educational Institution «Uzhhorod National University», Uzhhorod, Ukraine, ORCID: http://orcid.org/0000-0002-0906-7020, e-mail: vonaz2713@gmail.com*