https://orcid.org/0000-0003-4076-3830
https://orcid.org/0000-0002-5523-6120

UDC 004.942:930.24

N.O. KHYMYTSIA[*], S.V. HOLUB[**]

INTELLECTUAL ANALYSIS OF THE RESULTS OF THE CLIOMETRIC MONITORING

[*]Lviv Polytechnic National University, Lviv, Ukraine
[**]Cherkasy State Technological University, Cherkasy, Ukraine

Анотація. *Для підвищення адекватності результатів кліодинамічних досліджень використовується інформаційна технологія багаторівневого інтелектуального моніторингу в поєднанні з експертними оцінками результатів моделювання. Результати інтерпретувались у рамках класичних концепцій історичної науки. Це сприяє залученню до аналізу джерел систем штучного інтелекту (баз знань, експертних систем, когнітивних комп'ютерних моделей розуміння тексту, фреймових систем), в яких моделюються знання істориків. При формуванні масиву чисельних характеристик певних історичних періодів на сьогодні спостерігається тенденція до збільшення питомої ваги історичних джерел, створених колективним способом. Застосування технології інтелектуального моніторингу дозволяє автоматизувати процеси обробки історичних даних і підвищити ефективність наукових досліджень та адекватність висновків. Представлені результати застосування технології багаторівневого інтелектуального моніторингу для вирішення однієї з задач кліометрії. Розв'язувалась задача визначення подібності історичних періодів. Перелік ознак, що описують історичні періоди, визначено експертним шляхом. Тривалість часового проміжку вибрано за результатами обробки статистичних даних. Чисельні характеристики вибраних ознак формувались впродовж однакових часових проміжків і утворювали вектор ознак для кожного з історичних періодів. Вектори ознак історичних періодів піддавались кластеризації за результатами моделювання. Метод синтезу моделей вибирався окремо для формування кожного кластера за результатами випробування кожного алгоритму синтезу моделей моніторингової інтелектуальної системи. В більшості випадків моделі-кластеризатори будувались за багаторядним алгоритмом GMDH. Описано процеси формування масиву вхідних даних, синтезу моделей та визначення впливу ознак. Експериментально підтверджено доцільність використання нового методу визначення подібності історичних періодів.*

Ключові слова: *кліометрія, МГУА, моніторинг, кластеризація, інтелектуальний аналіз даних.*

Анотация. *Для повышения адекватности результатов клиодинамических исследований используется информационная технология многоуровневого интеллектуального мониторинга в сочетании с экспертными оценками результатов моделирования. Результаты интерпретировались в рамках классических концепций исторической науки. Это способствует привлечению к анализу источников систем искусственного интеллекта (баз знаний, экспертных систем, когнитивных компьютерных моделей понимания текста, фреймовых систем), в которых моделируются знания историков. При формировании массива многочисленных характеристик определенных исторических периодов сегодня наблюдается тенденция к увеличению удельного веса исторических источников, созданных коллективным способом. Применение технологии интеллектуального мониторинга позволяет автоматизировать процессы обработки исторических данных и повысить эффективность научных исследований и адекватность выводов. Представлены результаты применения технологии многоуровневого интеллектуального мониторинга для решения одной из задач клиометрии. Решалась задача определения сходства исторических периодов. Перечень признаков, описывающих исторические периоды, определен экспертным путем. Продолжительность временного промежутка выбрана по результатам обработки статистических данных. Численные характеристики выбранных признаков формировались в течение одинаковых временных промежутков и образовывали вектор признаков для каждого исторического периода. Векторы признаков исторических периодов подвергались кластеризации по результатам моделирования. Метод синтеза моделей выбирался отдельно для формирования каждого кластера по результатам испытания каждого алгоритма синтеза моделей мониторинговой интеллектуальной системы. В большинстве случаев модели-кластеризаторы строились по многорядному алгоритму GMDH. Описаны процессы формирования массива входных данных, синтеза моделей и определения влияния признаков.*

Экспериментально подтверждена целесообразность использования нового метода определения сходства исторических периодов.
Ключевые слова: *клиометрия, МГУА, мониторинг, кластеризация, интеллектуальный анализ данных.*

Abstract. *In order to increase the adequacy of the results of clinical studies, information technology of multilevel intelligent monitoring is used in conjunction with expert estimates of simulation results. The results were interpreted in the framework of the classical concepts of historical science. This contributes to attracting to the analysis of sources of systems of artificial intelligence (knowledge bases, expert systems, cognitive computer models of understanding of the text, frame systems), in which the knowledge of historians is modeled. When forming an array of numerical characteristics of certain historical periods, there is a tendency towards increasing the proportion of historical sources created by the collective method. The application of intelligent monitoring technology allows automating the processing of historical data and improving the efficiency of research and the adequacy of the findings. The results of the application of the technology of multilevel intelligent monitoring for solving one of the problems of climmetry are presented. The problem of determining the similarity of historical periods was solved. The list of features describing historical periods is determined expertly. The length of the time interval is selected based on the results of processing the statistics. Numerical characteristics of the selected signs were formed during the same time intervals and formed a vector of signs for each of the historical periods. Vectors of signs of historical periods were subjected to clustering by the results of modeling. The model synthesis method was selected separately for the formation of each cluster based on the results of testing each of the algorithms for the synthesis of models of the monitoring intellectual system. In most cases, model clusterizers were built on a multi-row GMDH algorithm. The processes of the formation of the input data array, the synthesis of models and the determination of the effect of the signs are described. The feasibility of using a new method for determining the similarity of historical periods has been experimentally confirmed.*
Keywords: *Cliometric, GMDH, Monitoring, Clustering, Data Mining.*

DOI: 10.34121/1028-9763-2019-4-87-92

1. Introduction

Intelligent monitoring – is an information technology providing knowledge of decision-making processes by organizing untested spots, processing and transforming their results. Application of the process and transformation of monitoring results, allows to obtain massive numerical characteristics monitoring objects and build a base of model knowledge in the form of a hierarchical combination of models that solve one of the standard tasks: grouping of objects, forecasting, identification.

In accordance with the methodology for the establishment of multi-level monitoring information systems (MLM) [1], the processing of monitoring results is carried out through the consistent application of statistical processing methods, inductive and other methods for synthesizing models that solve local data transformation tasks at each of the levels of monitoring. Stratification of the structure of the multi-level monitoring information systems allow solving tasks of identifying functional dependencies, classifying, forecasting and some others by constructing multilayered models [2]. The hierarchical combination of these models form a global functional dependence (GFD) [3].

For the synthesis of multilayered models that form the structure of GFD, the basic algorithms of the GMDH [4], genetic and other evolutionary methods of synthesis of models, neural networks with various topologies, and hybrid algorithms, which are formed by combining several methods into a single process of designing the algorithm of model synthesis is used.

Today MLM is used in many subject areas. In particular, information socioecological [5], sociogeygenic [6], medical monitoring [7] have already been created. Monitoring systems are used in public administration [8], in pedagogy [9, 10].

The variety of tools of historical science [11, 12] opens up wide opportunities for using information technology of multilevel monitoring based on inductive modeling methods.

Application of such methods allows realizing the prognostic function of historical science, which aims at predicting ways of development of historical processes, various variants of historical events. The prognostic function answers the question: what the historical reality will be and when certain events will take place. The result of this function is, above all, the hypothesis, historical forecasting, which is based on objective scientific data.

2. Analysis of research and publications

In practice, mathematical modeling in historical science has been used for more than 40 years. I. D. Kovalchenko, the founder of the Klimometric School in the USSR and the co-chairman of the International Commission on Quantitative History (INTERQUANT) widely applied mathematical modeling [13]. The scientist believed that this method makes it possible to analyze the historical process that could be implemented, but by virtue of some reasons has not happened. An American economist, one of the founders of the cliometriya and Nobel Laureate of the Year 1993, Robert William Fogel applied counter-factual modeling and proved that small innovations in industry contribute more to its evolution than large-scale technological discoveries [14]. The scientist conducted a fundamental study on what would be the US transport system if the railways were not invented.

At the beginning of the twenty-first century, especially during periods of social chaos, when the predictability of the course of history is significantly weakened and the possibility of unexpected variants of development increases, the relevance of multi-parametric modeling as a study of historical alternatives increases. Within this approach, the principle of synergetics as an interdisciplinary science that deals with the study of the processes of self-organization and the emergence, support of the stability and decomposition of structures (systems) of different nature is realized. Synergetics allows the historian to determine in which conditions small influences cause a complex system of large-scale changes and avalanche-like processes.

A new direction of modeling in historical research is the study of nonlinear processes. professor L. Borodkin introduces the direction – the cliodinamik that study the models of unstable historical processes [15, 16]. Under the guidance of L. Borodkin, the project was released which is devoted to the analysis of alternatives to historical development in 1929, when the so-called "big change" began in the USSR (the transition from the new economic policy to the beginning of the course on industrialization and collectivization).

The purpose of this study is to develop a method of cliodinamik monitoring, which will provide the identification of similar historical periods by clustering the vectors of their numerical features.

3. Using the template

The method of cliodinamical monitoring combines the use of processes of clustering historical periods with their numerical features and expert justification of clustering results. It involves the following steps:

1. The list of features that are significant for making decisions based on the results of monitoring of historical processes is determined.

2. An array of numerical characteristics is formed for meaningful signs. The point of observation is numerical characteristics within one year.

3. The clustering of monitoring points by modeling results is carried out.

4. The hypothesis about the similarity of historical processes that took place over the years that formed separate clusters is proposed.

5. Examination of hypotheses is carried out by their expert justification using historical research methods. If an expert way has succeeded in substantiating the given hypotheses, the historical periods included in one cluster are considered to be similar. If this does not work – the

following research is conducted and conclusions about the similarity of historical periods are not announced.

Clusterization of vectors of signs of historical periods was carried out on the results of the simulation [16]. At the initial stage, it is necessary to solve the problem of clustering historical periods, which are presented in the form of vectors of their numerical features. Moreover, each historical period is represented by one vector of numerical characteristics of these attributes. The list and limits of historical periods, as well as the list of characteristics, are determined by expert means. The results are presented as an array of data (1):

$$\begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} & y_{11} & y_{12} & \dots & y_{1m} \\ x_{21} & x_{22} & \dots & x_{2n} & y_{21} & y_{22} & \dots & y_{2m} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ x_{k1} & x_{l2} & \dots & x_{kn} & y_{k1} & y_{k2} & \dots & y_{km} \end{pmatrix}, \quad (1)$$

where xij is the j-th independent index of the i-th historical period, yij is the j-th dependent index of the i-th historical period, k is the number of vectors; n – the number of indicators, m – the number of historical periods.

It is necessary to develop a method of clustering historical periods in the form of a function

$$a:(X_i, y_i) \to r, \quad (2)$$

where r is the cluster number that forms the set of clusters R.

The process of clusterization for each vector of the signs of the historical period sets the cluster number $r \in R$. In this case, the power of the set R is not known in advance.

Socio-economic and demographic development of Ukraine was investigated during 1998-2012 based on the technology of multi-level monitoring.

The dependence of the gross domestic product of the country from the indicators presented in Table 1 was modeled. For synthesis of models, multi-row algorithm of GMDH [4] was used. A set of models was synthesized. After the tests, models were selected that met the criteria of accuracy, adequacy and stability, and the influence of the indicators from Table 1 that were included in the structure of these models was estimated. The impact of indicators was determined by the weighting factors calculated by the formula [1]:

$$W_i = \frac{F'_{x_i}}{\sum_{i=1}^{n} F'_{x_i}}, \quad (3)$$

where F'_{x_i} is a partial derivative of the model on its i-th variable,

n – number of indicators table. 1 that entered into the structure of the model.

The results of the study of these models are presented in Table 1.

Table 1 – Influence of signs

Indicator	Cluster 1 (2001–2012)	Cluster 2 (1999–2009)	Cluster 3 (1998–2002)
Time of observation	99,88	0,00	0,00
Income of the population, billion UAH.	0,00	20,60	90,66
Average monthly salary, UAH	0,06	19,50	0,00
The average monthly retirement pension, UAH	0,00	1,72	0,00
The average size of assistance to low-income groups of population per capita, natural	0,00	22,86	0,00

Continuation of table 1

Economically inactive population of able-bodied age, thousand	0,002	0,00	0,00
Fixed assets*, billion UAH	0,00	8,96	0,00
Budget (consolidated), UAH billions Revenues	0,00	1,09	0,00
Budget (consolidated), UAH billions Expenditures	0,00	3,05	0,00
Investments in fixed assets*, UAH billions	$1 \cdot 10^{-6}$	20,75	0,00
Imports of goods and services, billion dollars. USA	0,06	0,00	0,00
Direct foreign investments in Ukraine (at the end of the year), mln. USA	0,00	1,46	9,34

Expert estimation of historical periods is as follows:

1 cluster of 2001–2012: decrease of investments into fixed capital; strengthening the processes of industrialization; crisis in mechanical engineering, chemical industry; growth in agriculture, trade, transport; according to sources of financing, bank loans and other loans play an essential role; negative balance of trade balance; lack of investment; the super-profits of the export industries fell into the pockets of the oligarchic circles in full; increase in the amount of social assistance; uncontrolled migration of the population; growth of demographic load.

2 clusters of 1999–2009: the economic downturn, resulting in GDP declining to 40,8 %; violation of macroeconomic equilibrium, correlation between wage level and gross income / mixed income in the structure of GDP, did not have the character of a stable trend; industry was the leading sector of the real economy, while retaining the largest share in its structure (in 2008, it accounted for 46,1 % of the gross output of goods and services and 31,3 % of gross value added); more than 2/3 of the total industrial output accounted for in the industry producing raw materials and energy resources; the share of products of social orientation is 1/5 of the total volume of industrial production. The light industry almost disappeared (1999: 1,6 %, in 2008 – 0,9, in August 2009 – 0,8%); the machine-building industry even added in the rate of development during 1999–2007, but its share in the structure of industry through the crisis has fallen to 10,3%, which was 3–4 times lower than the level of developed countries.

3 clusters of 1998–2002: growth of real incomes of citizens was 5–6 times higher than GDP; there was a rise in wages; the real GDP growth was 4,1 – 9%; recorded increase in production in industry; the growth rate of export of goods exceeded the growth rate of imports; increased competitiveness of Ukrainian goods; currency offer exceeded demand**.** Thus, we proved the closeness of the points included in each cluster. This means that during these years historical events in Ukraine were caused by close processes and research objects.

The results of expert interpretation make it possible to conclude on the usefulness of the process of clustering historical periods using the new method.

4. Conclusions

It is presented a new methodology for historical research, which combines methods of intellectual monitoring and historical research. The efficiency of combining processes of intellectual analysis of historical data with expert interpretation of its results was experimentally proved by the example of solving the problem of classification of historical periods. The simulation results obtained by the new method should be used to identify the same types of periods of Ukrainian history (according to the mechanisms of influence on the economy that creates GDP).

REFERENCES

1. Holub S.V. Multi-level modeling in environmental monitoring technologies. Chernivtsi: B. Khmelnytsky National University of Chernivtsi, 2007. 220 p.
2. Holub S.V., Burley I.V. Multilayer data transformation in information systems of multi-level monitoring of fire safety. Collection of scientific works of Kharkiv University of Air Forces. 2014. Vol. 1, N 38. P. 246–251.
3. Holub S.V., Rudnitsky V.N., Dendarenko V.Yu., Pivneva S.V. Formation of the structure of the global function in information systems of multi-level monitoring of fire safety. *The Science Vector of the Togliatti State University*. 2012. Vol. 4, N 22. P. 41–45.
4. Ivakhnenko A.G. Inductive method of self-organization of models of complex systems. K.: Naukova Dumka, 1981. 296 p.
5. Holub S.V., Kolos P.O. Features of creation of information systems for crisis socioecological monitoring. *Systems of information processing.* 2011. Vol. 4, N 94. P. 273–277.
6. Golub S.V., Nemchenko V.Yu. Adaptive formation of duplicate levels in the structure of hierarchical systems of multi-level sociohygienic monitoring. *Inductive modeling of complex systems. Collection of scientific works.* 2011. Vol. 1. P. 41–48.
7. Holub S.V., Dzholos V.M. Using of heuristic monitoring system for forecasting postinfarction complications. *Visnyk of Zhytomyr State Technological University. Engineering.* 2004. Vol. 4, N 31, P. 109–114.
8. Holub S.V., Khymytsia N. O. The display of consolidated information of regional economic performance in the structure of multi-level models. *Visnyk of Volodymyr Dahl East Ukrainian National University*. 2012. Vol 8, N 179, Part 1. P. 122–128.
9. Holub S.V., Semenenko L.M. Information modeling as the basis of the didactic system of training young scientists. *Problems of general and pedagogical psychology, Collection of scientific works of the Institute of Psychology named after G.S. Kostyuk APS of Ukraine*. 2005. Vol. VII. P. 96–100.
10. Holub S.V., Semenenko L.M. Application of heuristic observation systems as a modeling pedagogical tool. *Vesnyk Kharkiv National Automobile and Road University. Collection of scientific works.* Kharkiv: KhNADU, 2005. P. 10–12.
11. Williamson S. Cliometrics history in the United States (translation A.N. Polevaia). *Economic History. Review.* 1996. Vol. 1. P. 75–107.
12. Khymytsia N., Lisina S., Morushko O., Zhezhnych P. Analysis of computer-based methods for processing historical information. *Advances in Intelligent Systems and Computing: Selected Papers from the International Conference on Computer Science and Information Technologies*. Springer International Publishing, Lviv, Ukraine. CSIT 2017. Vol. 1. P. 365–367.
13. Kovalchenko I.D. Methods of historical research. 2nd ed. M., 2003. 454 p.
14. Fogel R. Railroads and American Economic Growth: Essays in Econometric History. Baltimore, Johns Hopkins Press, 1964. 296 p.
15. Borodkin L.I. Historical information science at the bifurcation point: the movement to Historical Information Science. Circle of ideas: algorithms and technologies of historical Informatics: *Proc. of the IX conference of the Association "History and Computer"* / ed. L.I. Borodkin, V.N. Vladimirov. Moscow-Barnaul, 2005. P. 7–21.
16. Golub S.V., Burley I.V. Increasing the efficiency of clustering by the results of simulation in the information technology of operational fire monitoring. *Information Processing Systems*. 2014. Iss. 2, N 118. P. 253–257.

Стаття надійшла до редакції 08.08.2019