Security of Infocommunication Systems and Internet of Things

2023 Vol 1, No 1

https://doi.org/10.31861/sisiot2023.1.01004

Received 05 May 2023; revised 22 May 2023; accepted 05 June 2023; published 30 June 2023

Full-Stack Development of an Intelligent System for the Development of Population Migration

Dmytro Uhryn*, Yuriy Ushenko, Oleksandr Galochkin and Artur Hostiuk

Computer Science Department, Yuriy Fedkovych Chernivtsi National University, Chernivtsi, Ukraine

*Corresponding author (E-mail: d.ugryn@chnu.edu.ua)

ABSTRACT A generalised model of population migration is proposed. On the basis of the model of population migration, the article develops models of: a set of directions of population flows, which are formed on internal and external flows of the State; duration of migration, which is determined by its nature in time, including permanent or irreversible duration of migration, movement for a relatively short time, annual movement of people and pendulum duration of migration; type and form of migration. A model of indicators of actual migration (resettlement) that can characterise the overall level of mobility of the population of the territories, the scale, structure, directions and effectiveness of migration flows for a given period is developed and their groups are divided. It is proposed that the results of population migration should be presented in the form of a number of absolute and relative indicators for the purpose of further regression analysis of data, namely, those who arrived for permanent residence from other settlements; those who left for permanent residence to other settlements; migration balance or mechanical growth. Inter-rayon relations are characterised by the strength of migration flows. To obtain the results of migration, we take into account the strength of migration flows, which depend on the population of the territories between which the exchange takes place and on their location. The result of this exchange is expressed in the migration balance or by means of efficiency coefficients of migration ties. The intensity of migration exchange, independent of the population size of both the areas of origin and the places of settlement, is determined by the intensity coefficients of migration ties. The types of migration intensity coefficients are formed depending on the properties, namely the intensity coefficients of arrival (immigration), departure (emigration), reverse migration, and net migration. The intelligent geographic information system implements the lightgbm algorithm for population migration forecasting, which is a decision tree with gradient reinforcement. For data analysis, the migration forecasting system includes regression analysis and neural networks and is capable of predicting international migration or migration between different countries.

KEYWORDS geodata, population migration, full-stack development, intelligent system, decision-making.

I. INTRODUCTION

Recently, demand in making decisions related to population migration has increased, as this process affects many aspects of social and economic development. Migration can have both positive and negative effects on the economy, demographics, cultural diversity, national security, and other aspects of people's lives.

The analysis of migration processes is a result for increasing the scientific level of planning the development of individual regions and industries. The study of patterns and features of migration allows for the development of more accurate and effective strategies for the development of the territory [1-3, 5], to ensure a more rational distribution of resources [4-5] and to solve problems with employment [6-7] and social integration of migrants [8-9].

The study of migration processes is an important element of the development of the policy of migration and integration of migrants. Taking into account the dynamics of migrations and their properties can help in determining optimal strategies for attracting and integrating migrants into society, ensuring the rights and freedoms of migrants, preventing discrimination and exclusion of migrants from society.

Therefore, the study of population migration has important theoretical and practical significance, as it allows to understand the processes taking place in society and to develop effective improvement and integration strategies.

II. RESEARCH OBJECTIVE

Based on the developed and researched models and artificial neural network, to develop a full-stack intelligent system for the development of population migration.

III. MODELS FOR THE DEVELOPMENT OF AN INTELLIGENT SYSTEM

The main source of information on population migration is state statistics, which contain current migration records and population census materials [10, 11]. In order to obtain additional data, random surveys are conducted, the purpose of which is to find out the reasons for displacement [12, 13]. The study of population migration is based on a system of indicators, each of which reveals a certain aspect of the phenomenon. Modern conditions also affect migration processes, including climate change, military conflicts, diseases and pandemics, nomadism, tourist and work travel. The generalized model of population migration is proposed to be presented in the following expression:

$$MN =< NP, TM, TypM, FM >, \quad (1)$$

where NP – set of directions of population flows, TM – duration of migration, $TypM$ – type of migration, FM – form of migration.

The set of population flow directions is given by the model:

$$NP =< p_v, p_z >, \quad (2)$$

where p_v – internal flows of the state that occur between different units of a separate administrative division, $p_z = \{em, im, reem, rep\}$ – external flows of the state, carried out between states, regions of the planet, continents. Accordingly, the elements of the set em – emigration, which characterizes the movement of citizens of a particular state abroad; im – immigration, as the entry of foreign citizens into the state; $reem$ – remigration, which characterizes the return of former or valid citizens of the country to their homeland; rep – repatriation, which means the return of a population that was forcibly removed.

Duration of migration TM can be determined by its nature in terms of time and is suggested to be set in the form of a set:

$$TM =< TM_p, TM_t, TM_s, TM_m >, \quad (3)$$

TM_p – permanent or irreversible durations of migration, characterizing the movement of the population to a permanent or long-term place of residence; TM_t – movement for a short time; TM_s – annul movements of people related to work, recreation at resorts, nomadic migration; TM_m – pendulum duration of migration, which determines regular daily or weekly trips to work or study outside the place of residence.

Type of migration $TypM$ may include:

$$TypM =< M_{st}, M_{ssh}, M_{sm}, M_{ms}, M_{kp}, M_{td}, M_m, M_{pt} >, \quad (4)$$

M_{st} – migration of seasonal tourism; M_{ssh} – migration of seasonal agriculture production; M_{sm} – migration from rural areas to cities, which occurs in developing countries in the process of urbanization; M_{ms} – migration from cities to rural areas, characterized of developed countries (ruralization); M_{kp} – migration of nomads and pilgrimages; M_{td} – temporal and long-term migration; M_m – pendulum migration; M_{pt} – border and transit migration.

Form of migration FM is given by set:

$$FM =< FM_K, FM_n, FM_d, FM_p, FM_l >, \quad (5)$$

FM_K – managed or socially organized form organization; FM_n – unorganized form of organization; FM_d – voluntary form of organization; FM_p – forced form of organization; FM_l – legal or illegal form of organization.

Migration indicators can indicate the general level of population mobility in the territory, as well, as the scale, structure, directions and effectiveness of migration flows for a specific period [1, 3]. One of the most accurate indicators of the level of migration mobility is the determination of the number of migrations of persons of a certain age during their lifetime of in the population as a whole [2].

Indicators of actual migration (immigrants) can be divided into three groups:

$$P_{FM} =< P_Z, P_S, P_{MO}, P_{pow}, P_{intens}, P_{effectivm} >, \quad (6)$$

where general P_Z, which characterize the migration processes common to the territory; special (structural) P_S, which characterize the migration of specific socio-demographic groups, as well as indicators of interdistrict (international) exchange P_{MO}, that characterize migration relations between specific territories of migration exchange. They include power indicators P_{pow}, intensity P_{intens} and effectiveness of migration in fig.1.

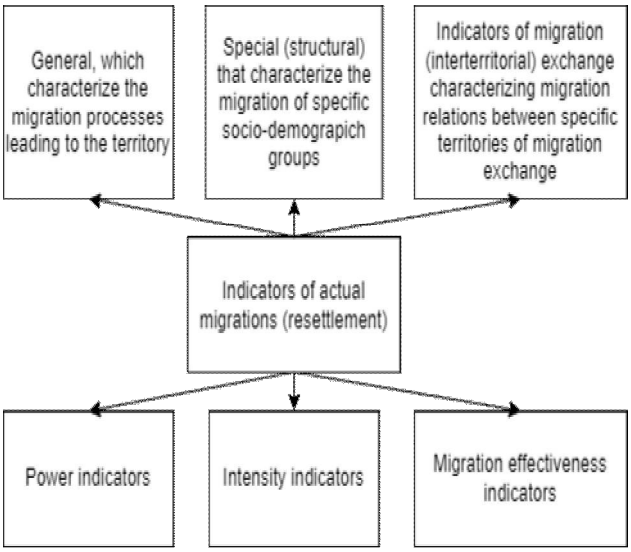

FIG. 1. Indicators of actual migration (resettlement)

The results of population migration are characterized by a number of absolute and relative indicators: a) the number of people arriving for permanent residence from other settlements (P_m); б) the number of those who left for permanent residence in other settlements (V_m); в) net migration or mechanical growth $S_m = P_m - V_m$.

The main characteristic of interdistrict relations is the power of migration flows [1, 2], which depend on the population of the territories between which the exchange takes place, as well as their geographical location. The result of the exchange is reflected through the balance of migration or coefficients of effectiveness of migration relations (CEMR). The intensity of migration exchange remains independent of the population of the regions of origin and destination, using the coefficients of the intensity of migration relations (CIMR). CIMR allows to take into account the impact of the population count only in the form of the specific weight of arrivals or departures

in the general flow of arrivals or departures to the total population of the territory that maintains migration relation with this region.

CIMR on arrival is calculated using the formula:

$$K_{ij} = \frac{P_{ij}}{P_{ji}} : \frac{S_j}{S_i} \times 1000, \quad (7)$$

CIMR on departure is calculated using the formula :

$$K_{ij} = \frac{V_{ij}}{V_{ji}} : \frac{S_j}{S_i} \times 1000, \quad (8)$$

where P_{ij} – the number of arrivals in the regions j from region i; V_{ji} – the number of departures from the region j in region i; S_i – average population of the region i for period; S_j – the total number of exit regions, which the region i supports migration relations.

The intensity of migration is defined as a statistical characteristic of population mobility, which shows the frequency of migration in different territorial and demographic groups. This characteristic is expressed using the coefficient of intensity of migration (CIM) [4], which can be calculated for the total population or for its structural elements, such as gender, age, nationality, etc.

We will describe four types of coefficients of intensity of migration, depending on its properties:

1. Coefficient of intensity of arrival (immigration) - it is a quantitative characteristic of persons who arrived in a certain region from another place. This ratio is calculated by dividing the number of arrivals by the average population in that region. This shows the population that was exposed to the risk of receiving migrants.

$$I_j = \frac{P_{ij}}{S_i} \times 1000. \quad (9)$$

2. Coefficient of intensity of departure (emigration) - it is a quantitative characteristic of persons who left a certain region (or country) for a certain period. This coefficient is calculated by dividing the number of departures by the average population size for the same period. This shows the population that was exposed to the risk of leaving.

$$I_j = \frac{B_{ij}}{S_i} \times 1000, \quad (10)$$

3. The coefficient of the intensity of return migration can be defined as the sum of the number of people who arrived in the region and those who left it, divided by the average population in this region.

$$I_j = \frac{P_{ij} + B_{ij}}{S_j} \times 1000, \quad (11)$$

4. The coefficient of the intensity of net migration is calculated as the difference between the number of people who arrived in the region and those who left it, divided by the average population of the region.

$$I_j = \frac{P_{ij} - B_{ij}}{S_j} \times 1000, \quad (12)$$

where P_{ij} – the total number of arrivals in the region j from all regions; B_{ij} – the total number of departures from the region j; S_j – average population of the region j for the period under study.

Coefficients of migration intensity make it possible to determine the level of population mobility in a specific territory, to compare the levels of population mobility in different regions according to their rank and size, to identify the dynamics of migration and to predict changes in the future. However, the use of this indicator is limited, as it depends not only on the intensity of migration relations, but also on the migration capacity and capacity of regions. Therefore, it cannot be used with absolute accuracy, and additional data must be used for analysis.

Regarding the efficiency of migration, this is another important characteristic that reflects the ratio between the number of departures and arrivals. The coefficient of effectiveness of migration is usually determined per 1000 arrivals and can be general or interdistrict. This indicator makes it possible to assess the effectiveness of the migration movement and its impact on the population of the regions.

The general migration coefficient is calculated according to the formula [2]:

$$m = \frac{M}{P \cdot k}; \quad (13)$$

where M – the number of migrants; P – average population; k – constant that is equal to a value from 1 to 100 (the result is expressed as s percentage) or up to 1000 (the result is expressed in ppm).

It is necessary to calculate special coefficients for different components of migration flows, since the characteristics of migrants directly depend on their composition. For example, age coefficients of migration intensity are defined as the ratio of the number of migrants of a certain age to the average number of the population of that age in the region of departure or arrival. Similarly, other coefficients can be calculated, such as the ratio of the number of men and women to the total population, the percentage of the able-bodied and disabled population, etc.

IV. IMPLEMENTATION OF INTELECTUAL SYSTEM

The developed web-based system uses geographic information system (GIS) elements such as geocoding, regression analysis, and neural networks to analyze migration data and explore patterns and drivers of migration. The GeoDjango web framework is used to work with geographic data, which provides access to spatial queries and provides efficient storage and processing of geospatial data [14-15]. The JavaScript library Leaflet [16-18] is used to display data on the map, which allows customize the appearance and functionality of the map, as well as aggregate data by regions and

countries. The lightgbm algorithm is used to predict immigration and emigration based on country and population age, which is a gradient-boosted decision tree-type machine learning model. The model is trained on a large set of immigration and emigration data and population age data for each country and can accurately predict future immigration and emigration patterns based on the input data provided.

The dataset comprises data regarding the count of immigrants, emigrants, and net migration, encompassing all nations across the globe. It is compiled from a wide range of reliable sources, including official government resources such as national statistical agencies, international organizations like the United Nations (UN), and other pertinent data repositories. The rigorous collection process ensures that the dataset is based on credible and up-to-date information, providing a comprehensive and accurate representation of global immigration and emigration patterns.

The first step in using lightgbm is to prepare the input data, including selecting appropriate features and variables for prediction, such as country and population age. Then it is necessary to divide the data into sets for training and testing. Next, the lightgbm algorithm is used to train a decision tree model, using a gradient boosting algorithm to optimize the accuracy of the model.

To analyze migration data, the system uses regression analysis and neural networks [19-21]. Regression analysis helps model the relationship between migration and demographic, economic and environmental variables to identify key drivers of migration and predict future migration patterns. Neural networks (fig.2) are used to model complex relationships between variables and improve the accuracy of predictions over time. Once the model is trained, it can be used to predict migration based on new data, such as country and population age, using pre-trained data.

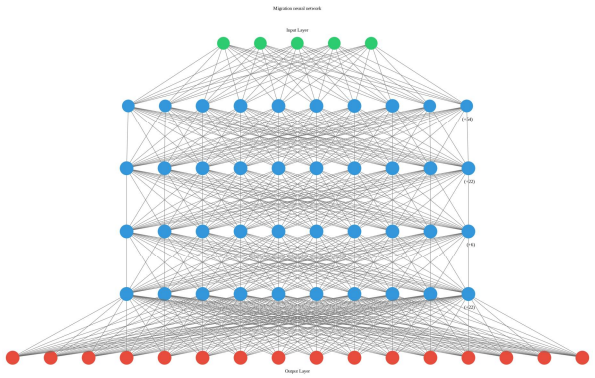

FIG. 2. Visualization of neural network of internal population migration

The lightgbm algorithm is particularly effective for predicting immigration and emigration due to its ability to handle large and complex datasets. It can also automatically select the most important features for prediction and handle missing data.

However, there are limitations and potential problems with the accuracy of predictions. Inaccurate or incomplete data can lead to inaccurate predictions, limiting the usefulness of the system. Therefore, constant updating and monitoring of the model is necessary to ensure its relevance and effectiveness over time.

Despite the inherent complexity and unpredictability of human migration patterns, the migration forecasting system can successfully forecast internal migration within a country, including migration between regions accounting for factors that influence people's decisions (fig.3).

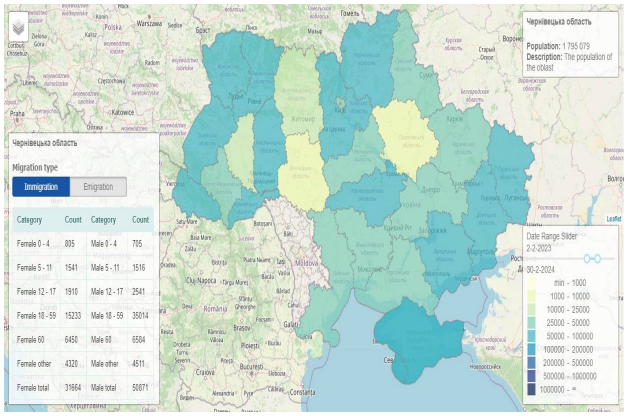

FIG. 3. Interactive map on internal population migration

One of the methods of internal migration analysis is to study the age characteristics of migration between regions. This approach makes it possible to determine trends in the movement of the population by age category, for example, young people migrating to cities for the purpose of getting an education or getting acquainted with new employment opportunities. For example, by analyzing data on the number of migrants in different age groups between different regions during a certain period of time, it is possible to predict future trends in migration.

In addition, the migration forecasting system can be used to forecast international migration or migration between different countries (fig.4).

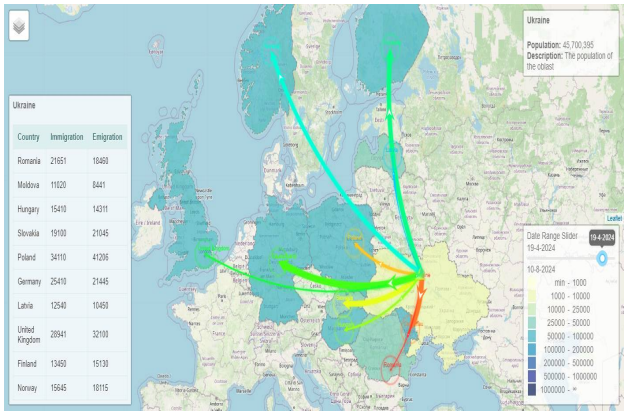

FIG. 4. Interactive map on international population migration

V. CONCLUSION

The article deals with the construction of models for describing population migration. The means influencing migration for the development of an intelligent system for predicting population migration are identified. In general, a migration forecasting system based on GIS, regression analysis, neural networks, and the lightgbm algorithm is a powerful tool for predicting migration patterns within and between countries. It has the potential to provide valuable

insights into the patterns and drivers of migration and help governments and organizations plan and prepare for future demographic changes.

AUTHOR CONTRIBUTIONS

D.U., Y.U. – conceptualization, methodology, investigation; D. U., O. G. – writing-original draft preparation; D. U., A. H. – software, validation, formal analysis, investigation, resources; D. U., Y. U. – writing-original draft preparation, visualization, supervision writing-review and editing.

COMPETING INTERESTS

The authors declare no competing interests.

REFERENCES

[1] UKRSTAT. (2014, Jan). Migration movement of the population [Online]. Available: https://ukrstat.gov.ua/operativ/operativ2014/ds/mr/mr_u/mr0114_u.html

[2] I.S. Blahun, L.I. Sysak, and O.O. Soltysik, Modeling of sustainable development of the region. Ivano-Frankivsk: Vasyl Stefanyk Precarpathian National University, 2006.

[3] E.M. Libanova (Ed.), O.V. Pozniak, O.B. Makarova, V.G. Sarioglo, and L.G. Tkachenko, External labour migration of population. Kyiv: Ukrainian Center for Social Reforms; State Committee of Statistics of Ukraine, 2009.

[4] O.V. Poznyak and T.O. Hnatiuk, "Human development in Ukraine: the historical dimension of the transformation of state social policy" in State Migration Policy and its Impact on the Population Structure, Kyiv: Ptoukha Institute for Demography and Social Studies of the National Academy of Sciences of Ukraine, 2014, pp. 92-108.

[5] V.G. Saryoglo, Great data' as a source of information and tools for official statistics: potential, problems, perspectives. Statistics of Ukraine, 2016, vol. 4, pp. 12-19.

[6] O. Malinovska, Labour migration of Ukrainian population: what to expect at the nearest future?. [Online]. Available: http://migraciya.com.ua/news/migrantworkers/ua-labour-migration-ukrainewhat-to-expectinthe-near-future-part-1

[7] U. Sadova (Ed.), "Regional Migration Policy and Mechanisms for its Implementation," National Academy of Sciences of Ukraine, Institute for Regional Research, Lviv, 2011.

[8] U. Sadova (Ed.), O. Grinkevich, L. Semiv, R. Tesliuk, M. Bil, V. Bidak, O. Rindzak, O. Levytska, and O. Mulska, "Ukrainian migration in the context of the global and national challenges of the 21st century," Lviv, 2019.

[9] O. Ovchinnikova, Models of Regional Migration of the Population. Intelekt XXI, 2017, vol. 4, pp. 27-32.

[10] M. Romanyuk, Overseas Labor Migration and Remembrance in the Context of National Security of Ukraine. Rehionalna ekonomika – Regional Economy, 2016, vol. 4, pp. 22-30.

[11] A. Sadovenko, L. Maslovska, V. Sereda, and T. Tymochko Eds. Sustainable Development of Society: A Textbook, 2nd ed. K., 2011.

[12] G. Abel, Quantifying global international migration flows. Science, 2014. [Online]. Available: https://science.sciencemag.org/content/343/6178/1520

[13] M. Bell, International migration and data. Polish Academy of Sciences, 2015. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1002/psp.1848

[14] M. Diener, Python Geospatial Analysis Cookbook: Over 60 recipes to work with topology, overlays, indoor routing, and web application analysis with Python. Packt Publishing Ltd, 2015.

[15] Django documentation, Geospatial fields. 2023. [Online]. Available: https://docs.djangoproject.com/en/3.2/ref/contrib/gis/fields/

[16] LightGBM. (2023). LightGBM Documentation. [Online]. Available: https://lightgbm.readthedocs.io/en/latest

[17] UNHCR, Global Trends: Forced Displacement in 2021. [Online]. Available: https://www.unhcr.org/refugee-statistics

[18] Leaflet, Leaflet - a JavaScript library for interactive maps. [Online]. Available: https://leafletjs.com

[19] A. J. Kardika, A. Khoirunnita, Salman, Saharuddin, and I. Muliana, Development Web-GIS of Commodity Information System for Agriculture, Establishment and Forestry in Marangkayu District. International Journal of Education and Management Engineering (IJEME), vol. 12, no. 5, pp. 1-8, 2022. doi: 10.5815/ijeme.2022.05.01.

Dmytro Uhryn

Graduated from Yuriy Fedkovych Chernivtsi National University, Chernivtsi. Currently, he is Doctor of Technical Sciences, associate professor Yuriy Fedkovych Chernivtsi National University. He has currently published more than 140 publications. His research interests are data mining, information technologies for decision support, swarm intelligence systems, industry-specific geographic information systems.

Yuriy Ushenko

Prof., Computer Science Department, Chernivtsi National University, Chernivtsi, Ukraine. Research Interests: Data Mining and Analysis, Computer Vision and Pattern Recognition, Optics & Photonics, Biophysics.

Oleksandr Galochkin

Graduated from Yuriy Fedkovych Chernivtsi National University, Chernivtsi. Currently, he is Doctor of Technical Sciences, associate professor Yuriy Fedkovych Chernivtsi National University. He has currently published more than 140 publications. His research interests are data mining, information technologies for decision support, swarm intelligence systems, industry-specific geographic information systems.

Artur Hostiuk

Student, Computer Science Department, Chernivtsi National University, Chernivtsi, Ukraine. Has publication in student scientific conference. Research Interests: Data Mining, Artificial Intelligence and Analysis.

FULL-STACK розробка інтелектуальної системи розвитку міграції населення

Дмитро Угрин*, Юрій Ушенко, Олександр Галочкін та Артур Гостюк

Кафедра комп'ютерних наук/Відділ комп'ютерних технологій/Навчально-науковий інститут фізико-технічних та комп'ютерних наук, Чернівецький національний університет імені Юрія Федьковича, Чернівці, Україна

* Автор-кореспондент (Електронна адреса: d.ugryn@chnu.edu.ua)

АНОТАЦІЯ Запропоновано узагальнену модель міграції населення. На основі моделі міграції населення розроблено моделі: множини напрямів потоків населення, що формуються на внутрішніх і зовнішніх потоках держави; тривалості міграції, що визначається за її характером по часу, яка включає постійні або безповоротні тривалості міграції, переміщення на відносно короткий час, щорічні переміщення людей та маятникова тривалість міграції; типу і форми міграції. Розроблено модель показників фактичної міграції (переселень), що можуть характеризувати загальний рівень рухливості населення територій, масштаби, структуру, напрями і результативність міграційних потоків за той або інший період та поділено їх групи. Запропоновано результати міграції населення подавати у вигляді ряду абсолютних і відносних показників з метою подальшого проведення регресійного аналізу даних, а саме прибулих на постійне проживання з інших населених пунктів; вибулих на постійне проживання в інші населені пункти; сальдо міграції або механічний приріст. Міжрайонні зв'язки характеризуються потужністю міграційних потоків. Для отримання результатів міграції враховано потужність міграційних потоків, що залежать від чисельності населення територій, між якими відбувається обмін, та від місця їх розташування. Результат цього обміну виражається в сальдо міграції або за допомогою коефіцієнтів результативності міграційних зв'язків. Інтенсивність міграційного обміну, незалежна від чисельності населення як районів виходу, так і місць вселення, визначається за допомогою коефіцієнтів інтенсивності міграційних зв'язків Сформовано типи коефіцієнтів інтенсивності міграції в залежності від властивостей, а саме коефіцієнти інтенсивності: прибуття (імміграції), вибуття (еміграції), міграції по звороту, чистої міграції. В інтелектуальній геоінформаційній системі реалізовано алгоритм lightgbm для прогнозування міграції населення, що є деревом рішень із посиленням градієнта. Для аналізу даних система прогнозування міграції включає регресійний аналіз і нейронні мережі та здатна прогнозувати міжнародну міграцію або міграцію між різними країнами.

КЛЮЧОВІ СЛОВА геодані, міграція населення, Full-stack розробка, інтелектуальна система, прийняття рішень.