Construction Optimizing of Search and Data Processing Scenario Based on a Graphical Model

Oleksandr V. Koval[1,2], Valeriy O. Kuzminykh[1], Iryna I. Husyeva[1], Beibei Xu[3] and Shiwei Zhu[3]

[1] National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, 03056, Ukraine
[2] Institute for Information Recording of National Academy of Sciences of Ukraine, M. Shpaka str. 2, Kyiv, 03113, Ukraine
[3] Information Institute, Qilu University of Technology (Shandong Academic of Sciences), Jinan, 250316, China

Abstract

The article discusses the method of constructing a graph describing the scenario of analytical activity based on the Cartesian product operation of the knowledge sources graph of the subject area and the graph of data processing procedures, which makes it possible to construct a connected graph that reflects the possible scenarios of the analyst's search for data according to the task. The weighted evaluations of the efficiency of the optimal path search on the graph are determined, which enables the analyst to choose the most relevant to the task paths on the graph and discover the most efficient data collection scenario. Quantitative methods of assessing the stability of the analytical process by numerical indicators are proposed. As an advantage of this method, one can point out the ability to quantitatively assess the stability of the analytical process of the current structure of the analyst's sequence of actions, which must be performed to achieve the goal of the task. Based on this type of evaluation, certain recommendations can be formulated regarding the list of actions of the analytical process, which will allow for achieving a defined goal in accordance with a specific task for conducting the analysis. At the same time, the construction of scenarios of the analyst's activity based on the graph of the scenario description can be especially efficient when the information in individual sources is partially or completely duplicated but has a different level of actualization. In several cases related to the construction of information and analytical systems and decision-making systems, the approach shown in the article can be used in the initial processing and consolidation of data from disparate sources.

Keywords

data collection, ontology, graph analysis, scenario, branched network

1. Introduction

Today, there are many methods of selecting the required scenarios regarding quality and efficiency from a set of possible or acceptable scenarios. These scenarios are built to analyze the composition and nature of the factors influencing the scenario planning process. When planning in practice, there is a large number of situations that can be divided by the number of influencing factors. These factors can be the basis for deciding the methods and procedures for forming scenarios of the analyst's actions.[1]

When describing ontologies to solve the problems of choosing the optimal scenarios in terms of quality and efficiency from a set of acceptable scenarios, a method of description based on graphs is applied. As a rule, hierarchical graphs provide the most complete options for displaying possible paths. At the same time, estimates of various costs specific to a specific case of conducting analytical actions can be used. These can be the costs of processing, transforming, and consolidating information in the nodes of the graph and the costs of transition between nodes. At the same time, total scores can be determined for a certain path on the graph and any part of this path. In turn, this makes it possible to compare these total estimates to determine the optimal ways, from the point of view of efficiency, for further inclusion in the scenario [2].

ITS-2023: Information Technologies and Security, November 30, 2023, Kyiv, Ukraine
avkovalgm@gmail.com (O.V.Koval); vakuz0202@gmail.com (V.O.Kuzminykh); iguseva@yahoo.com (I.I.Husyeva); xubeibei1987@163.com (Beibei Xu); zhusw@sdas.org (Shiwei Zhu)
0000-0003-0991-6405 (O.V.Koval); 0000-0002-8258-0816 (V.O.Kuzminykh); 0000-0002-8762-3918 (I.I.Husyeva); 0000-0003-1430-5334 (Beibei Xu); 0000-0002-6651-8449 (Shiwei Zhu)

In a graphical form, any ontology can be presented as a graph, with the nodes determining certain knowledge of the subject area under consideration and the edges - the connections between them, reflecting the sequence of user actions to achieve the goal. Permanent hierarchical connections of the type "part and whole" and "class and subclass" simultaneously determine the structure of actions and the structure of certain knowledge of the related subject area. Such a model can be built using approaches based on the product of a graph describing the data storage structure or the knowledge structure of the subject area and the graph describing the hierarchy of possible actions (methods) related to data processing [2].

Ontological description is a basis for constructing the subject area using graphic description. This approach gives a visual representation of the structure of the subject area in the form of its components and their connections.

One of the most common ways to solve the tasks of searching and processing data according to a request when they are stored on a distributed hierarchical network is to build a script based on a certain ontology, which a corresponding directed graph can describe. Such a graph should reflect both the data structure in the places of storage and placement by the database hierarchy. Various tasks in different areas of activity most often consist of performing a series of typical actions of the level of complexity corresponding to the request. Describing possible scenarios in the form of graphic structures greatly simplifies the application of automated data collection methods and optimization of such procedures within the information and analytical systems framework. At the same time, when constructing a graph based on which the most efficient scenarios could be selected, one of the main tasks is the need to construct a graph that would reflect both the data structure and the procedures for data extraction and primary processing to determine its relevance.

Today, many methods make it possible to build and select the most effective scenarios according to certain criteria from a set of scenarios corresponding to the given task. The construction of such methods in several cases can be based on the ontology and analysis of factors influencing the formation and execution of scenarios. In practice, the planning and implementation of the scenario determine the features of the paths that distinguish a scenario by factors established by the values of the criteria based on which decisions are made and the procedures for forming the necessary scenarios are determined. One of the rather complex and currently relevant tasks that require the construction of efficient scenarios for its solution is the construction and further optimization of scenarios for conducting analytical activities based on extensive information. In several cases, tasks related to optimizing the collection and consolidation of streaming information in information and analytical systems and decision-making systems are based on the processing of branched information and, by their scope, are big data processing tasks [3].

2. Construction of a graph for data collection

In the common case, building a general scenario model is necessary, which becomes the basis for solving the problem of analytical activities.

Such a model can be built using approaches based on the product of the graph describing the hierarchy of the data storage structure or knowledge structure of the subject area and a graph describing the hierarchy of possible actions (methods) related to data processing [4-7].

Knowledge structure of the subject area can be presented in the form of a graph adjacency matrix (Figure 1):

$$G = (V_1, E_1),$$

where $V_1 = \{v_1, v_2, \dots, v_n\}$ – set of nodes;

$E_1 = \{e_1, e_2, \dots, e_m\}$ – set of edges.

0	1	0	0
0	0	1	1
0	0	0	1
0	0	0	0

Figure 1: Graph adjacency matrix G (data hierarchy)

The structure of procedures for possible actions is presented in the form of a graph adjacency matrix (Figure 2):

$$H = (V_2, E_2),$$

where $V_2 = \{v'_1, v'_2, \dots, v'_n\}$ – set of nodes;

$E_1 = \{e_1, e_2, \dots, e_m\}$ – set of edges.

0	1	1	0	0
0	0	0	1	0
0	0	0	0	1
0	0	0	0	0
0	0	0	1	0

Figure 2: Graph adjacency matrix H (data processing hierarchy)

The adjacency matrices are obtained based on the classical formula:

$$a_{ij} = \begin{cases} 1, if\left(v_i, v_j\right) \in E; \\ 0, if\left(v_i, v_j\right) \notin E. \end{cases}$$

The corresponding graph models built based on adjacency matrices have graphical representations for the knowledge structure hierarchy graph of the subject area (Figure 3) and the hierarchy graph of possible actions (processing procedures) (Figure 4). These are generally directed graphs.

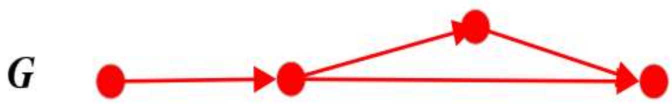

Figure 3: Graph of sources knowledge of the subject area

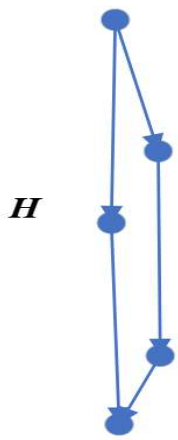

Figure 4: Graph hierarchy of data processing procedures (methods).

For further construction of scenarios, it is necessary to build a graphic model that unites the hierarchy of the knowledge structure of the subject area and the hierarchy of possible processing actions [8].

This model is presented in the form of a graph $J(V, E) = G \times H$, which will be built based on the Cartesian product of graphs.

That is, based on the use of the following ratios:

$$V = V_1 \times V_2$$

$$E = \begin{cases} 1, if\,(v_1 = v_1') \wedge (v_2\ adjacent\ v_2'); \\ 1, if\,(v_2 = v_2') \wedge (v_1\ adjacent\ v_1'); \\ \quad 0, in\ the\ opposite\ case. \end{cases}$$

Then, the adjacency matrix of the graph $J(V, E)$ will have the following form (Figure 5).

No	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
1	0	0	1	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0
4	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0
5	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0
6	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
8	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0
9	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0
10	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
12	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0

Figure 5: The adjacency matrix of the graph *J = G x H.*

Based on the adjacency matrix of the graph $J = G \times H$, , we will construct the following graph (Figure 6).

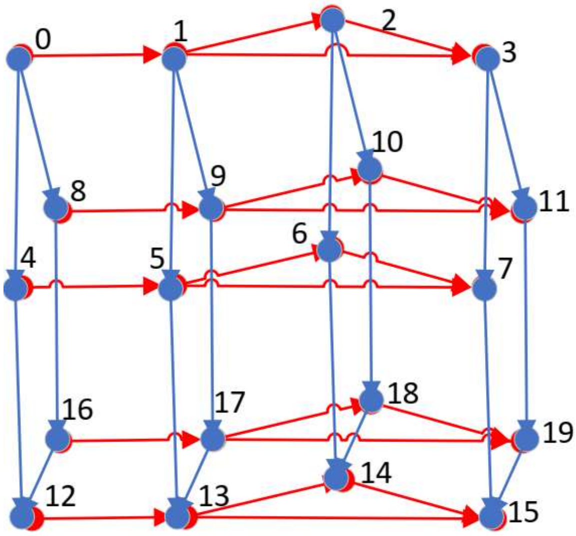

Figure 6: Graphical model *𝒥 = G x H*, which unites the hierarchy of the knowledge structure of the subject area and possible actions (processes) structures.

This model, built based on the Cartesian product of graphs, differs from the existing ones in that the structure hierarchy combines the knowledge and possible actions that describe the analytical process.

This approach allows to:

– build various variants of the data processing scenario;

– optimize the structure of possible scenarios;

– provide the analyst with information about possible alternative solutions when performing tasks while using the information and analytical system;

– reasonably determine connections between the knowledge of the subject area and actions of a specific analytical process;

– increase the reliability of the search for optimal analytical solutions in the hierarchical analysis of information that determines the necessary knowledge.

This approach to building a general model of possible data collection and processing scenarios can be widely applied when the form of information storage is identical for all possible nodes corresponding to its locations. Then the choice of the path in the forming of the scenario, according to the constructed graph of the hierarchy of the knowledge structure of the subject area and possible actions structure, can be selected and optimized, if necessary, according to the cost estimate for each path of the graph. Determining the required quality and efficiency scenarios for situations where the planning process is based on one factor is possible by applying the following strategies that use the hierarchical structure of the graph describing the relevant ontology [9].

The following main assessments of the optimality of scenarios can be distinguished:

- by priority, according to the hierarchy,
- with an assessment of the costs of using each scenario,
- taking into account expert and intuitive assessments.

Determining optimal scenarios regarding quality and efficiency in compliance has several features that require allocating two different planning procedures in hierarchical systems with several factors.

An important element while considering the influence of a certain number of factors on the determination of optimally necessary scenarios is the method of considering the advantages of certain factors on the value of the scenario evaluation. At the same time, the choice can be made both by determining the preference of one of the factors and by finding solutions based on considering the total evaluation of the factor for the entire selected path on the graph that defines the selected scenario.

To solve such problems within the framework of information and analytical systems, it is necessary to build algorithms that can provide the opportunity to optimally choose a scenario based on the hierarchical structure of the data description ontology, based on which the scenarios are built.

At the same time, the task of choosing the most efficient scenario is set by identifying and using the most appropriate scenarios for the conditions of a specific request [10-11].

3. Estimates of the optimal path on the graph

The total estimation of time consumption and resource consumption for each node of the graph and the total time consumption and resource consumption in relation to transitions between nodes can be used as a basis for choosing the optimal path on the graph, which, under such consideration, determines the optimal scenario [12]:

$$P_a = \sum_k (t_k + r_k),$$

t_k – processing time at node *k*;

r_k – resource costs on each of *k* nodes;

$k \in K$ – node numbers of the selected script execution path on the graph.

The evaluation criterion can be defining for such consideration as a total estimate of time and resource costs for each node of the graph and total time and resource costs for transitions:

$$P_d = \sum_{i,j} (T_{ij} + R_{ij}),$$

where T_{ij} – time spent on each of the *ij*-th edges of the graph;

R_{ij} – resource costs on each of the *ij*-th edges of the graph;

$i, j \in K$ – numbers of nodes corresponding to the edges of the selected path in the graph.

The total criterion is then equal to:

$$P = \sum_k (t_k + r_k) + \sum_{i,j} (T_{ij} + R_{ij}),\ k, i, j \in K.$$

The final estimates can be significantly influenced by the analysts' evaluations, which determine the quality and completeness of the results obtained for a specific task. These influences can significantly alter the results of assessments:

$$P = \sum_k (c_t t_k + c_r r_k) + \sum_{i,j} (c_T T_{ij} + c_R R_{ij}),\ k, i, j \in K,$$

where $c_t, c_r, c_T, c_R \geq 0$ - the impact coefficients on the time spent on each of the *k*-th nodes and on each of the *ij*-th edges of the graph.

As a rule, the model, which considers the costs related to both the aggregate estimate of work costs for each node and the cost estimate for the transition to the next node in the process of transition from the initial node to the final one, can only consider the costs on the edges of the graph.

At the same time, the costs in the nodes will be included in the costs of the edges of the graph.

$$P_{ij} = t_j + r_j + T_{ij} + R_{ij},\ i, j \in K.$$

The final estimates are thus affected by both the time and resource costs for each node and edge of the graph that describe the path on the graph corresponding to a given scenario and the resulting estimates. Sometimes, depending on the specific task, only time or resource consumption may be considered in the evaluation.

4. Stability conditions of the data processing process

Evaluation of the results at each step of scenario construction from the point of view of its optimality determines whether it is possible to increase the analyst's performance of the task:

$Q_i > 0$ – node *i* is considered selectable,

$Q_i = 0$ – the node is not considered,

Q_i – additional results were received by the analyst for node *i*.

The analyst should assess based on knowledge, experience, or obtained results. The analyst determines the Qi value to estimate the improvement in results.

Only those edges and nodes of the graph for which the condition $Q_i > 0$ is fulfilled are considered. All edges and nodes for which this condition is not fulfilled are not considered further. Further, this leads to a significant sparseness of the graph, which describes the possible direction of the analyst's execution of the scenario.

In turn, this, as a rule, leads, where necessary, to the determination of the data processing stability according to the graph's defined nodes, which is considered a model of possible scenarios of the analyst's analytical activity process [13-14].

The condition for ensuring the possibility of using all nodes for processing analytical data can be defined as:

$$\forall v_i \in V \Leftarrow \varpi_i(\tau) = 1, \tau \in [0, t],$$

where $V = \{v_i\}$ – set of analytical data processing nodes;

$\varpi(\tau)$ - is a Boolean function that takes on the value 1 if the data processing node takes part in the work and 0 - otherwise;

τ – is the current working time of analytical data processing nodes,

t – is the total time of performing analytical actions according to the scenario.

Then the condition for performing actions between data processing nodes along user action routes to achieve the goal of the task can be defined as follows:

$$\forall v_i, v_j \Leftarrow \exists v_j \in \Gamma_i, i, j = 1,2, \dots, n,$$

where Γ_i is a set of reachable nodes of the graph from node v_i,

n – is the number of analytical data processing nodes participating in the analytical process.

These conditions formalize the stability of the state of the analytical process in relation to the sequence of the analyst's execution of actions to obtain the necessary results based on the actions described by the corresponding directed graph.

This condition consists of the availability of knowledge about the process of analytical activity for edges and nodes and the presence of scenario routes that reflect the sequence of actions of the analyst to achieve the goal of the study:

$$\forall G(V, L), v_i \in V, l_{ij} \in L \Longleftarrow \begin{cases} \forall v_i \in V \Leftarrow \varpi_i(\tau) = 1, \tau \in [0, t); \\ \forall v_i, \forall v_j \Leftarrow \exists v_j \in \Gamma_i, i, j = \overline{1, n}. \end{cases}$$

The determination of the condition of stability of the process of analytical activity is based on certain knowledge describing the process of analytical activity and the possibility of using a set of other paths on the graph to achieve the required results:

$$\forall G(V, L), v_i \in V, l_{ij} \in L \Longleftarrow \begin{cases} \forall v_i \in V \Leftarrow \varpi_i(\tau) = 1, \tau \in [0, t); \\ \forall v_i, \forall v_j \Leftarrow \exists v_j \in \Gamma_i^{\geq 2}, i, j = \overline{1, n}. \end{cases}$$

A bridge is an edge of a graph connecting two subgraphs. If such an edge is removed, the graph ceases to be connected and turns into a two-component one.

A node of connectivity is a node of a connected graph; after its removal, together with the edge's incident to it, the graph is transformed from one-component to two-component.

The number of node connectivity $\chi(G)$ is the smallest number of nodes, the removal of which together with the edges incident to them leads to a disconnected or single-node graph.

The number of edge connectivity $\lambda(G)$ is the smallest number of edges, the removal of which leads to a disconnected graph.

Therefore, the current structure is on the limit of the stability of the analytical process, if the graph of the resulting structure is connected and contains bridges ($N_L \geq 1$) or connection nodes ($N_V \geq 1$):

$$\{K = 1\} \wedge [\{N_V \geq 1\} \vee \{N_L \geq 1\}],$$

where K is the number of components of the graph,

$K = 1$ means that the graph is connected;

N_V, N_L– the number of connectivity nodes and graph bridges, respectively.

A bridge or connectivity node connecting two subgraphs in the structure means that all routes of user actions to achieve the research goal from the nodes of one subgraph to the nodes of another will include this bridge or connectivity node. This event significantly reduces the stability of the analytical process. The analysis of the structures shows that if the graph describing the analytical process is on the limit of stability, then the system is operational and performs the defined number of functions. However, if at least one bridge or connectivity node fails, the system will go into an unstable state.

The areas of stable and unstable analytical processes can also be represented in Cartesian space in coordinates $\lambda(G)$ and $\chi(G)$ (Figure 7). Depending on the parameters $\lambda(G)$ and $\chi(G)$ of the graph structure, a point will be determined that will characterize the state of the analytical process.

In the graphic representation, the stability limit of the analytical process will be a set of points lying on two straight lines $\lambda(G) = 1$ and $\chi(G) = 1$.

It is important to determine how far the analyzed structure of the sequence of user actions to achieve the research goal is from the limit of stability of the analytical process or, on the other hand, what is the reserve of the specified stability of the analytical process. At the same time, the stability margin is determined according to the connectivity of the structure. In this case, the stability reserve is characterized by the number of irrelevant user actions to achieve the study's goal or the failure of nodes, which can bring the analyzed structure to the limit of a stable analytical process.

The margin of stability of the analytical process can be determined as follows:

- the margin of stability of the analytical process is a number Z_L, is equal to the distance from the point with coordinates $(\chi(G), \lambda(G))$ in the space of stability areas of the analytical process to a straight line $\lambda(G) = 1$, which describes the imaginary part of the limit of stability of analytical activity:
$$Z_L = \lambda(G) - 1.$$
- the peak margin of stability is the number Z_V, which is equal to the distance from the point with coordinates $(\chi(G), \lambda(G))$ in the space of stability areas of analytical activity to the straight line $\chi(G) = 1$, which describes the boundary:
$$Z_V = \chi(G) - 1.$$

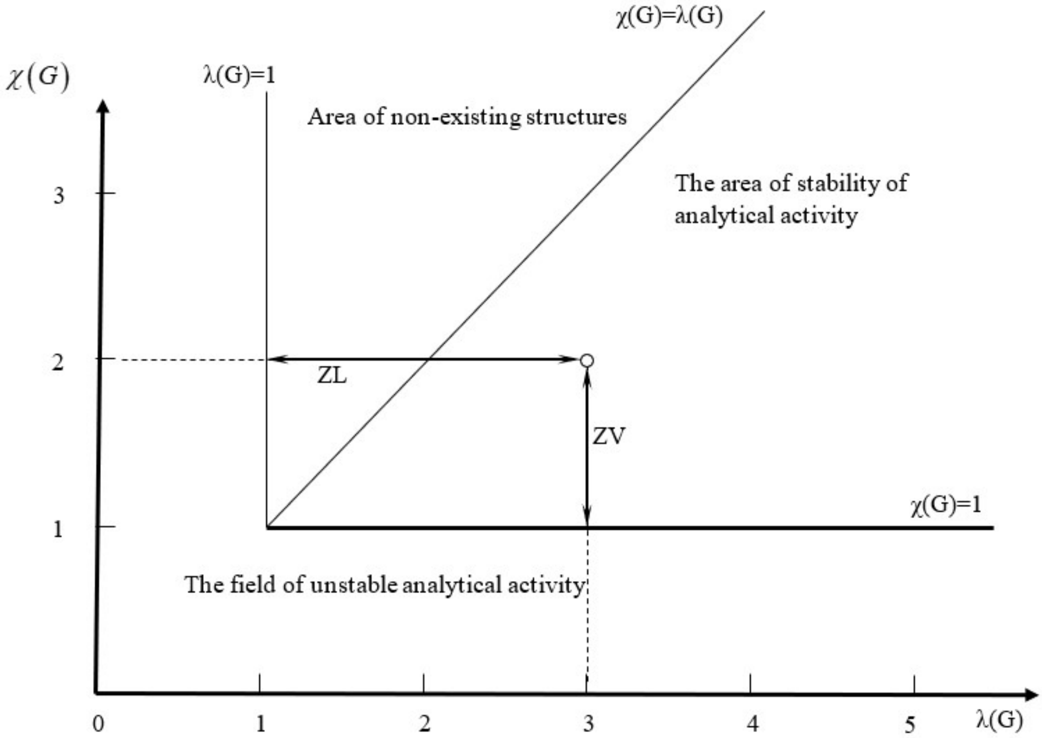

Figure 7: Geometric interpretation of stability areas

The essence of the margin of stability of the analytical process is as follows:

ZV - characterizes the number of nodes of the graph, after removal of which, together with the edge's incident to them, the structure passes from a stable state to the limit of stability. This value characterizes the maximum number of failures of nodes containing certain knowledge, which can be compensated by using alternative nodes;

ZL– characterizes the number of edges of the graph, after removing which, the structure passes from a stable state to the limit of stability of the analytical process. This value characterizes the maximum number of removed edges of the graph, which reflects the sequence of user actions to achieve the research goal, which can be compensated by using alternative routes.

Thus, based on signs of stability of the analytical process, after determining the considered parameters, it is possible to discover the state of the analytical process, namely, whether the system is in a stable or unstable state [15]. The measure of stability of the analytical process is determined by the margin of stability, which can be decided analytically using the proposed dependencies. It is possible to define and justify the general requirements for the description, formalization, and sequence of the organization of the analytical process for information-analytical systems and decision-making support systems, reflecting the actions of the analyst based on the description of the subject area in the form of an oriented graphic structure.

5. Conclusions

The article shows the method of constructing a graph describing the scenario of analytical activity, which is built based on the Cartesian product operation of the knowledge sources graph of the subject area and the hierarchy of data processing procedures (methods). This approach makes it possible to build a fully connected graph that reflects the possible scenarios of the analyst's search for data according to the task.

For a general approach to building an optimal scenario for the collection, primary processing, and consolidation of branched data, procedures for evaluating the efficiency of paths on a directed graph are provided, which, in turn, gives a certain opportunity to the user of the analytical system to build a certain scenario that best meets the purpose.

Based on the developed signs and indicators of the stability of the analytical process, the concepts of the limit and margin of stability of the analytical process are introduced. Quantitative methods of assessing the stability of the analytical process by numerical indicators are proposed. As an advantage of this method, it is possible to quantitatively assess the stability of the analytical process of the current structure of the analyst's sequence of actions to achieve the goal of the task based on simple signs. Using such estimates in the general case makes it possible to determine and formalize the requirements for the organization and execution of the analytical process as a sequence of actions of the analyst in accordance with the description of the subject area in the form of an oriented graph.

The formation of analyst activity scenarios based on the graph of the scenario description, which is built by using the operation of the Cartesian product of the knowledge sources graph of the subject area and the graph of the hierarchy of procedures (methods) of data processing, can be especially effective when the information in individual sources is partially or completely duplicated and has a different level of actualization. At the same time, the types and forms of data storage are a priori, as a rule, significant. Then the use of performance evaluations and determination of the data processing stability conditions gives the scenario the highest quality. The described approach can be applied to construct various information and analytical systems for data collection and processing.

References

[1] Dodonov A.G., Senchenko V.R., Koval A.V. Analytics and knowledge in computer systems. Kyiv: IPRY of the National Academy of Sciences of Ukraine, "KPI named after Igor Sykorsky", 2020. 315 p.

[2] Oleksandr Koval, Valeriy Kuzminykh and Maksym Voronko. "Standard Analytic Activity Scenarios Optimization based on Subject Area Analysis", in Selected Papers of the XIX International Scientific and Practical Conference "Information Technologies and Security" (ITS 2019) CEUR Workshop Proceedings. 2019. Vol. 2577. R. 37–46. http://ceur-ws.org/Vol-2577/paper4.pdf

[3] Chernyak, Leonid Large data — new theory and practice// Open systems. DBMS. — Moscow: Open Systems, 2011. — No. 10. — ISSN 1028-7493

[4] Crowe M. J. A History of Vector Analysis – The Evolution of the Idea of a Vectorial System. — Courier Dover Publications, 1994. — C. 32. — 270 c. — ISBN 0486679101.

[5] Imrich, Wilfried; Klavžar, Sandi. Product Graphs: Structure and Recognition. — Wiley, 2000. — ISBN 0-471-37039-8.

[6] J.A. Bondy, U.S.R. Murty (2008). Graph Theory. Springer. c. 654 c. ISBN 978-1-84628-969-9.

[7] Jungnickel, Dieter (2008). Graphs, Networks and Algorithms. Springer. c. 650 c. ISBN 978-3-540-72779-8.

[8] France R., Rumpe B. Model-driven development of complex software: A research roadmap // 2007 Future of Software Engineering. – IEEE Computer Society. 2007. pp. 37–54.

[9] Yu. Ishchenko. Ontology // Philosophical encyclopedic dictionary / V. I. Shinkaruk (chief editor) and others. — Kyiv: Grigory Skovoroda Institute of Philosophy of the National Academy of Sciences of Ukraine: Abrys, 2002. — P. 449. — 742 p. — 1000 copies. — BBK 87ya2.— ISBN 966-531-128-X.

[10] Ontology and metaphysics: Textbook. T.3 / B. Mondin; Translated from Italian. B. Zavidnyak. — Zhovkva: Missioner, 2010. — 284 c. — (Student of systematic philosophy).

[11] Koval, O.V., Kuzminykh, V.O., Husyeva, I.I., Beibei, X., Shiwei, Z. Improving the Efficiency of Typical Scenarios of Analytical Activities CEUR Workshop Proceedings, № 3241, pp. 123-132. 2022. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85139866439&partnerID=40&md5=e78a2c0fb37daaa43c442d81c177bf07

[12] Graph Analysis and Visualization: Discovering Business Opportunity in Linked Data Published by John Wiley & Sons, Inc. 10475 Crosspoint Boulevard Indianapolis, IN 46256

[13] Lee, C Y. "An Algorithm for Path Connections and Its Applications." IEEE Transactions on Electronic Computers 10, no. 3 (September 1961): 346–65.

[14] Beibei Xu, Valeriy Kuzminykh, Shiwei Zhu, Junfeng Yu, Mingjun Zhang, Sisi Li, Dmytro Lande.Research on library construction driven by multi-technology integration. Third International Symposium on Computer Engineering and Intelligent Communications. Xi'an, China (2023). Proceedings Volume 12462, https://doi.org/10.1117/12.2660850

[15] Thacker BH et al. Concepts of model verification and validation. – Los Alamos National Lab., Los Alamos, NM (US), 2004. No. LA-14167.