DOI: 10.35681/1560-9189.2021.23.1.235408

UDC 621.391

O. V. Koval[1], V. O. Kuzminykh[1], S. Y. Svistunov[1], Xu Beibei[2], Zhu Shiwei[2]

[1]National Technical University of Ukraine
«Igor Sikorsky Kyiv Polytechnic Institute»
37, Pr. Peremogy, 03056 Kyiv, Ukraine
[2]Information Research Institute of Shandong Academy of Sciences
19th Keyuan Road, Jinan, Shandong Province, China (250014)

Data collection for analytical activities using adaptive micro-service architecture

A structure of the adaptive query processing subsystem with micro-service architecture, which is being developed within the framework of creating an information and analytical system for assessing the level of international activity, has been considered. When building a request processing system, the use of adaptive micro-service architecture is proposed. An extensive system of network hosts for micro-services, focused on the possibility of more efficient extraction of information from the sources. The subsystem provides the use of micro-services as stand-alone containers, each of which is associated with specific information sources. The system provides the adaptation of parameters, for both individual micro-services and scenario parameters. Formation and management of containers involves the use of modern specialized platforms Docker and Kubernetes, which allows optimize the management of containers and control the load on each host.

Key words: *quality assessment, scenarios of analytical activity, information collection, microservice architecture, adaptation, big data.*

Introduction

One of the current trends in the development of universities and research institutions is the forecasting of performance indicators, by which they are evaluated in comparison with similar institutions both inside the country and internationally. Evaluation is made in the form of various ratings and in the form of monitoring the efficiency of the activities. The ability to assess the current situation and predict the value of performance indicators for the future is an important criterion in determining the ranking of the institution in the world and an important factor in shaping the directions of development.

To successfully advance in the international educational and scientific space and increase the interest from foreign researchers, faculty and students, it is necessary to have a clear understanding of the level at which they are in terms of international recognition [1].

This is especially important when investing in science by the state or funds, both international and domestic, when it is necessary to take into account the actual quality of work that can be attributed to international activities.

Modern development of information technologies is characterized by a number of trends which require new understanding and development of new efficient architectural solutions and approaches to the design and implementation of software systems. This is especially important for decisions aimed at supporting analytical activities. This is due to the peculiarity of the functioning of modern analytics systems, which objectively require the development of intelligence in both decision-making processes and software systems and technological components to support such activities. Taking into account the dynamics of specific environmental requirements, the complexity of the system integration problems dictate the need to develop methods and tools for designing a distributed information system for analytical tasks, which is based on distributed software architecture for analytical activities and covers a wide range of information sources [2, 3].

The concept of «international activity» of universities and research institutes is a rather «vague» criterion and currently has no clear definition. Experts note that depending on the goals of the analysis set by the university or research institute, both the number of criteria used for evaluation and the method of calculating the complex indicator vary [4].

An important question is how the international activities of universities and other world-class research organizations differ from other types of institutions.

To date, a number of such basic characteristics have been identified. It is a highly qualified teaching and research staff, which has a significant number of publications in international journals indexed in bibliographic databases; outstanding results of internationally recognized research; availability of international patents; large amounts of funding for international research; availability of foreign students and PhD-students; academic freedom; clearly defined management structures in accordance with international standards; conducting research through international grants and programs [5, 6].

It is the reflection of these characteristics in the form of digital indicators presented in many methods of assessing international activities.

Problem Statement

The variety of data types on scientific and technical international cooperation is significant and reaches several hundred titles [7–9]. This raises the dilemma of choosing a basic set: a broad list facilitates a more detailed assessment of individual participants of international activities, but makes it difficult to compare with other actors in international scientific and technical cooperation, as most data values, and hence the indicators developed based on them, may be uncertain, due to unavailability of data. Therefore, the choice of the basic set should be made taking into account the purpose of the indicator, the availability of data and the technical capabilities for the extraction and processing of such information [10].

This leads to the need to consider all the information necessary to assess the level of international activity as structured information, the essence of which reflects the most important elements of the analysis of the level of international activity, to further determine the ratio of these assessments for each participant.

That is why the problem of choosing criteria for evaluating international activities, developing methods for searching and extracting numerical values of criteria from available sources of information, developing methods for calculating a comprehensive assessment criterion taking into account the contribution of each criterion remains relevant and in demand.

When developing an information and analytical system for analyzing the level of international activity, it is necessary to expand the approach to assessing the international scientific activity of engineering, scientific and educational organizations and structure those indicators which to some extent determine the final total value of the indicator. This is especially important in cases where, on the basis of these assessments, decisions are made about the prospects for continued funding of certain areas in research or educational organizations, the need to expand scientific cooperation in a particular scientific field.

The basis for solving problems of developing an information and analytical system for analysis the level of international activity is the development of effective methods for extracting the necessary information from various sources, where it is partially, to some extent relevant and reliable [11, 12]. To do this, it is necessary to develop efficient methods and architectural solutions for software tools for data collection of analytical activities.

Modern solutions in the software development for data collection of analytical activities

Scenarios for information collecting and analyzing

In general, there are many different definitions of the term «scenario». Some researchers understand the scenario as a static picture of the future, others — as a sequence of events over time; also highlight many differences in other characteristics of the scenario. The general characteristics of the scenarios should be highlighted [13]:

— hypothetical nature of the scenario;
— description of alternatives for the future;
— chain of events;
— causal relationship;
— the presence of internal consistency;
— basis for action.

As a result of generalizing these characteristics, we take the following as the most general definition of a scenario: a scenario is a step-by-step description of alternative hypothetically possible scenarios in the future, reflecting different perspectives on past, present and future, and can be a basis for action planning [14].

The successful application of the scenario approach by a number of different companies, as well as its growing popularity in public strategic management has given its considerable popularity and prevalence. One of the most important features of the scenario approach is its versatility, flexibility, the ability to apply a variety of methods at different stages of scenario research and process management. The development of a scenario approach using a number of analytical methods will increase the reliability and validity of the results. This increases the efficiency of its application in the strategic management of complex information and analytical systems.

Today, there are a number of methods for selecting the necessary quality and efficiency of scenarios from a set of possible or acceptable scenarios, which are based on an analysis of the composition and nature of the factors influencing the process of scenario planning. In practice of planning there are situations which differ in the number of factors on the basis of which decisions are made and determine the fundamental differences in the procedures for forming the necessary scenarios. One of the most complex and relevant tasks today, which requires the construction of efficient scenarios for its solution, is the task of building and further optimizing scenarios for the collection and analysis of various processes of organizations. The task of optimizing scenarios for collecting and analyzing information on the network is one of the most important tasks for processing large amounts of data (big data) [15].

Scenarios can be seen as a way to collect and analyze a complex information environment in which there are many meaningful, interconnected and alternative sources of information. In this case, the scenario itself may include a possible set of events which determine the development of certain factors that affect the outcome of the activity [16].

From the point of view of developing a system for analysis the international activities of organizations, a typical scenario of information retrieval at the request of the analyst can be understood as a sequence of actions (steps) of the information retrieval system selected by the analyst to solve such a task. Typical scenarios can be changed and transformed while searching for information, followed by saving new versions and versions of versions. After conducting analytical research, new copies of versions can be saved and become separate versions of standard scenarios for further use by the analyst to solve similar analytical problems.

One of the main problems for the system of analysis the international activities in the collection of primary information is its incompleteness and improper relevance in the vast majority of primary information sources. This leads to the need for duplication in determining the sources of information, analysis of information quality during information retrieval and the need to build complex algorithms for the sequence of analysis and processing of information based on typical scenarios of information analysis at the request of the analyst [17]. The complexity of such procedures requires not only the adaptation of the structure of the steps of typical scenarios to the specific task set by the analyst, but also the adaptation of the parameters of individual elements that implement certain steps of the scenario to these features both during and after the analysis [18].

Efficient implementation of such complex in structure and significant in terms of data processing systems is possible only with the use of methods and processes of parallel information collection, which can be qualitatively implemented on distributed networks using micro-service architecture.

Micro-service software architecture

Micro-service architecture of information systems is a set of independently deployed services. The system constructed in this way has the ability to evolve in parts, primarily due to the fact that each micro-service is largely autonomous. In addition, this architecture allows flexible scaling of information system components, thereby ensuring optimal use of existing server equipment. The scaling of each micro-service is performed independently of other micro-services. Allocation of computing resources became possible taking into account the real user activity and demand for a particular functionality of the system.

Micro-service architecture is an approach to application development that involves abandoning a single, monolithic structure. This is a way to develop software applications by creating separate independent software modules. Each of them is responsible for a specific task, can be changed, supplemented and expanded. Moreover, these software modules can run on different servers and interact with each other using network tools, such as HTTP [19].

One of the important advantages of micro-services is that we can use different technologies to solve the same problem. For example, use different parsing libraries or different data processing tools in each micro-service. This makes it possible to build adaptive to the type of task performed software that implements certain elements of machine learning. It is possible that some technology and libraries will get out of control. So you need to choose a basic set of tools and use others only when it is really needed [20].

In addition, this approach makes it possible to build and use patterns of schemes for collecting and primary processing of information as part of the scenarios for search and processing of information for specific tasks of assessing the level of international activity [21].

Today there are many approaches and tools for implementing micro-service architecture. Among them, the most developed and widely used are the platforms Docker [22] and Kubernetes [23].

Docker is a platform for developing, delivering and launching container applications. Docker allows create containers, automate their launch and deployment, manages the life cycle. It allows run multiple containers on a single host machine.

Containerization is similar to virtualization, but is not the same. Virtualization works as a separate computer, with its own virtual hardware and operating system. Thus within one operating system it is possible to start one more operating system. In the case of containerization, the virtual environment runs directly from the core of the main operating system and does not virtualize the equipment. This means that the container can only be run on the same operating system as the main one. However, as containers do not virtualize equipment, they consume much less resources.

Docker is one of the most famous tools for working with containers.

Containers are a way to pack the application and all its auxiliary elements into a single image [24]. This image runs in an isolated environment that does not affect the main operating system. Containers separate the application from the infrastructure. Therefore, developers do not need to worry about the environment in which the application will work, whether there will be settings and dependencies. The developers simply create the application then pack all the dependencies and settings into a single image. You can then run this image on other systems without worrying that the program will not start. Containers are a good alternative to hardware virtualization. They allow you to run applications in an isolated environment, but consume much less resources.

A container can be thought of as a set of processes isolated from the main operating system. Applications run only inside containers and do not have access to the main operating system. This increases the security of applications because they will not be able to accidentally or intentionally harm the main system. If the application in the container terminates with an error or crashes, it will not affect the main operating system.

If you need to place containers on different hosts with different operating systems, a more functional and efficient container management system is the Kubernetes platform.

Kubernetes was developed by Google as an open-source container management platform designed to deploy, scale, and manage container applications. Kubernetes has become a widely accepted container management standard supported by key market players such as Google, Amazon Web Services (AWS), Microsoft, IBM, Intel, Cisco, and Red Hat.

Kubernetes makes it easy to deploy and use applications in micro-service architecture, creating a level of abstraction on top of a host group. As a result, development teams can deploy applications and allow the Kubernetes platform to:

— control how many resources the application or team consumes;

— evenly distribute the load of applications on the host infrastructure;

— automatically send load balancing requests between different instances of the application;

— monitor resource consumption and limits to automatically stop and then restart programs that consume too many resources;

— move an instance of the application from one host to another if the host is running low on resources or has crashed;

— automatically deploy additional resources when adding a new host to the cluster.

The structure of the software implementation of the data collection subsystem for the analysis of international activities

To solve a complex task as the development of a data collection subsystem for the analysis of international activities, an architecture based on the principles of scenario approach using adaptive micro-service architecture with machine learning elements is used. This subsystem is designed for parallel data collection on various aspects of international activities from different in structure, form of storage and location information sources in real-time. This subsystem allows collect a variety of information about both individual scientists and the groups, and organizations built on any criteria. The schematic diagram of the implementation of such a system is presented in Fig. 1.

The general process of the information and analytical system for assessing the level of international activity can be divided into the following main steps:

1) request processing and selection or scenario formation;

2) based on the scenario form the micro-services structure to perform the tasks concerning inquiry;

3) collection and processing of information upon request;

4) consolidation and preservation of results for further analysis;

5) analysis of micro-services and adaptation of their parameters;

6) evaluation of the scenario and its adaptation for specific analysis tasks;

7) analysis of the received information and results of the request.

The first step is the analysis of the request, if necessary, its clarification and formalization of the request in terms of the possibility of its implementation within the information and analytical capabilities of the system functions. For a refined and formalized query, a typical scenario is selected from the standard scenario repository, and, if necessary, a new scenario is formed, usually based on one of the standard scenarios [22].

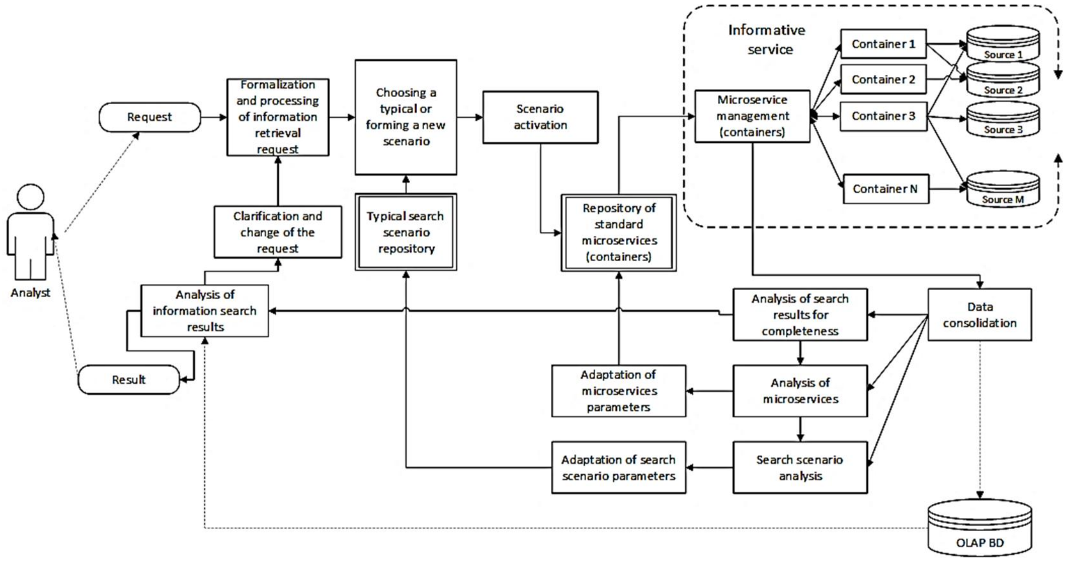

Fig. 1. Schematic diagram of the information collection subsystem of the information and analytical system for assessing the level of international activity

In the second step, the connection is made in accordance with the scenario of the necessary micro-services, as elements of the functions of collecting and processing information according to the typical scenario or its version. This scenario activation procedure is performed using a standard micro-service repository. Thus micro-services are developed as separate containers. To do this, the tools of the software platform Docker for rapid development, testing and deployment of microservices are used [23]. Docker packs the software into standardized blocks called containers. Each container includes everything that is needed to run the program: libraries, system tools, code and runtime. With Docker, one can quickly deploy and scale micro-services in any environment.

The third step is the collection and processing of information on request. This uses the Kubernetes platform, which allows you to easily implement and use applications in micro-service architecture, creating a level of abstraction on top of a group of hosts [21]. The Kubernetes platform allows to:

1) deploy containers and all operations to run the required configuration. These include restarting stopped containers, moving them to allocate resources to new containers, and other operations;

2) scale and launch multiple containers simultaneously on a large number of hosts;

3) balance many containers during start-up. To do this, Kubernetes uses an API whose task is to logically group containers. This makes it possible to identify their pools, set the placement, as well as evenly distribute the load.

The fourth step involves the consolidation of information obtained from various sources. Consolidation is seen as a set of methods and procedures aimed at combining data from different sources, ensuring the required level of relevance, completeness and conversion into a single format in which it can be uploaded to the database [25].

To save the received information after its consolidation, the use of OLAP database is provided, which gives a number of advantages for further processing depending on the need of the analyst who provided the request [26].

The fifth step determines the procedures for analyzing the work of micro-services and adapting their parameters (Fig. 2). Adaptation of micro-services parameters is considered as correction of values of parameters of standard micro-services from the repository of micro-services that leads to emergence of new versions of these micro-services which will be fixed on standard scenarios as a result of which performance those versions appeared.

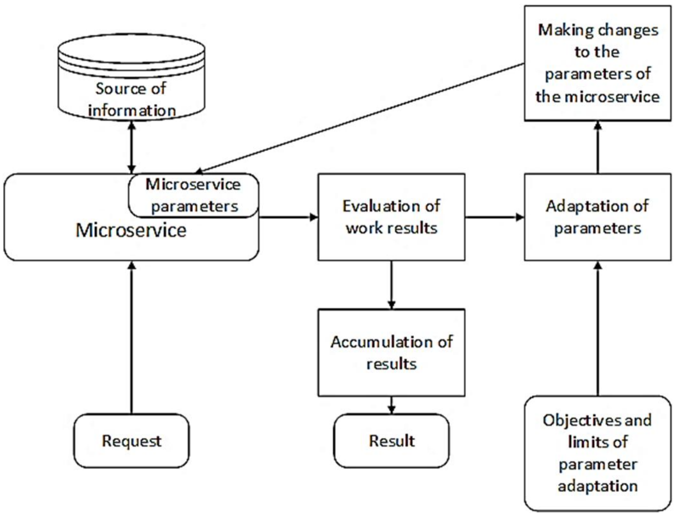

Fig. 2. Adaptation of micro-service parameters

The following mathematical model of the problem statement for adaptation of parameters of micro-services is considered.

$A_j = \{a_{1j}, \ldots, a_{mj}\}$ — parameters j-th automaton $j = \{1, \ldots, n\}$

with restrictions

$$a_{ij}^{\min} \le a_{ij} \le a_{ij}^{\max} \text{ for } i = \{1, \ldots, m\} \text{ and } j = \{1, \ldots, n\}.$$

In the general case, the restrictions can be both in the form of boundaries and functional.

The function that determines the quality of the j-th machine for the information collection subsystem can be set as

$$Q_j(A_j) = \frac{k_{uj}}{k_{zj}},$$

where k_{uj} — successful attempts to find relevant information of the j-th machine; k_{zj} — total number of attempts of the j-th machine M.

Then $\sum_{j=1}^{n} k_{zj} = M = \text{const}$ — total number of attempts for all the machines at the current iteration.

Depending on the results of the micro-service on the current iteration, the number of attempts to perform the micro-service on the next iteration is determined

$$k_{zj}^{new} = k_{zj} + \Delta k_{zj}$$

at

$$\Delta k_{zj} = \varphi\,(k_{zj}, Q_j(A_j)).$$

In the simplest case, for the information collection subsystem, the value of Δk_{zj} can be determined taking into account the average value of all machines in the current iteration.

For example, the definition of the value of Δk_{zj} can be built on the basis of comparing $Q_j(A_j)$ with the average value for all micro-services in the current iteration of the search

$$Q_j^*(A_j) = \frac{\sum_{j=1}^{n} Q_j(A_j)}{n}.$$

Then if

$$Q_j(A_j) \le Q^*_j(A_j) \text{ then } \Delta k_{zj} \le 0,$$

and when

$$Q_j(A_j) > Q^*_j(A_j) \text{ then } \Delta k_{zj} > 0.$$

Experimental data have shown that it is quite effective to reduce or increase the values of k_{zj} depending on the values of $Q_j(A_j)$ and $Q^*_j(A_j)$ by 15–20 % with subsequent normalization to the total value of M.

The sixth step is to assess the performance of the scenario and its adaptation to the specific tasks of the analysis of the level of international activity. To build new versions and instances of typical scenarios, it is proposed to use an algorithm based on a linear stochastic automaton [27]. Structural adaptation based on the usage of linear stochastic automata is based on the assessment of the most efficient sources of information for a particular request and methods of information processing possible for these sources [28]. This will allow to correctly, in relation to the highest relevance, to choose the sources of information according to the request and methods of processing this information.

For a system of assessing the level of international activity in the scenario structure, the most critical are alternative methods of extracting information that are directly related to the relevant sources of information.

For the first group, an important parameter of the quality of work that can affect the adaptation of the scenario is the number of relevant results obtained over time.

In terms of information retrieval, we can determine:

$u(t)$ — the state of the system at the current time t, which determines the probabilities of choosing alternative micro-services for information retrieval or alternative micro-services for information processing ($M_1, \ldots, M_n$) on the current iteration;

$u(t + 1)$ — the state of the system at the next moment (in the next iteration), which determines the probability of choosing alternative micro-services ($M_1, \ldots, M_n$) at the next iteration;

$x(t)$ — the input data of the system on the current iteration, which determines the results of the evaluation of the sample size V from the selected on the current iteration of the alternative micro-service;

$y(t)$ — the data source of the system on the current iteration, which determines the selected alternative micro-service (M_1, …, M_n).

The implementation of such a machine to select an alternative micro-service can be constructed in such a way that the change in state of the machine $u(t + 1)$ is defined as a regular dependence, and its output $y(t)$ is defined as a stochastic process.

An algorithm based on the use of a stochastic automaton consists of the following steps.

1. The initial state of the stochastic automaton $u(t)$ for $t = 1$ is determined as a vector

$$P(t) = \{p_1(t), p_2(t), p_3(t),..., p_i(t),..., p_{n-1}(t), p_n(t)\},$$

where $p_i(t)$ — probability of using i-th micro-service from a group in which there are n alternative micro-services

At the same time dependence is fulfilled $\sum_{i=1}^{n} p_i(t) = 1$.

2. A random variable w is generated, which is evenly distributed on the interval [0,1].

3. Random value w determines the interval corresponding to one of the n alternative micro-services of information.

Determine $p_0 = 0$, then, if

$$\sum_{i=0}^{z-1} p_i(t) < w < \sum_{i=1}^{z} p(t)$$

then the implementation of the machine will correspond to the micro-service with the number z.

4. V units of information (documents) are sampled or processed by the micro-service under number z.

5. The quality of the micro-service R_i is evaluated by the number z for V units of information.

6. According to a certain algorithm for each micro-service, the probabilities are recalculated.

The conversion factor is determined

$$k_R = f(R_i).$$

For example, the probability factor of a stochastic automaton can be made as shown in the table. The values of R_i and k_R used in the table were obtained and performed well in testing the algorithm.

R_i	0–0,2<	0,2–0,4<	0,4–0,6<	0,6–0,8<	0,8–1,0
k_R	0,5	1,0	1,5	1,7	2

Then $p_z(t+1) = p_z(t)k_R$.

The calculation is performed for the vector $P(t + 1)$.

With

$$D(t) = 1 - p_z(t),$$

$$D(t+1) = 1 - p_z(t+1),$$

$$p_i(t+1) = \frac{p_i D(t+1)}{D(t)} \quad \text{for } i = 1,...,n \text{ and } i \neq z.$$

7. Step 2 to continue the execution of the scenario with the selected alternative micro-service, which is associated with a specific source of information or method of processing the extracted information.

The seventh step concerns the analysis of the received information according to inquiry. Depending on the assessment of the quality of the received information and its compliance with the request, the analyst, in case of his dissatisfaction, clarifies or changes the request, and otherwise the processing the request is complete.

Evaluation of algorithm efficiency

To assess the efficiency of the algorithm, step-by-step tests were performed on a test environment. A test environment was built, which included ten sources of information, each of which was associated with a corresponding micro-service, which extracted relevant information on request. Two thousand records were generated in each source. Randomly, there are 10 to 30 relevant records corresponding to the query in each source. The information was removed step by step for 1000 requests in one step. The number of requests to information sources for each of the ten microservices at each step was distributed according to the characteristics of the algorithms for the four options for data collection analyzed.

Fig. 3 shows graphs for four analysis options: without adaptation (WA), with adaptation of micro-service parameters (PA), with adaptation of scenario (SA) and with two-level adaptation (SPA) of both micro-service parameters and information retrieval scenario. The two-level adaptation showed almost a twofold effect compared to the sequential extraction of information and a significant increase in efficiency relative to the partial, both parametric and adaptation of scenarios.

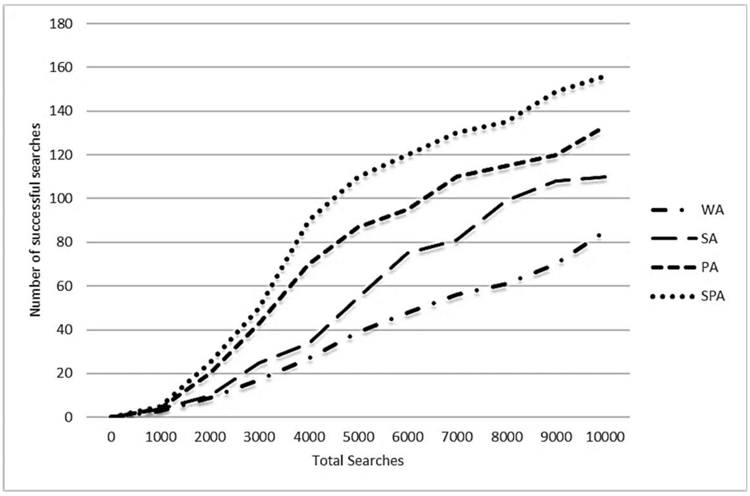

Fig. 3. Test results

Conclusions

1. The general structure of the adaptive subsystem for processing the inquiries which is developed within the limits of creation the information and analytical system for assessment the level of the international activity is considered. The use of micro-service architecture is the basis for building a query processing subsystem. Placement of micro-services is implemented on a distributed network. An extensive system of hosts on which micro-services are hosted, focused on the possibility of more efficient extraction of data from a specific source from which the micro-service receives information. The subsystem provides the use of micro-services as stand-alone containers. Each of the containers is designed to perform certain operations to extract the necessary information from a specific information source with parameters that can be adapted to a specific request. For the formation and further management of containers, the use of Docker and Kubernetes platforms is envisaged, which allows, in addition, to control the load on each host.

2. Each request of the analyst after its verification for feasibility and formalization is implemented in the form of a scenario based on the use of a repository of standard scenarios that can be used without changes, modified or further developed and included in the repository, depending on the specific request of the analyst. The subsystem provides the ability to adapt such scenarios due to the availability of alternative micro-services that can be included in the scenario, or change the sequence of use of micro-services.

3. The analysis of test results of the offered approach to realization of adaptive micro-service architecture for the decision of problems of parallel collecting of the information from many sources is given.

4. The adaptive information collection subsystem for query processing presented in the article, which is developed within the framework of creating an information and analytical system for assessing the level of international activity, can be used in the development of various information and analytical systems focused on real-time network information collection.

1. Михайлов О.В. Критерии объективной оценки качества научной деятельности. *Вестник российской Академии наук*. ISSN: 0869-5873. 2011. Т. 81, № 7. С. 622–625

2. Alex Guazzelli, Michael Zeller, Wen-Ching Lin and Graham Williams PMML: An Open Standard for Sharing Models: available. URL: https://journal.r-project.org/archive/2009-1/RJournal_2009-1_Guazzelli+et+al.pdf

3. Kuzminykh V., Koval O., Khaustov D. Using stochastic automation for data consolidation. *Research Bulletin of NTUU «KPI». Engineering*. 2017. No. 2. P. 29–36.

4. Адинда ван Гален, Рената Гилесен. Домашняя интернационализация: голландский подход. Международное высшее образование. Русскоязычная версия информационного бюллетеня *International Higher Education* (Бостонский колледж, США). 2015. Т. 78. С. 16–18.

5. Джамиль Салми. Создание университетов мирового класса. Международный банк реконструкции и развития / Всемирный банк, 2009, 134 с. ISBN 978-5-7777-0396-5.

6. Арчаков А.И., Карпова Е.А., Пономаренко Е.А. Международные критерии эффективности научно-исследовательской деятельности коллективов и отдельных ученых в области биологии и медицины. URL: https://giduv.com/nauka/mezhdunarodnye-k

7. Luukkonen Terttu & Tijssen Robert & Persson Olle & Sivertsen Gunnar. (1993). The Measurement of International Scientific Collaboration. Scientometrics. 28. 15–36. 10.1007/BF02016282

8. Lane J. Let's make science metrics more scientific. *Nature*. 2010. No. 464. P. 488–489.

9. Nassi-Calò L. Evaluation metrics in science: current status and prospects. Rev. 25 (2865). *Latino-Am. Enfermagem.* 2017. H. 25–28.

10. Шаховская Н.Б., Пасечник В.В. Хранилища и пространства данных/монография. Львов: «Львовская политехника», 2009. 244 с.

11. Ахо А.В., Хопкрофт Дж., Ульман Дж.Д. Структуры данных и алгоритмы. Москва: Изд. дом «Вильямс», 2019. 400 с. ISBN 978-5-6041393-6-3.

12. Белоногов Г.Г, Богатырев В.И. Автоматизированные информационные системы. Москва: Сов. радио, 1973. 328 с.

13. Van Notten Ph. Writing on the wall: scenario development in times of discontinuity. Florida: Boca Raton, 2005. 209 p

14. Van Notten Ph. Scenario development: a typology of approaches. Think Scenario. Rethink Education. OECD, 2006. – P. 69–84.

15. Натан Марц, Джеймс Уоррен. Большие данные: принципы и практика построения масштабируемых систем обработки данных в реальном времени. Москва: Диалектика-Вильямс, 2018. 368 с. ISBN 978-5-907114-02-9

16. An introduction to the owl web ontology language, available. URL: http://www.cse.lehigh.edu/~heflin/IntroToOWL.pdf]

17. Kuzminykh V., Koval, O., Voronko M. Standard Analytic Activity Scenarios Optimization based on Subject Area Analysis. *CEUR Workshop Proceedings*. ISSN: 1613-0073. 2019. Vol. 2577. P. 37–46. URL: [http://ceur-ws.org/Vol-2577/paper4.pdf]; https://herbertograca.com/2017/01/26/microservices- architecture/

18. Додонов О.Г., Коваль О.В., Глоба Л.С., Бойко Ю.Д. Комп’ютерне моделювання інформаційно-аналітичних систем. Київ: ІПРІ НАН України, 2017. 239 с.

19. Ford Neal, Parsons Rebecca, Kua Patrick. Building Evolutionary Architectures: Support Constant Change. O'Reilly Media, 2017. 332 p. ISBN-10 1491986360; ISBN-13 978-1-491-98636-3.

20. Koval O., Kuzminykh V., Otrokh S., and Kravchenko V. Optimization of Scenarios for Collecting Information Streaming Wide-Area Network. 2019. 3rd International Conference on Advanced Information and Communications Technologies (AICT). Lviv, Ukraine, 2019. P. 213–215. doi: 10.1109/AIACT.2019.8847832.

21. Ричардсон Крис. Микросервисы. Паттерны разработки и рефакторинга. Санкт-Петербург: Питер, 2019. 544 с.: ил. ISBN 978-5-4461-0996-8

22. Pethuru Raj; Jeeva S. Chelladhurai; Vinod Singh. Learning Docker. Packt Publishing, 2015. 240 c. ISBN 978-1-78439-793-7.

23. Лукша М. Kubernetes в действии = Kubernetes in Action. Москва: ДМК Пресс, 2018. 672 с. ISBN 978-5-97060-657-5.

24. Ньюмен С. Создание микросервисов = Building Microservices. Санкт-Петербург: Питер, 2016. 304 с.

25. Valeriy O. Kuzminykh, Alexander V. Koval, Mark V. Osypenko Methods of Machine Training on the Basis of Stochastic Automatic Devices in the Tasks of Consolidation of Data from Unsealed Sources. CEUR Workshop Proceedings. ISSN: 1613-0073. Vol-2067 urn:nbn:de:0074-2067-8. 2017. P. 63–68.

26. Codd Edgar F. Providing OLAP to User-Analysts: An IT Mandate. *Computerworld*. Vol. 27, No. 30. ISSN 0010-4841.

27. Растригин Л.А., Рипа К.К. Автоматная теория случайного поиска. Рига: Зинатне (ЛатАН), 1973. 344 с.

28. Kuzminykh V.O., Otrokh S.I., Shvets Ye.Yu. Analysis of information in the tasks of consolidating from open sources. *Телекомунікаційні та інформаційні технології*. 2018. № 1(58). С. 70–76. ISSN 2412-4338.

Received 01.03.2021