The architecture of the multy-agent search system of the company

A G Yanishevska, P V Pesterev
Omsk State Technical University, 11, Mira Ave., Omsk, 644050, Russian Federation

Abstract. Fast and convenient search of information on a large number of servers simultaneously to a large number of users at enterprises allows for the use of search engines. Since such a system is complex and consists of a large number of applications, it is effective to use the theory of multi-agent systems to create it. For a more systematic search, we suggest using the methods of ranking the results obtained. The article describes the developed mathematical model for determining the page indicator based on ranking factors for determining the position of the page in search results. In the proposed architecture, the simulation results are processed.

Keywords: multi-agent system, enterprise search system, system reliability estimation, channel load estimation.

1. Introduction

CAD is an automated system that provides the implementation of information technology for the performance of design functions [1], it is an organizational and technical system of personnel and a complex of software, technical or other means of automation of design processes [2]. Any organization or enterprise planning to implement the hardware and software system of CAD [3, 4] should ensure the availability of computers for the implementation of application servers, services and user data. Keeping a large amount of information on servers leads to the need for its fast, systematic and convenient search. To solve this problem, the introduction of a search system [5, 6], which is a computer system for information search, interacting with the user through the interface is effective. Since such a system is complex and consists of a large number of applications, it is effective to apply methods of the theory of multi-agent systems [7-12]. Multiagent system is a system of several interacting intelligent agents (programs) designed to perform tasks for a long time specified by another program or user.

2. Formulation of the problem

As it can be concluded that the creation of an enterprise search system based on a multi-agent system is an urgent task for today, it is proposed to develop the architecture of a multi-agent search system of the enterprise and conduct experimental studies involving the analysis of traffic congestion and calculation of system reliability.

3. Theory

The proposed architecture [13] consists of an interface agent responsible for communication between the user and the multi-agent system; an agent of processing requests, necessary for processing user-entered requests; an ontology agent providing a uniform representation of all agents of the multi-agent system about contained ontologies and storing all information about them; a registration agent that provides registration of new agents and services in a multi-agent system; the coordinator of the working groups interacting with the registration agent by receiving a request from him for the presence or absence of the group in which the new agent or service requests registration; the agent
directory and the service catalog required to centrally store information about all agents and services present in the system, respectively, and to search for and form all agents and services available for interaction; coordinating agents that are the link between the interface agent and a group of search agents related to the same ontology as the coordinating agent itself; search agents, which are a mechanism for accessing MAS to databases and knowledge bases of the enterprise, ensuring the implementation of requests from the coordinators of requests and returning to them the results of their implementation; agents for the processing of results interacting with the coordinating agent and the interface agent and serving to form and rank the results of the SERP according to the degree of relevance and significance, and also calculate the data page indicators based on user actions through the program interface for subsequent iterations of the search. The scheme of the multi-agent information retrieval system at the enterprise is shown in figure 1.

The presented multi-agent information retrieval system, thus, consists of four parts: interface, system-wide applications, search applications and coordinators. The interface part is located on a user computer with a lower bandwidth in comparison with high-performance servers on which other applications reside. This is a consequence of the fact that the servers use wide data transmission channels. Figure 2 presents the general scheme of the architecture of the developed multi-agent information retrieval system at the enterprise. Solid lines show channels with high capacity, dotted lines - channels with an average throughput.

![Figure 1](image1.png)
Figure 2. General scheme of the architecture of the developed multi-agent information retrieval system at the enterprise

The search for information through the presented multi-agent information retrieval system in all databases requires only one request through the interface part to the coordinator, which in this case is the query processing agent. Then the request is redirected to the system-wide part to determine the coordinator agents from the current ontology. And after that they are sent to search applications. The obtained data is returned to the coordinating agents, then to the processing agent of the results with subsequent transfer to the interface part.

If, instead of using a multi-agent information retrieval system, to directly connect to each database, the number of requests will be directly proportional to the number of databases themselves. Accordingly, it can be concluded that the use of multi-agent information retrieval system at the enterprise leads to a decrease in the load on the data transmission channels. More specific calculations can be made only when choosing specific means of transmission and coding of messages.

At the stage of building the general architecture of the application, it is possible to calculate the reliability of the hardware, since there is no need to define a specific implementation of the system. To calculate the hardware, it is necessary to take into account such groups of system equipment with applications as a user computer, servers with coordinators, servers with databases, data transfer channels between them. Reliability of the system is the ability of the object to maintain, within the established time limits, the values of all parameters characterizing the ability to perform the required functions under specified conditions of use, maintenance, storage and transportation. The concept of reliability is associated with the inadmissibility of failures in work. Reliability is the property of the object, which ensures continuous preservation of the operational state for a certain time or operating time. Thus, the probability of failure-free operation can be defined as the likelihood of failure of a technical device under specified operating conditions for a certain operating time. Proceeding from the fact that the probability of failure is an accidental event, we can conclude that the time of the failure is also random. Consequently, the probability of failure-free operation can be expressed:

\[ p(t) = p(t_0 \geq t) \]  

where \( t_0 \) is the time of failure, \( t \) is the specified time of work.

The probability of failure is the opposite event, which means that it can be calculated:

\[ q(t) = q(t_0 < t) \]

The system can contain serial connections, parallel or mixed. Failure of one of the elements in a serial connection leads to the failure of the entire system. Accordingly, the probability of failure-free operation of a system with a serial connection can be expressed:

\[ P_e(t) = \prod_{i=1}^{n} p_i(t) \]  

where \( p_i(t) \) is the probability of failure-free operation of the \( i \)-th element, \( n \) is the number of elements of the system.

In the case of equal probability values of fail-safe operation of the elements of a system with a serial connection, the formula for calculating the failure-free operation can be presented:

\[ P_e(t) = p^n(t) \]

where \( p(t) \) is the probability of failure-free operation of each of the elements.
If there are $k$ consecutive subsystems in a system with a serial connection, each of which has $n$ equiprobable elements, then, based on formulas (3) and (4), derive the following formula for calculating the probability of failure-free operation:

$$P_c(t) = \prod_{j=1}^{k} p_j^n(t).$$  

(5)

where $k$ is the number of consecutive subsystems, $n_j$ is the number of equipotential elements of the $j$-th subsystem.

When a parallel-connected system is running, the presence of at least one work item means the operation of the system. Accordingly, the probability of failure formula can be expressed:

$$Q_c(t) = \prod_{i=1}^{n} q_i(t).$$  

(6)

where $q_i(t)$ is the probability of failure of the $i$-th element.

Proceeding from the fact that the probability of failure and failure-free operation are opposite events, we can conclude that the following expression is true:

$$p(t) + q(t) = 1.$$  

(7)

Starting from formula (7), formula (6) can be transformed:

$$Q_c(t) = \prod_{i=1}^{n} (1 - p_i(t)).$$  

(8)

It follows that the probability of failure-free operation of a system with parallel connection of elements can be represented:

$$P_c(t) = 1 \cdot \prod_{i=1}^{n} (1 - p_i(t)).$$  

(9)

When used in a system with parallel connection of equipotential elements, formula (9) can be transformed:

$$P_c(t) = 1 \cdot (1 - p(t))^n.$$  

(10)

To assess the probability of trouble-free operation of a system with a mixed connection, it is divided into groups consisting of elements with parallel and serial connection, after which the probability of failure-free operation of each of them is calculated.

To simplify the calculation of the reliability of the system, we will assume that the information search starts and ends with the user's computer, and the search is successful in the event that at least one database returns a response. As the data transmission channels in this case, the computer is to sit down with the network equipment. Figure 3 presents a structural diagram of the reliability of the described multi-agent information retrieval system at the enterprise.
For the sake of simplicity in calculating the probability of failure-free operation of the system, we assume that all elements of the system are equally reliable.

Let us first consider the section shown in figure 4.

**Figure 3.** Structural diagram of reliability of the multi-agent information retrieval system at the enterprise

**Figure 4.** The section of the structural diagram of the reliability of the multi-agent information retrieval system

The site consists of three series-connected elements. Hence it follows that the probability of failure-free operation of the section can be represented:

\[ P_1(t) = p^3(t), \]  

(11)

where \( p(t) \) is the probability of failure-free operation of each of the elements of the presented section of the structural diagram of the reliability of the multi-agent information retrieval system at the enterprise.

Next, consider the section shown in figure 5.

**Figure 5.** The site of the structural diagram of the reliability of the multi-agent information retrieval system in the enterprise

The section consists of \( n \) parallel-connected groups of elements, each of which has three series-connected elements. Consequently, the probability of failure-free operation of a section can be described:

\[ P_2(t) = 1 - (1 - p^3(t))^n. \]  

(12)

Next, consider a section with a consecutive connection of three groups of elements (Figure 6).
The data transmission channel
The data transmission channel
The data server
The data server
The data transmission channel
The data transmission channel
The data transmission channel
The data transmission channel
The coordinator
The coordinator
The coordinator
The coordinator
The data server
The data server
The data server
The data server
The database server
The database server
The database server
The database server
The data transmission channel
The data transmission channel

**Figure 6.** Site of the structural scheme of reliability of the multi-agent information retrieval system at the enterprise

The probability of failure-free operation of the plot is calculated:

\[ P_3(t) = p^4(t) \left(1 - \left(1 - p^3(t)\right)^n\right). \]  

(13)

Next, consider the probability of failure-free operation of \( k \) such parallel-connected sections (Figure 7).

The probability of failure-free operation of this section can be calculated:

\[ P_4(t) = 1 - \left(1 - p^4(t)\left(1 - \left(1 - p^3(t)\right)^n\right)^k\right). \]  

(14)

where \( k \) is the number of large parallel sections, shown in Fig. 5.

Thus, the general formula for the probability of failure-free operation of the system is as follows:

\[ P_s(t) = p^6 \left(1 - \left(1 - p^4(t)\left(1 - \left(1 - p^3(t)\right)^n\right)^k\right)^k\right). \]  

(15)

The probability of failure-free operation of the developed architecture can be compared with the probability of failure-free operation of the classical three-tier architecture, depicted in figure 8.
The client

Table 2. The results of an assessment of the reliability of systems based on an experiment

| The probability of failure-free operation of the element p(t), % | The number of databases | The probabilities of failure-free operation, % |
|-------------------------------------------------------------|-------------------------|---------------------------------------------|
|                                                             | The number of coordinators in the system. Architecture is based on multi-agent system | Three-tier architecture |
|                                                             | 2          | 3     | 4     | 5     | 10     |
| 99                                                           | 5         | 94    | 94.14 | 94.15 | 94.15  | 94.15  |
| 99                                                           | 50        | 94    | 94.15 | 94.15 | 94.15  | 94.15  |
| 97                                                           | 5         | 82.2  | 83.17 | 83.28 | 83.3   | 83.3   |
| 97                                                           | 50        | 82.2  | 83.17 | 83.28 | 83.3   | 83.3   |
| 95                                                           | 5         | 70.98 | 73.04 | 73.42 | 73.49  | 73.51  |
| 95                                                           | 50        | 70.98 | 73.04 | 73.42 | 73.49  | 73.51  |
| 93                                                           | 5         | 60.59 | 63.66 | 64.44 | 64.63  | 64.7   |
| 93                                                           | 50        | 60.59 | 63.66 | 64.44 | 64.63  | 64.7   |
| 91                                                           | 5         | 51.16 | 55.01 | 56.23 | 56.61  | 56.79  |
| 91                                                           | 50        | 51.16 | 55.01 | 56.23 | 56.61  | 56.79  |
| 89                                                           | 5         | 42.75 | 47.1  | 48.73 | 49.35  | 49.7   |
| 89                                                           | 50        | 42.75 | 47.1  | 48.73 | 49.35  | 49.7   |
| 85                                                           | 5         | 28.94 | 33.48 | 35.67 | 36.73  | 37.7   |
| 85                                                           | 50        | 28.94 | 33.48 | 35.67 | 36.73  | 37.7   |

Also, to carry out an experimental comparison of the probabilities of failure-free operation of the systems examined, an application was written that simulates them. Disabling of elements was made randomly by means of a random number generator. The results of the experiment are presented in Table 2.

Table 2. The results of an assessment of the reliability of systems based on an experiment

| The probability of failure-free operation of the element p(t), % | The number of databases | The probabilities of failure-free operation, % |
|-------------------------------------------------------------|-------------------------|---------------------------------------------|
|                                                             | The number of coordinators in the system. Architecture is based on multi-agent system | Three-tier architecture |
|                                                             | 2          | 3     | 4     | 5     | 10     |
| 99                                                           | 5         | 94    | 94.14 | 94.15 | 94.15  | 94.15  |
| 99                                                           | 50        | 94    | 94.15 | 94.15 | 94.15  | 94.15  |
| 97                                                           | 5         | 82.2  | 83.17 | 83.28 | 83.3   | 83.3   |
| 97                                                           | 50        | 82.2  | 83.17 | 83.28 | 83.3   | 83.3   |
| 95                                                           | 5         | 70.98 | 73.04 | 73.42 | 73.49  | 73.51  |
| 95                                                           | 50        | 70.98 | 73.04 | 73.42 | 73.49  | 73.51  |
| 93                                                           | 5         | 60.59 | 63.66 | 64.44 | 64.63  | 64.7   |
| 93                                                           | 50        | 60.59 | 63.66 | 64.44 | 64.63  | 64.7   |
| 91                                                           | 5         | 51.16 | 55.01 | 56.23 | 56.61  | 56.79  |
| 91                                                           | 50        | 51.16 | 55.01 | 56.23 | 56.61  | 56.79  |
| 89                                                           | 5         | 42.75 | 47.1  | 48.73 | 49.35  | 49.7   |
| 89                                                           | 50        | 42.75 | 47.1  | 48.73 | 49.35  | 49.7   |
| 85                                                           | 5         | 28.94 | 33.48 | 35.67 | 36.73  | 37.7   |
| 85                                                           | 50        | 28.94 | 33.48 | 35.67 | 36.73  | 37.7   |

The formula for calculating the fail-safe operation of this architecture is as follows:

$$P_c(t) = p^6(t) \left(1 - \left(1 - p^3(t)\right)^2\right).$$

4. Results of experiments

In Table 1 presents the results of a theoretical comparison of the probabilities of failure-free operation of the developed multi-agent information retrieval system and the system based on the classical three-star architecture.

Table 1. Theoretical results of the reliability assessment of systems

| The probability of failure-free operation of the element p(t), % | The number of databases | The probabilities of failure-free operation, % |
|-------------------------------------------------------------|-------------------------|---------------------------------------------|
|                                                             | The number of coordinators in the system. Architecture is based on multi-agent system | Three-tier architecture |
|                                                             | 2          | 3     | 4     | 5     | 10     |
| 99                                                           | 5         | 94    | 94.14 | 94.15 | 94.15  | 94.15  |
| 99                                                           | 50        | 94    | 94.15 | 94.15 | 94.15  | 94.15  |
| 97                                                           | 5         | 82.2  | 83.17 | 83.28 | 83.3   | 83.3   |
| 97                                                           | 50        | 82.2  | 83.17 | 83.28 | 83.3   | 83.3   |
| 95                                                           | 5         | 70.98 | 73.04 | 73.42 | 73.49  | 73.51  |
| 95                                                           | 50        | 70.98 | 73.04 | 73.42 | 73.49  | 73.51  |
| 93                                                           | 5         | 60.59 | 63.66 | 64.44 | 64.63  | 64.7   |
| 93                                                           | 50        | 60.59 | 63.66 | 64.44 | 64.63  | 64.7   |
| 91                                                           | 5         | 51.16 | 55.01 | 56.23 | 56.61  | 56.79  |
| 91                                                           | 50        | 51.16 | 55.01 | 56.23 | 56.61  | 56.79  |
| 89                                                           | 5         | 42.75 | 47.1  | 48.73 | 49.35  | 49.7   |
| 89                                                           | 50        | 42.75 | 47.1  | 48.73 | 49.35  | 49.7   |
| 85                                                           | 5         | 28.94 | 33.48 | 35.67 | 36.73  | 37.7   |
| 85                                                           | 50        | 28.94 | 33.48 | 35.67 | 36.73  | 37.7   |

**Figure 8.** Structural diagram of system reliability based on classical three-star architecture
the element p(t), % | 2 | 3 | 4 | 5 | 10
---|---|---|---|---|---
99 | 5 | 95.98 | 96.06 | 96.06 | 96.06 | 96.06 | 97.03
99 | 50 | 95.98 | 96.06 | 96.06 | 96.06 | 96.06 | 97.03
97 | 5 | 93.67 | 94.08 | 94.11 | 94.11 | 94.11 | 90.37
97 | 50 | 93.67 | 94.08 | 94.11 | 94.11 | 94.11 | 90.37
95 | 5 | 89.66 | 90.29 | 90.34 | 90.35 | 90.35 | 89.44
95 | 50 | 89.66 | 90.29 | 90.34 | 90.35 | 90.35 | 89.44
93 | 5 | 84.15 | 84.87 | 84.94 | 84.94 | 84.94 | 78.12
93 | 50 | 84.15 | 84.87 | 84.94 | 84.94 | 84.94 | 78.12
91 | 5 | 80.63 | 82.19 | 82.44 | 82.48 | 82.49 | 74.94
91 | 50 | 80.63 | 82.19 | 82.44 | 82.48 | 82.49 | 74.96
89 | 5 | 78 | 79.95 | 80.28 | 80.34 | 80.35 | 70.84
89 | 50 | 78.01 | 79.95 | 80.28 | 80.34 | 80.35 | 70.86
85 | 5 | 54.19 | 57.5 | 58.42 | 58.68 | 58.78 | 55.59
85 | 50 | 54.22 | 57.51 | 58.43 | 58.68 | 58.78 | 55.64

5. The discussion of the results
Based on the data presented in Table. 1 and 2, the following conclusions can be drawn:

– Theoretical calculations show that the multi-agent information retrieval system and the system based on the classical three-tier architecture have comparable levels of reliability, while experimental studies show that the multi-agent information retrieval system has a slightly higher level in comparison with the classical three-tier architecture, conclude that the level of reliability of the developed system is slightly higher;

– there is no dependence of reliability of systems on the number of databases;

– the probability of failure of a multi-agent search system becomes less when using three or more coordinators, while the difference in reliability is noticeable between systems with two and three coordinators, but with a further increase in their number, the reliability remains practically unchanged.

These experiments have an approximate result, because the probability of failure of different elements of the system is different in reality. This can be attributed not only to different instances of elements, but also to entire classes. For example, the reliability of servers and computer technology is higher in comparison with the reliability of data transmission channels.

6. Conclusions
The proposed model of multi-agent information retrieval system at the enterprise meets the main features of modern CAD, providing a convenient, systematized search for information in various databases simultaneously to a large number of users. Experimental researches show that the use of multi-agent information retrieval system at the enterprise leads to a decrease in the load on the data transmission channels, and the reliability of such a system is insignificantly higher in comparison with the system with the classical three-tier architecture. At the same time, the reliability of the system increases with an increase in the number of coordinators in it.

7. References
[1] GOST 34.003-90 2009 Information technology. Set of standards for automated systems. Terms and Definition
[2] GOST 23501.101-87 1988 Computer aided design systems. Basic provisions
[3] Golev V 2000 Three components of the CAD system of the enterprise: hardware *CAD and Graphics* **11**

[4] T M Zubkova *et al* 2018 *J. Phys.: Conf. Ser.* **1015** 052031

[5] Segev El 2010 Google and the Digital Divide *The Biases of Online Knowledge*

[6] E Birialtsev *et al* 2017 *J. Phys.: Conf. Ser.* **913** 012010

[7] Gurevich L A, Vakhitov A N 2005 Multiagent systems *Introduction to Computer Science* 116–139

[8] A Gwiazda *et al* 2017 *IOP Conf. Ser.: Mater. Sci. Eng.* **227** 012052

[9] Srinivasan S, Jagjit Singh, Vivek Kumar 2011 Multi-agent based decision Support System using Data Mining and Case Based Reasoning *IJCSI International Journal of Computer Science Issues* **8(4)** 340-349

[10] V Gyurjian *et al* 2008 *J. Phys.: Conf. Ser.* **119** 022025

[11] I V Kovalev *et al* 2016 *IOP Conf. Ser.: Mater. Sci. Eng.* **122** 012018

[12] Xin-Jian Zhang, Ai-Ju Wei, Ke-Zan Li 2016 Successive lag synchronization on dynamical networks with communication delay *Chinese Physics B* **25(3)**

[13] Pesterev P V 2018 Structure and algebraic model of enterprise information retrieval system based on multi-agent system *Information technologies in science and production: materials Vseros. young. nauch.-teehn. Conf* 193-198