ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 11, Issue 11, November 2024

Network Forming on the Basis of Numerous Aspects in Differing Environment

Dmitriy Bystrov

P.G. Associate Professor, Department of Power, Tashkent State Technical University named after Islam Karimov, Tashkent,100095, Uzbekistan

ABSTRACT. The paper presents the architecture of the system to retrieve information for a university environment. The architecture based on multi-agent concept. Main properties of the proposed system are its user orientation and learnability, which is possible by adding new terms to the ontology and relating them with terms, stored in the System. The queries could be presented in the natural language. The System provides transparency of the information resources. The architecture of the system is composed of 3 type agents: user agent, communication agent and information retrieving agents. Communications between the Communication Agent and the other agents are realized via the Communication Module. The module works on the SOAP protocols, which are high level messaging protocols based on XML. The developed system works on a distributed environment, where component agents collaborate via XML Web Services and SOAP protocols.

KEY WORDS: Multi-Agent System, information retrieval, heterogeneous sources, integration of different types information

I. INTRODUCTION

In performing routine tasks, in developing rapport, for decision support and decision making, people spend much time in finding, filtering and processing necessary information. The effort is increasing, if the information is located in heterogeneous sources with different presentation, and access to it requires different methods. To retrieve different types of information from multiple sources it is necessary to create integrated system, where for all types of information, system must be able to process user queries for all types of information, in depending of the context and structure of the queries [1]. The information retrieval system must be transparent for the user. One of the first works on integrated retrieval systems is [2]. The approach proposed in the work provides the flexibility of integrating of any information retrieval model with any type of database management system. This framework makes possible the systematic investigation of the potential of using information retrievals models as a general tool for supporting management decisions. An architecture for an interactive information retrieval system has been proposed in [3]. An integrated retrieval system has two main design problems. First, the enormous amount of information available in multiple sources produces a complex information environment that may be not easy for users to be able to cope with it. Second, problem is different representation of the information in the different sources. Delegating some of these functions to intelligent agent could increase productivity managers, and allows to them concentrating on more creative tasks.

There is growing interest to designing of intellectual agent systems in last years. Some design problems of the information retrieval systems based on agent architecture are discussed in [4-11]. In [4] intelligent agents-software assistants which take care of specific tasks for managers are defined. The design and implementation details of information retrieval system based on multi-agent architecture are presented. A multi-agent architecture [6] provides the basic infrastructure for creation of agents and organizing their communication. Different agents are needed for each different component. The system architecture is elaborated on four types of agents, each of which has a specialized function. These agents are the user, broker, resource and backup agents.

Motivations and issues for a multi-agent approach to personal information retrieval and how those issues are addressed in multi-agent system of Calvin is described in [7]. Calvin's agents individually implement high level abstractions of the information retrieval task, integrating information from different sources. The paper shows that, how personal information agents can have benefit from the open, flexible environment of a multi-agent system. Architecture of the multi-agent system called ISAME is presented in [8]. ISAME is designed for intelligent information retrieval from heterogeneous distributed sources. The system constitutes a virtual library that supplies a set of dynamic information

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 11, Issue 11, November 2024

sources available under electronic formats, as well as services for facilitating and optimizing information retrieval. ISAME support four categories of agents: user agents, information agents and support agents. Main functions of the agents and the communication between the user and the ISAME environment is described. Implementation details of the system are given.

A comprehensive multi-agent architecture for an intelligent information retrieval system with distributed heterogeneous data sources is given in [9].The architecture is composed of five agents, data sources and a user profile base.
The design and implementation details of the multi-agent system are described in [10]. Running on a UNIX platform, the system is implemented in the conventional TCP/IP client-server environment.

An approach for building research models accessing distributed and heterogeneous knowledge resources using multi-agent information retrieval is presented in [11]. Task-oriented scenarios for intelligent information retrieval are discussed. Implementation of the algorithm for decision support process is based on agent technology.

In this paper we propose a multi-agent system to perform the retrieval and integration of information from the operational data bases, document repositories, belonging to an application domain and from the web. Main properties of the proposed System are its user orientation, modularity, extensibility and learn ability, which is possible by adding new terms to the ontology and relating them with terms, stored in the System. Because of the dictionary of the NLP includes about 16000 words, queries could be presented in the natural language without any constraints. The system provides transparency of the information for user.

The paper is organized as follows. In section 2, the architecture of the multi -agent system is presented. In section 3, the multilevel structure of the functional and software components of the system is given. Some realizations details are given in section 4.

II. THE COMPONENTS OF THE MULTI-AGENT INFORMATION RETRIEVAL SYSTEM

The architecture which we have proposed for the Multi Agent Information Retrieval System with distributed heterogeneous data sources is composed of 3 type agents:
User Agent
Communication Agent
Information Retrieval Agent
The role of the User Agent (UA) consists of presenting a user query given in natural language, analysing the query, determining search terms; submitting the results of the retrieval process. The agent provides interactive communication between the user and the System. The User Agent consists of the following components: Web Based Interface Module; Natural Language Processor and Knowledge Base. The agent provides connection with a user via Web Based Interface Module; which receives a query in the NL and sends it to the Natural Language Processor (NLP) and then sends the results of the query processing to the user. The NLP carries out syntactical and partially semantic analysis using own Knowledge base. The Knowledge base consists of the grammar rules of the natural language (we used Turkish grammars) and the dictionary (NLD). As we say above the dictionary may be extended during the System life cycle. The NLP parses a query to set of the separate words. These words are called the set of the search terms, on which the system performs retrieving process. Then the search terms are sent to the Communication Agent.

The Communication agent (CA) carries out interface and transport functions between the agents. The agent includes the following modules: Query Extension Module (QEM), Queue Module (QM), Conversion Module (ConM), Communication Module (CM) and Ontology Scheme (Ont).

The Query Extension Module carries out of fuzzy extension of the search terms. The extended search terms contains all the terms, associated with the user's submitted terms by fuzzy means. The fuzzy associations between the terms are defined by the ontology scheme. Initially, the scheme may be contained small part of this association. But during the new terms are being got by the system, an associations between new terms and existing terms are built. In this way, the ontology scheme is extended dynamically.

Communication between the Communication Agent and the other Agents is realized via the Communication Module. The Module works on the SOAP protocols, which are high level messaging protocols based on XML. To control data

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 11, Issue 11, November 2024

flow between the systems modules, The Queue module is used. The module also is served as buffer for data, waiting its serving. Search terms are identified with own ID. It is allow determining correspondence between the search terms and retrieving information.

The Conversion Module converts information with different representation to unique form (to the XML format).
The Information Agents. There are 3 types of the information agents: web agent, data base agent and document agent. Hearing the network Web Agent, Document Retrieval Agent and Data Retrieval Agent receive message about of retrieval process. Then every agent begins to retrieve information autonomously.
The document agent performs following functions:
Receive and send data and control information from/to the Communication Agent;
Retrieve relevant documents;
Recognize new terms and adding this to the list of index terms,
Make appropriate modifications on the ontology;
Initialize and update local metadata;
Provide data transfers between the general metadata storage and the local document metadata storage.

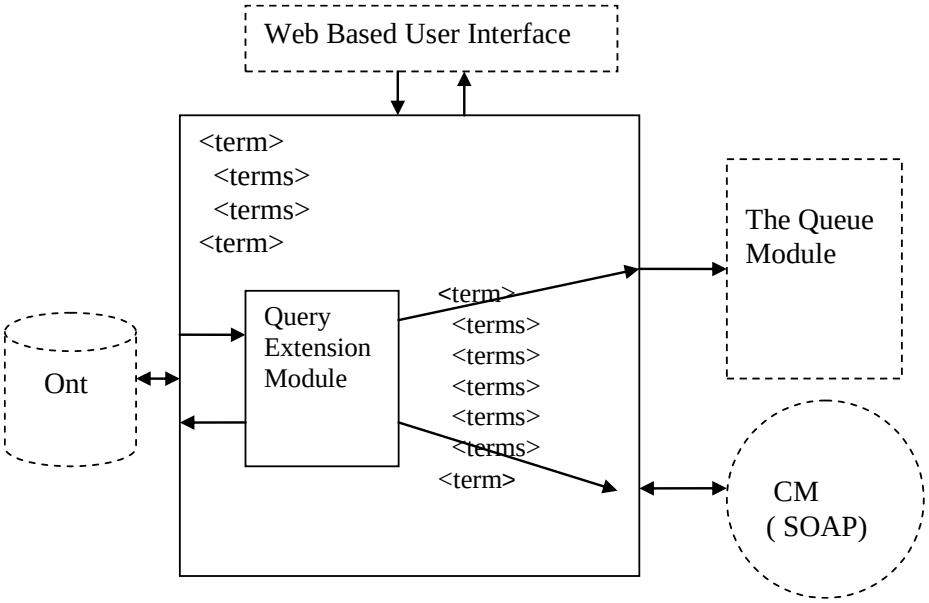

Fig.1. Query Extension Process.

The web agent includes three modules: (Fig. 2)
Selection module
Rank and Filtration Module
Communication Module
The Selection Module (SM). Its possible two ways to retrieve information from the web:
a)Retrieve information on the set of web pages, predetermined by an expert (user) previously. The expert determines more interesting pages, related to an application area. The list of the pages may be updated by the user, if it's necessary.
b) Retrieve information by a search engine (SE) over the all web.
Searching method is determined by a user, when he submits his query to the System.
After the converting the search terms to the appropriate form, Web agent sent the terms to the Goggle Search Engine (SE). Extracted URL's for suggested pages by the Search Engine is passed to the Rank and Filtration Module.
The Rank and Filtration Module (RFM). It is important problem to determine relevantness retrieving pages to the user demands. The goal of the module is determining more relevant web pages among all extracting pages. The Module eliminates low related pages from the list of retrieving pages by using statistical data, which may be formed and updated by the user.

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 11, Issue 11, November 2024

The Communication Module. (CM) Communication Module provides information and control data transfer between the web agent and the agent's environment. The module allows receiving a search terms from the Communication Agent and to send results of the retrieving process or controlling message to CM.
The Data Base agent retrieves data from the operational data bases of the domain area. Receiving the search terms from the mediator agent, the communication agent sends the terms to the metadata block. The Metadata block includes in particularly, the types and locations of all the terms which are used in the database scheme. Three types of the terms have been determined: data base name, table name, column of the table name.
After determining of the term, SQL statements are formed. The SQL statements are implemented with parameter, which are given dynamically.

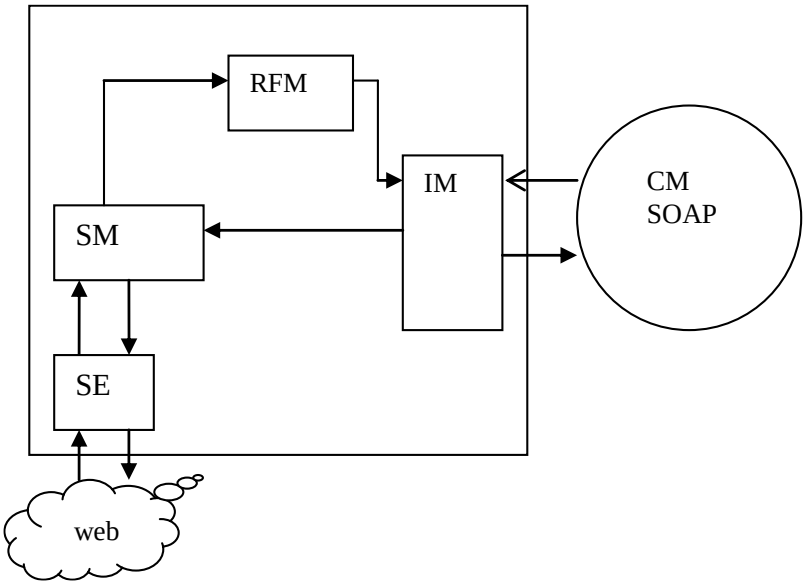

Fig.2. The web agent modules.

Ontology Scheme and its Extension. Ontology is a vocabulary of entities, classes, properties, functions and their relationships. Ontologies are meant to provide an understanding of the static domain knowledge. Ontologies could be useful for: retrieving the appropriate information from documents; integrating the information from various sources; ensuring consistency and correctness by formulating constraints on the content of information; creating libraries of interchangeable and reusable models; supporting inference to derive additional knowledge from a set of facts [12-15]
In our system we use domain ontology, which consists of set of terms frequently used in a university environment. The scheme also includes types of the relations between pairs of the terms and fuzzy grades of the relations. The ontology scheme is represented as two dimensional relation ONT (*rid, st,dt, rt, gr).* Where *rid* is identifier of the relation*; st,dt-* are attributes, describing the first and second elements of a semantic pair, correspondingly. *rt* is a relation 's type between the pair of the corresponding pair. Degree of the semantic relation is described by attribute *gr*.
The following types of the relations are used in the System:
APO- x is a part of y;
SIN- x is synonymy of y;
FR- functional and/or semantic relation between x and y; the relation is determined by the properties of the application area)
AR- x is attribute of y;
SubC- x is subclass of y;
SupC- x is superclass of y;
Below a fragment of the ONT table is given:
The system allows adding new terms to the ontology scheme. If a query consists of terms, which are not in the ontology scheme, the user may be add these terms to the scheme. Adding new term to the scheme is realized by interactive way. According to the list of the clusters, an user determines clusters, which the new word is more closest. Same word may be included to more than one cluster. Second step of the extension process includes the determining the fuzzy closest between the words existing in the ontology scheme and words, are added to his cluster. The user selects the term

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 11, Issue 11, November 2024

(or terms) from the list, and determines the type of the relations and its degree. Such new term find its location in the scheme. This step may be repeated more than one time, if there are exists large number of the words in the cluster.

Table 1.
The Ontology table (the sample).

Rid	St	Dt	Rt	Ggr
00021	Department	Faculty	APO	1
00038	Student	Grant	AR	0.6
00039	Grant	Student	AR	0.8
00120	Undergraduate Student	Student	SubC	1
00121	Undegradute Student	Master student	SIN	1
00200	Course	Lector	FR	1
00301	Salary	Employee	ATR	1

III. SYSTEM ANALYSIS

To understand interrelations of the systems components, it is useful to look the architecture of the system as hierarchy of the components. The system is divided of the following levels of the components:
Information Resources Level. The System allows collecting and processing information from heterogeneous sources. These sources are: Data resources (Operational Data Base, Distributed Data Base Systems, Data Mart or Data Warehouse) of the Organization; Web sources; Document resources (documents, specifications, normative, instructions), which are used in the Organization and might be useful for decision of the management problems.
Information Resources Control Systems Level. Database Management System (or Distributed Database Management System, if distributed data environment is exists) is used to control structured data resources (Data Base). The search browser, like Goggle, may be used to retrieve information from web. But to increase the degree of relevantness of the retrieving document, and don't loss time for retrieving irrelevant information, the list of the internet sites, which have relation to the problem domain are determined previously. Searching on the web is carrying out on this list. The list may be updated time by time.

Network Level. Because of the system works in a distributed environment, data and control information passing between distributed components are realized on the network via TC/IP protocols.
SOAP Level. One of the design problems is to provide data connection with outer modules. Because of development of different modules by different software tools, SOAP protocols, which have high level communication protocols capturing XML, allows solving this problem.

XML level. XML services are responsible for the data transfers between the system modules. Transporting queries to the related modules, formatting the results of the queries in suitable form for the user, generating report on base retrieving information, are also realized with XML services. Query processing on integrated, and aggregated information, coupling different sources are important services of the XML (Fig. 3).
Knowledge base level. Knowledge Base consists of knowledge about language knowledge and knowledge about problem area;

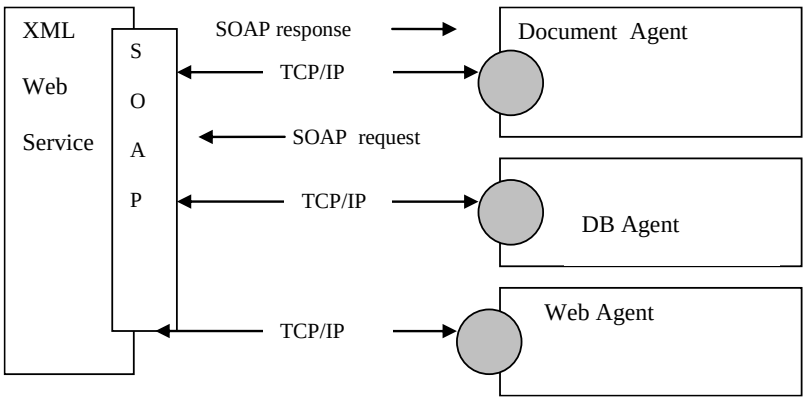

Fig.3. Interagents Communication via TCP/IP protocols.

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 11, Issue 11, November 2024

IV. RESULTS

The Prototype of the System have been developed and tested in the department of Computer Science of the Canakkale OnSekiz Mart University. The mediator agent and the ontology scheme are located in the server computer. The other computers are connected to the system as clients. All agents and users of the system are communicated through TCP/IP links. As server the operation system Windows Server 2003 was used; as clients Windows XP Professional. The following programming languages and development tools were used for designing of the systems models:
Programming Language: Visual Basic 6, ASP.Net, Visual C#.Net, Java, Prolog.
Development Tools: Visual Studio 6.0, Visual Studio.Net, SQL Server 2000, XML, SOAP, Internet Information Services 6.0The characteristics of the server are following: Processor: AMD Athlon XP 32000+, 2.2GHz; RAM: 1 Gb; Network card: 100Mbit Fast Ethernet. The characteristics of the some client machines used are: Processor: Intel Pentium 4, 2GHz; RAM:512 MB;100Mbit Fast Ethernet. The results of the test implementation demonstrated the efficiency of proposed system.

V. CONCLUSION

In this paper the multi-agent information retrieval agent, working in the distributed environment and retrieving information from heterogeneous sources is described. The properties of the system are: a) the system is domain oriented; the terms and relations between the terms are determined on the characteristic of the domain area; b) the system is transparent to the user; user does not know, where the demanding information will be retrieved; c) the system is extensible and learnable- new information resources may be adding to the system without any modifications of the Systems architecture. The system allows adding new terms to the ontology system, updating the contents of the scheme. It provides the learning of the system through its life cycle.
We continue our research on the system on integration of the retrieving information. It will be help to the user by developing of rapports. For this goal, the rapport generation agent must be adding to the System. The module will be work with interaction with the user.

REFERENCES

[1]. Adrian A. Oopgood, *Intelligent Systems for Engineering Scientists*, Second edition, CRS Pres,2000
[2]. David Robins. Interactive Information Retrieval: Context and Basic Notions, *Informing Science*, volume 3,No2, 2000
[3]. Melanie Gnasa, Jens Woch. Architecture of knowledge based Interactive Information Retrieval System, *http://ir.iai.uni- bonn.de/research/ downloads/* konvens-GnasaWoch-2002poster.pdf
[4]. Hyacinth S. Nwana . Software Agents: An Overview. *Knowledge Engineering Review*, Vol. 11, No 3, pp.1-40, Sept 1996.
[5]. Donna S. Haverkamp, Susan Gauch. Intelligent Information Agents: Review and Challenges for Distributed Information Sources. *Journal of the American Society for Information Science;* Apr. 1998;49,4; AI/INFORM Global, pg.3046
[6]. Katia P. Sycara. Multi Agent Systems, *AI Magazine*, 0738-4602-1998,pp.79-92
[7]. Travis Bauer and David Leake. Calvin: A Multi- Agent Personal Information Retrieval System, In Agent Oriented Information Systems 2002; Proceeding of the Fourth International Bi-Conference Workshop,2002
[8]. Sophie –Julie Pelletier, Samuel Pierre, Hai Hoç Hoang. Modeling a Multi-agent System for Retrieving Information from Distributed Sources, *Journal of Computing and Information Technology-CIT* 11,2003,1,1-10
[9]. Neal G.Shaw, Ahmad Mian, Surya B. Yadav. A Comprehensive agent-based architecture for intelligent information retrieval in a distributed heterogeneous environment. *Decision Support Systems*,32 (2002),401-415
[10]. Hui-Min Chen, Design and Implementation of the Agent-based EVMs System, *Multi-agent Based Search Support for Unfamiliar Databases*, June 30, 2000
[11]. A.Aldea et all. An Ontology-Based Knowledge Management Platform, http://www.isi.edu/info-agents/workshops/ijcai03/papers/DIsern-article-ijcai.pdf.
[12]. Nern H.-J., Atanasova T, Pautzke, F., Multi-Agent Approach for Task Related Decision Supported Information Retrieval, *Euromedia 2002*, April 2002, Modena, Italy
[13]. Ki-Young Lee and et all. Adaptive –Document Ontology Construction based on Fuzzy Logic and Fuzzy Function, *Proceedings World Multiconference on Systemics,Cybernetics and Informatics, SCI*, vol. XII, 2001
[14].D.Bystrov and etc. Fuzzy Systems for Computational Linguistics and Natural Language // NISS2020: 3rd International Conference on Networking, Information Systems & Security,2020, Morocco. https://doi.org/10.1145/3386723.3387873
[15].D.Bystrov and etc. Approach of decision making system modeling. *AIP Conference Proceedings* 2552, 060014 (2023); https://doi.org/10.1063/5.0116771