A Decentralized Multi-Agent Ontology-Based System for Information Retrieval

Emmanuel Solidakis[1], Nikolaos Konstantinou[1], Emily-Sirin Pashou[1], Anthi Papakonstantinou[1] and Nikolas Mitrou[1]

esolid@telecom.ntua.gr, nkons@cn.ntua.gr, el00094@mail.ntua.gr, el00027@mail.ntua.gr, mitrou@softlab.ntua.gr

[1] National Technical University of Athens, Division of Communications, Electronics and Information Systems, Heroon Polytechniou 9, 15773 Athens, Greece

Abstract. The search of information in heterogeneous systems is always a crucial task. Continuously changing fields of knowledge offer various standards, methods and technology evolvements to satisfy the users' queries. In this paper we present a catalog based system based upon data collection and harmonization from distributed fields of knowledge that enables the online search and discovery of information from different thematic regions. The key idea is a multi-agent platform that enables the connection and interoperability between independent agent entities, oriented however at offering manageability and accuracy. Users can access data and thematic regions not known to them beforehand by the use of the automatic multi-agent system. Our approach is ontology-based, motivated by scenarios of seeking tourist information.

1 Introduction

Many of the complex problems of real world applications are being solved by distributed agent systems because of the flexibility they offer in design and implementation. The term 'agent' refers to a computer system that is capable of flexible autonomous actions in order to achieve its design objectives [1].

The engineering of agent systems, the development tools, languages, methodologies and standards have evolved significantly, have embodied various fields of research and offer a powerful tool for designing robust solutions. FIPA[1] contributes in the development of agents' standards and has recently (8th June 2005) been accepted by IEEE as its eleventh standards committee. Agent scientists have done a respectable work in using semantics ontologies complied to W3C[2] standards such as RDF(S) or DAML+OIL[3]. Methodologies and standards are offered for creating FIPA/JADE compliant ontologies and assigning them to agents.

[1] The Foundation for Intelligent Physical Agents, http://www.fipa.org
[2] World Wide Web Consortium, http://www.w3.org/
[3] The DARPA Agent Markup Language, http://www.daml.org/

In this paper we introduce a decentralized multi agent system offering dynamic data-independent, semantic level search across ontological units. In chapter 2, related work is cited presenting the state-of-the-art in distributed information retrieval systems. Chapter 3 sets the scenario that our presented system copes with. Chapter 4 elaborates on the architecture of the system and explains the roles of each agent as well as the novelty added to the existing background. In chapter 5, a use case scenario is described while in the last chapter we focus on the future work and conclusions.

2 Related Work

Current tendencies in large online systems providing information often require the collaboration among numerous partners. Systems offering access to voluminous bodies of information such as libraries, biological information centres[4] or geo-portals[5] often consist of a network of data providers. In these systems, the data needs not be gathered in a single server as long as it is shared through a homogenized access mechanism. Users retrieve information by searching data by keyword-based queries. Through portals, users can access federated data in a network whose participants can be from international to local or private agencies. Standards have been developed to ease this process, such as Z39.50 [13] or the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH[6]). To this approach, systems provide access through client applications (harvesters) to network accessible servers (repositories).

Agent systems have been developed to provide infrastructures as above. Despite the fact that agent technology is not widespread to commercial applications [2], major steps have been taken towards this direction [15]. It has been proved that modern agent platforms like JADE[7] are scalable to a large extent, responding efficiently to respectively huge numbers of agents and messages [3,11]. Modern technologies have been incorporated to agent platforms fully exploiting technological novelties [4]. Moreover, peer to peer applications for content management systems [5] indicate that agent oriented solutions are essential to scaleable systems. Knowledge-relative fields as Decision Support Systems (DSS) have used agents as well [6], providing solutions that require huge amounts of data. Methodologies and toolkits have been created for building distributed multi-agent systems [16].

Current agent-based approaches lack in intelligence of the system behaviour. These systems [7,14] either use metadata retrieval based on simple keyword search or exploit RDF schemas. Our contribution is the semantic retrieval of information covering varying and dynamically changing thematic regions.

---

[4] National Biological Information Infrastructure (NBII) Metadata Clearinghouse: http://www.nbii.gov/datainfo/metadata/clearinghouse/

[5] Federal Geographic Data Committee (FGDC) Clearinghouse information page: http://www.fgdc.gov/clearinghouse/clearinghouse.html

[6] Open Archives Forum – OAI-PMH Online Tutorial, http://www.oaforum.org/tutorial/

[7] The Java Agent DEvelopment Framework (JADE), developed by TILAB, available at http://jade.tilab.com/

3 Motivating Example

Our work is set in the area of imposing semantic-level queries to knowledge bases covering different thematic regions. Our aim is to retrieve full and accurate information from distributed content and sites offering geospatial services.

Such systems cannot answer complicated questions such as: ‘What ecology-protected areas are ideal for exploitation from tourism enterprises in northern Greece?’, ‘Where is the suitable place for urban design in Greek islands to prevent from natural disasters and simultaneously meet the best real-estate standards?’ The use of semantic web and ontologies is essential in order to allow intelligent queries in a constantly changing environment and provide precise and right answers to questions like the above. The use of a multi-agent platform is also necessary to provide us with capabilities for searching in a distributed environment. Current systems and implementations permit search among only fixed areas of interest.

4 System Description

In this section we provide the functional description and information on the system architecture and analyze its structural model. Our system includes three types of Agents: Catalogue Agents (CAs), Thematic Agents (TAs) and User Agents (UAs). The connection between the independent modules is shown in Fig. 1 in abstract mode.

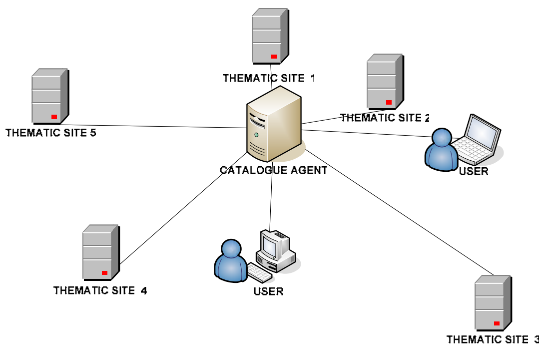

Fig. 1. System Abstract. Note that the system may contain more than one CA.

The CAs contain information about the data and the services offered by the TAs. The role of each CA is to be the intermediate between the UAs and the TAs during the initial search phase.

TAs are located in corresponding Thematic Servers containing content of the specific region that they cover. The UAs are installed in the terminal devices of every user of the system. The device may be a desktop PC, a laptop, a palmtop or a mobile phone and it should be registered to the JADE-agent system. Fig. 2 examines more closely the connection among the system elements.

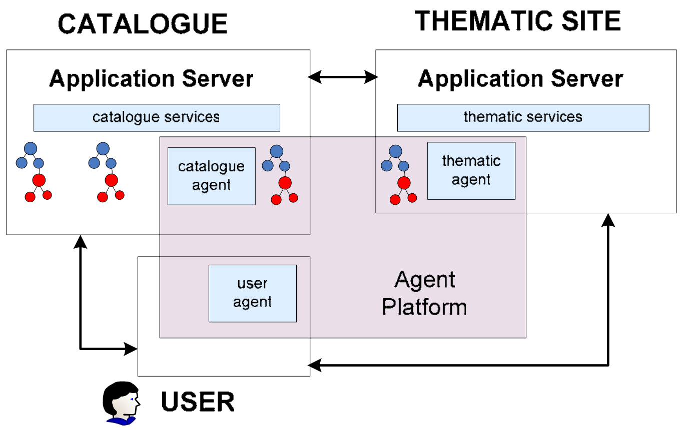

Fig. 2. System Architecture. The basic entities of our system and their interaction

4.1 Catalogue Agents

Every CA offers yellow pages services to the system and has stored locally various profile ontologies for various thematic regions (e.g. Tourism, Cadastre, Marine, Enterprise, Real Estate). Every such thematic region corresponds to one or more TAs, whose ontologies have relations to the ontology of the CA. This is why the ontologies of the CA can be described as core-ontologies of the ontologies of the thematic agents. Thematic ontologies conform to the core ontology and enhance it adding special information varying under each TA's individuality.

4.2 Thematic Agents

Each ontology of the TAs contains localised information about a thematic region. This ontology must be compatible to the corresponding core-ontology that lies on the Catalogue Server. Each service offered by the TA is registered to the yellow pages. A UA can send a query on the thematic area covered by the TA and the TA supplies him with results. Also, a UA can ask for specific ontology classes.

4.3 User Agents

Every UA is initially unaware of the content that can be found in the system. After sending a query to a CA, he receives a list containing the core-ontologies as described by the CA. After choosing among the returned ontologies, the ontology content is sent to the user. Then, the UA creates the dynamic GUI through which dynamic queries can be formed. The queries' structure varies according to the ontology structure. The formed query is broadcasted to all TAs serving this thematic area.

4.4 Dynamic query generation

In current systems, users retrieve data from predefined areas of interest. They can even search among the data but are restricted to specific actions and only keyword search is supported. What is achieved with the implemented system is access to dynamically changing fields of knowledge without prior information of the source or the structure of data. Once the proposed multi-agent system starts graphical interfaces and queries are adapted automatically to users' needs. This goal could not have been achieved without the use of semantics. Current query languages for ontology content (RQL [8], RDQL [9], OWL-QL [10]) are convenient.

5 Use Case Scenario

What is demonstrated below is a scenario about a user who whishes to go on a holiday and would prefer to visit a place near museums of history. User's query is formatted as: "Which place is ideal for holidays considering that there are some historical museums in close vicinity?". Moreover, "Are there any events taking place nearby?".

The user will receive the following GUI containing the thematic regions lying in the ontology of the catalogue services. By selecting the thematic region 'Holidays', the user receives a second GUI in which he can browse and query the contents of the core ontology of the region. In Fig. 3 we see user's main window, the window describing 'Holidays' and a query on the properties of 'Museums'. The user wants to visit historical museums, so he specifies 'History' in the corresponding field.

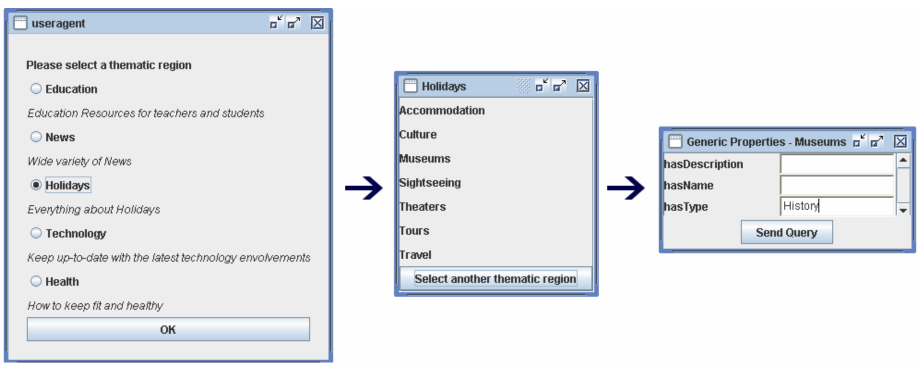

Fig. 3. Stepwise query generation through the graphical interface

The results that the user receives come from various TAs serving the 'Holidays' thematic region e.g. for this region there are two TAs, the Athens agent and the Thessaloniki agent. Thus, the user is informed at once about every historical museum in every agent's contents. The user can also ask about a desired concept in Athens and receive results, as he can do for Thessaloniki.

6 Conclusions and Future Work

We are presently refining the work presented here in several dimensions. First, a key issue that remains open is the synchronization of the ontologies describing the thematic regions. The CAs should automatically update their core- and meta- ontologies when a new TA is added. The same need appears when the existing TAs are updated so that our system remains up-to-date. We also plan to perform checks so that knowledge described in the ontologies should remain consistent. A mechanism of storing the history of the ontologies would be useful in this direction.

References

1. Jennings, N.R, Sycara, K. Wooldridge, M.J.: A roadmap of agent research and develoment, International Journal of Autonomous Agents and Multiagent Systems, 1, pp. 275-306, 1998
2. Jörg P. Müller and B. Bauer: Agent-Oriented Software Technologies-Flaws and Remedies
3. Chmiel, K. et al (2004). Testing the Efficiency of JADE Agent Platform, Proceedings of the 3rd International Symposium on Parallel and Distributed Computing, Cork, Ireland, IEEE Computer Society Press, Los Alamitos, CA, pp. 49-57
4. Y. Wang, L. Cuthbert, F. J. Mullany, P. Stathopoulos, V. Tountopoulos, D. A. Sotiriou, N. Mitrou, M. Senis: Exploring agent-based wireless business models and decision support applications in an airport environment, Journal of Telecommunications and Information Technology, Warsaw, Poland
5. D. Tsesmetzis, E. Solidakis, V. Stathopoulos, N. Mitrou: Distributed Search in P2P Networks through Secure-Authenticated Content Management Systems (CMSs), IEEE Computer Society, Fourth International Conference on Peer-to-Peer Computing (P2P'04)
6. Lawrence A. West, Traci J. Hess: Metadata as a knowledge management tool: supporting intelligent agent and end user access to spatial data, 2001
7. Juan L. Dinos, J. Fernando Vega-Riveros: A Document Ontology and Agent-Based RDF Metadata Retrieval, Nineteenth National Conference on Artificial Intelligence, 2004
8. G. Karvounarakis et al.: RQL: A declarative Query Language for RDF. WWW2002, May 7-11, 2002 Honolulu Hawaii, USA
9. A. Seaborne. RDQL – A Query Language for RDF.
10. R. Fikes, P. Hayes, I. Horroks: OWL-QL - A language for deductive query answering on the Semantic Web
11. "JADE - A FIPA-compliant agent framework", Fabio Bellifemine, Agostino Poggi, Giovanni Rimassa, CSELT internal technical report. Part of this report has been also published in Proceedings of PAAM'99, London, April 1999, pp.97-108.
12. I. Horroks, P. F. Patel-Schneider, Frank van Harmelen: From SHIQ and RDF to OWL the making of a Web Ontology Language
13. Information Retrieval (Z39.50): Application Service Definition and Protocol Specification, National Information Standards Organization, Bethesda, Maryland, USA
14. Nicholas Gibbins, Stephen Harris, Nigel Schatbolt: Agent-based Semantic Web Services, Journal of Web Semantics, Elsevier, 2003
15. J.Dale, S.Willmot and B.Burg: Agentcities: Challenges and Deployment of Next-Generation Service Environments. Proc. Pacific Rim Intelligent Multi-Agent Systems, Tokyo, Japan, August 2002
16. Nwana H., Ndumu D., Lee, L. & Collis J., ZEUS: A Toolkit for Building Distributed Multi-Agent Systems, Applied Artificial Intelligence Journal, Vol. 13(1), p129-185.