Information Retrieval in a Cloud using Ontologies and Multi-Agent System

Mohammed Najah Mahdi[1]

[1]*College of Graduate Studies*

[2]*College of Information Technology*

University Tenaga National –Malaysia[2]

moh12004@gmail.com

Abdul Rahim Ahmad[2]

[1]*College of Graduate Studies*

[2]*College of Information Technology*

University Tenaga National –Malaysia[2]

abdrahim@uniten.edu.my

Roslan Ismail[2]

[1]*College of Graduate Studies*

[2]*College of Information Technology*

University Tenaga National –Malaysia[2]

roslan@uniten.edu.my

Abstract

The existence of heterogeneous systems in the now existing global village brings about a lot of complexity in the search and retrieval of information. However, continuous changing fields of knowledge offer a number of standard procedures and methods made possible by the up spring of technology to help users solve their queries. This research examines the application of knowledge management procedures which perform tasks in a multi agent platform to accomplish user requirements and queries. It highlights how well distributed agent systems help to solve complex problems facing people in the real world as a result of their flexibility. Multi agent systems employ domain ontology to search for pages in the internet which contain relevant information as required by different domains of interest. In the field world wide web in particular, online searching has been embraced by a huge number of people bringing the need of a smooth flowing system that facilitates online activities in a cloud environment. The paper recommends to make use of multi-agent system and ontologies in a cloud environment to enhance the performance of information retrieval.

Keywords: *Information Retrieval, Cloud Computing, Ontologies, Multi-Agent System, Semantic web.*

1. INTRODUCTION

Starting from its origin in the early nineties where it was mainly used to share research documents, the World Wide Web has rapidly grown to encompass diverse information resources ranging from personal pages and accounts, online libraries, markets, government documents, news and research publications. All these pages highlight the fact that internet pages are growing on regular basis in large numbers. It is impossible for a person to retrieve thousands of web pages related to one particular topic, filter the relevant ones, analyze their content and integrate it with regard to their needs [12]. This is why knowledge management tools have integrated tools which automatically update features, monitor and assess how information, services, and technologies are evolving, declining and maturing [22]. According to cambridge university (2009) "Information retrieval (IR) is finding material (usually documents) of an unstructured nature (usually text) that satisfies an information need from within large collections (usually stored on computers)". The ability to search and obtain information from the internet efficiently and effectively enables man to appreciate the existence of ontology's and multi agent based systems [16].

Traditionally, search engines enable users to obtain information by direct input of keyword combinations. However, this method of searching has proven to be rather problematic because the number of pages retrieved after searching a particular concept may not be manageable while others may be irrelevant [5]. Ontology's have proven to be enabling search engines to improve performance in the semantic web. This is because they interweave human understanding of symbols with their machine process ability [22]. They have developed over time in artificial intelligence to facilitate their knowledge base. Typically, ontologies are viewed as representations of concepts that are relevant to particular domains of interest [15]. This is because they provide a semantic view that systematically arranges web pages with required data about concepts in syntactical ways.

2. RELATED WORK

The process of classifying information or information in groups is also one approach that can be used to handle large volumes of information. Many studies in the past have mentioned that the use of ontologies has a significant impact on information retrieval. Qin and Paling (2001) made use of ontologies in their work where they converted the controlled vocabulary of the gateway to educational materials which was mainly to improve semantic information retrieval from the World Wide Web. This shows that the contribution of ontologies has been since 2001 to improve information retrieval from the web.

In [21], ontology-based information extraction is applied to improve the results of information retrieval in multimedia archives, through a domain specific ontology, multilingual lexicons and reasoning algorithms to integrate cross-modal content annotations. This shows that ontologies can be applied to vivid data types rather than just text alone.

3. PROBLEM STATEMENT

The number of online repositories available in the World Wide Web today is skyrocketing and the amount of information available are very high as compared to earlier. This has further increased the need to retrieve product information from these repositories in order to acquire information of their choice. For instance, everybody constantly require information which are reasonable and are of good quality prompting the need to describe them in particular terminologies and languages they are mostly familiar with. However, there exist numerous and diverse descriptions which vendors use but at times do not match up with what people search for increasing difficulty in locating them and eventually this creates more delay in data retrieval. Ontologies have become very crucial in World Wide Web because of their proven usefulness in navigation, user effectiveness and parametric search. Class taxonomies and numerous tools that group ontologies are rapidly growing to fulfill user needs in the process of e-commerce.

4. REVIEW OF CURRENT WORK

Ontology is knowledge represented on the basis of conceptualization that intends a description of object and concept sets and relations between them [19]. On a rather formal perspective, ontology entails organized taxonomy, their definition and attributes. They can be constructed for product information retrieval because product ontology provides the participants in e-commerce activities with a clear and accurate product model which contains rich semantic information, so it can be the basis of the realization of ontology based product information retrieval. This is because product ontology defines the structure of the concept of the product classification strictly, and each concept is defined with all of the corresponding attributes, this makes computers easier to deal with the product data automatically [8]. Ontologies make it possible to search information as they can classify information and ontologies are well known to boost the process of information retrieval and makes it easier for the user to retrieve the information they are looking for. examples of such ontologies are UNSPSC and eCL@ss [4].

Ontologies are now central to a number of applications ranging from scientific knowledge portals, information management systems, integration systems and general internet services. They range from taxonomies, classifications, database schemas to fully axiomatized theories. Web ontology language is a language that was released by the w3c for representing ontology. It is developed from description logic and DAML + OIL and they are developed using integrated, graphical, ontology authoring tools where DAML and OIL are DARPA agent markup language, ontology interface language respectively. DAMN + OIL is a successor language to DAML and (Ontology Interface Layer) OIL that combines these features from both. However, it was in turn superseded by when ontology language (OWL). From this it can be easily concluded that due to the popularity of OWL, it will replace all the other ontologies and become the standard ontology language in the semantic web [1].

4.1 Benefits of ontology and MAS System

- Speed up the process of research making it less time consuming. This ease of access of required information or in this case product for purchase makes it easy for consumers to locate information and services that suit their needs.
- Improves efficiency and effectiveness. The use of ontologies down cut the occurrence of redundancies which limit the ability of to access the information in accordance to their specification.
- Forms a basis of interoperation as they allow distributed but autonomous and heterogeneous resources to function in a world-wide cooperative environment. This is made possible by for example splitting tasks between numerous search teams.
- The use of ontologies also facilitates also the sharing and reuse of knowledge without being task specific [7]. This means that different users can use the same ontology for different purposes without re-defining the existing relationships and properties.

Cloud computing is a model that enables convenient, on demand network access to a shared pool of configurable computing resources such as applications and network servers. These resources can be quickly provided for use by consumers with minimal management efforts as long as they are not needed simultaneously. Multi agent systems are computer systems that have multiple problem solving skills and entities and are made up a number of agents that are able to interact with one another in a cloud environment. An agent is a computer system that can act flexibly in dynamic and uncertain environments. Research in multi agent systems is mainly concerned with the development of methods and algorithms that enable multiple systems to interact and work with one another.

Cloud computing offers three types of services which are Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS) [23]. SaaS delivery model provides the service provider's application running in the cloud infrastructure. The PaaS delivery model provides the service provider's platform to create the application running in the cloud infrastructure. The IaaS delivery model provides the service provider's hardware resources and it may be computational or storage resources. All these services together offer tremendous services which are highly distributed in peer to peer nature.

4.2 Ontology and MAS in Cloud

The utilization of our ontology is projected as a robust of MAS Architecture developed by Protégé software system to tackle the problems of security goals of confidentiality, correctness assurance, availableness and integrity to make sure the protection of CDS. There are 3 main steps, 1) domain, purpose and scope setting, 2) vital terms acquisition, categories and sophistication hierarchy conceptualization and 3) instances creation [17].Multi agent coordination includes a cluster of intellectual agent interacting by means of the atmosphere and with them. As of the figure shown below we are able to perceive that a gaggle of agent's kind a relationship and equally different teams and that they communicate between them.

Autonomous agents can create Clouds smarter within the communication among user along with a lot of economical to allocate dealing out and cargo space to application. Inside large-scale information centres, agents will search, filter, question and revise the enormous volume of information that area unit hold on. Cloud agents operating on our software package can offer intelligent information access services; performance services processor-to-application transfer methods, and energy-efficient use of Cloud International Conference on EGovernance & Cloud Computing Services.

Multi agent systems that area unit capable to handle with ever changing configuration, non-uniformity, and instability, will offer a talented advance for addressing this demand. Security and trust area unit 2 terribly vital problems in Cloud computing. Services provided by the cloud computer system are improved by the utilization of MAS. Agent based models and algorithm for the trust and security in Cloud infrastructure additionally might be terribly helpful.

5. CLOUD COMPUTING

Cloud Computing is a term used to describe both a platform and type of application. As a platform, it supplies, configures and reconfigures servers; the servers can be physical machines or virtual machines. Cloud Computing also describes applications that are extended to be accessible through the internet and for this purpose large data centres and powerful servers are used to host the web applications and web services [3].

Cloud Computing is one of the most talked about technologies in recent times and has garnered attention from the media as well as analysts because of the opportunities it offers. Market research and analysis from IDC insights suggests that the market for Cloud Computing services was $16 billion in 2008 and will rise to $42 billion per year by 2012 [14]. It is estimated that the economic advantages of Cloud Computing results in a three to five times expense reduction for business applications and more than five times for consumer applications [18]. According to a Gartner press release from June 2008, Cloud Computing will be “no less influential than e-business” [13].

Clouds are metaphors for the Internet and are an abstraction for the complex infrastructure it conceals. There are some important points in the definition to be discussed regarding Cloud Computing. Cloud Computing differs from traditional computing paradigms as it is scalable, can be encapsulated as an abstract entity which provides different levels of services to the clients, driven by economies of scale and the services are dynamically configurable [11].

The work related to ontology in a cloud is still limited. Kang and Sim (2011) suggested a system termed as Cloudle where they used ontology in cloud to retrieve information from a search engines. The paper focused more on an enhanced ontology approach whereas this paper suggests using MAS along with ontology in a Cloud.

6. METHODOLOGY

Ontologies are very crucial in providing controlled vocabulary concepts that contain explicitly defined machine understandable concepts and semantics. This is why they are the next big venture in the generations to come as they provide better cooperation between humans and machines. In order to increase the performance of the cloud, it is important to focus on a ranking scheme using vector tags and a sharing scheme using ontological approach for accessing relevant information.

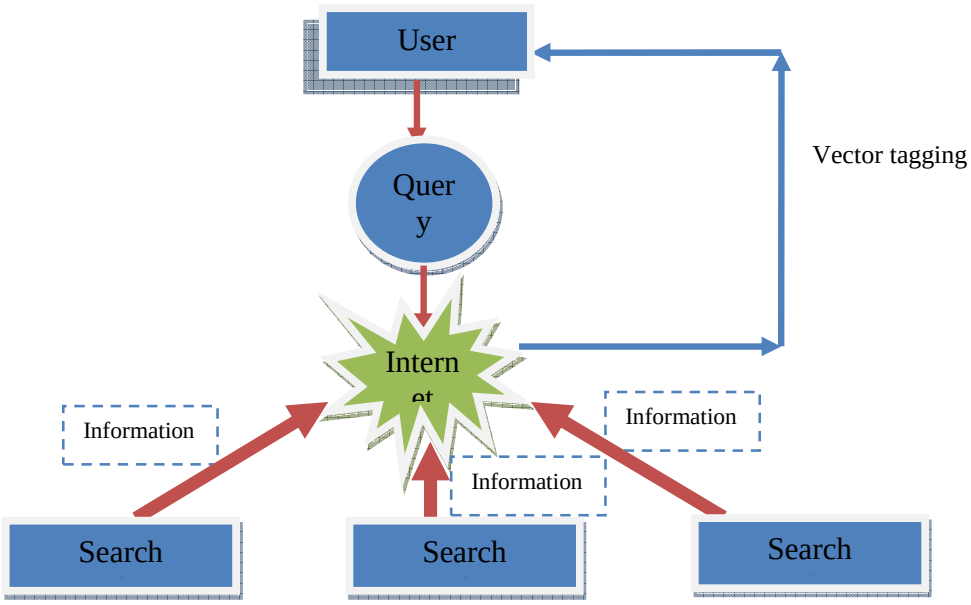

FIGURE 1: *information process into machine process query.*

The process of information retrieval deals with representation, storage and access to information sources. Therefore, information retrieval changes the user's demand for information into machine process able query [10]. The query is presented in a set of terms which describe the type of information required. In the case of e-commerce, when a buyer wants to order goods from a supplier, he or she describes the product type, quality and quantity. This is then followed by price negotiations, confirmation and delivery details. This whole process is made possible by the set ontologies created.

6.1 Multi Agent System Problem Solving Process

The diagram below illustrates how different types of agents are used in a multi agent system to solve a query.

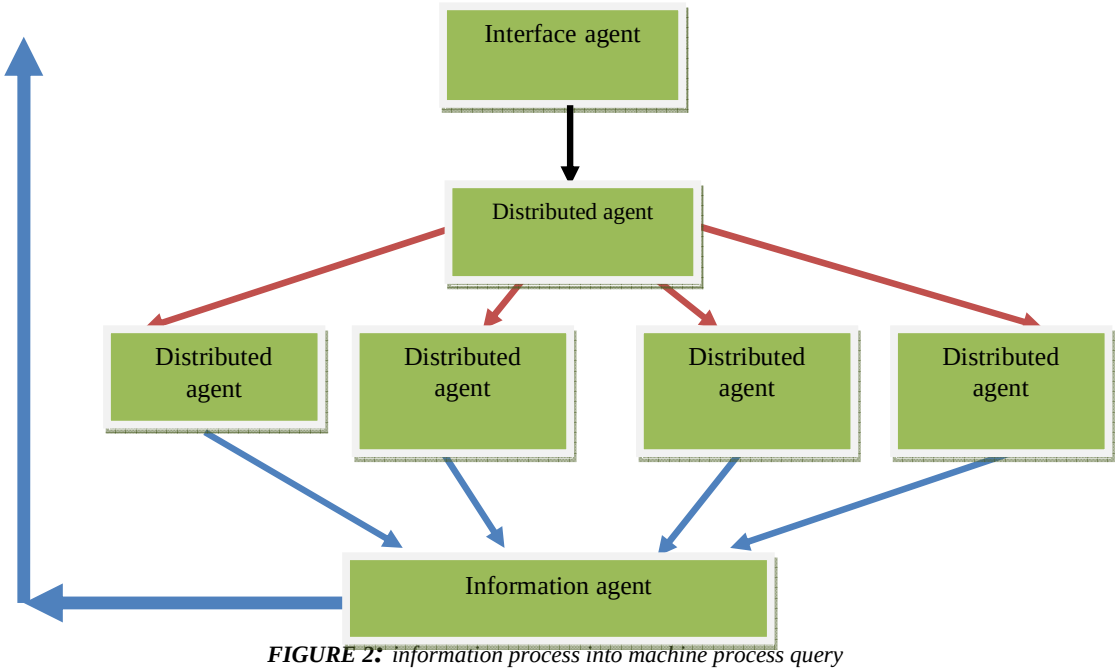

FIGURE 2: *information process into machine process query*

7. FINDINGS

The process of information retrieval starts when a user makes a query through the interface agent which sends requests to distributed agents for information searching [2]. In the case of e-commerce, the process begins with the search of information or services. After data input, distributed agents decompose the data into smaller parts and subparts to various information agents to search the web [6]. Agents then retrieve the requested information from a wide range of product databases and send them to the information agent. The information agent then assembles the information obtained and assembles it correctly before sending it to the interface agent which will later present the solution to the user to answer his or her query.

Multi agent systems make use of ontology for the purpose of intelligent and dynamic information retrieval [9]. In this case, ontology is used to locate and retrieve requested information because it enables different agents in a multi agent system to work cooperatively in a coherent manner to obtain requested information [24]. After information is located, ontology analyses and manipulates it in a way that any redundancy is eliminated to present only valuable and accepted information.

8. LIMITATIONS OF THE CURRENT SYSTEMS

This paper recommends an ontology-enhanced Cloud service search mechanism with the help of Multi agent systems with the aim of consult cloud ontology intended for reasoning about the relations among Cloud service.

The usage of ontology mapping would be good technique to further improve information retrieval which was not covered in this paper.

9. CONCLUSION

The biggest growth in the internet and the area that will prove to be one of the biggest agents of change is World Wide Web. World Wide Web enables people to comfortably gather information in the comfort of their homes. This kind of service also enables dealers to market their information and services over a great volume of people worldwide. However, all these would have not been possible if ontologies were nonexistent.

Ontologies are expressive knowledge models that increase the systems expressiveness and intelligence. This research shows how ontologies can be used in multi agent systems in intelligent information retrieval process. They can be used to support crucial information retrieval processes like posing queries by the user, problem decomposition and task sharing within different agents to ease the process of information sharing, result analysis and the general presentation. Deploying these technologies in a cloud environment would enhance the rate of information retrieval too.

10. RECOMMENDATION

Efficiently composed cloud web services when incorporated with multi agent features can give new form for cloud wide information retrieval procedures. The use of ontology will intelligently enable to search and retrieve information from the cloud environment. By implementing their use, search engines will greatly establish its position on the vast World Wide Web.

REFERENCES

1. Bailin, S.c. Truszkowski, W. (2001), negotiation between agents supporting intelligent information management, workshop on ontologies in agent systems
2. Benabod, R. (2012), Maamri, R, Sahnoun, Z, semantic web service discovery based on agents and ontologies, international journal of innovation, management and technology, vol. 3
3. Boss, G., Malladi, P., Quan, D., Legregni, L. & Hall, H. (2007) . Cloud Computing. IBM HiPODS Report. IBM
4. Cakula, S, Salem, A. (2013), e-learning developing using ontological engineering, WSEAS transactions on information science and applications
5. Calegari, S, Pasi, G. (2010), ontology based information behavior to improve web search, future internet article
6. Cao, l, Bazzan, A Gorodetsky, VMitkas, P, Weiss, G (2010), agents and data mining interactions,springer, Germany
7. Cherifi, H, Zain, J, El-Qawasmeh, E. (2011), digital information and communication technology and its applications: international conference, springer, London

8. Dignum, Y. (2010), trends in practical applications of agents and multiagent system: 8th international conference on conferenece on practical applications of agents and multi agent systems, springer, Berlin
9. Elst, L. Abecker, A. (2013), ontologies for information management: balancing formality, stability, and sharing scope
10. Fensel, D. (2011), Foundations for the web of information and services: a review of 20 years of semantic web research, springer, London
11. Foster, Ian, et al. "Cloud computing and grid computing 360-degree compared." Grid Computing Environments Workshop, 2008. GCE'08. Ieee, 2008.
12. Foud, k, Khalifa, A, Nagdy, N, Harb, H. (2012), web-based semantic and personalized information retrieval, international journal of computer science issues, vol. 9
13. Gartner (2008) Gartner Cloud computing influential Model accessed from www.gartner .com
14. Gleeson, E. (2009). Computing Industry set for a shocking change, http://www.moneyweek.com/investment-advice/computing-industry-set-for-a-shocking-change-43226.aspx,
15. Gutta, A, Sajja, S. (2012), intelligent farm expert multi agent system, international journal on computer science and engineering
16. Jedrzejowicz, P. (2010), Agent and multi agent systems: techmologies and applications, springer, New York
17. Jennings, Nicholas R., et al. "Decentralized Data and Information Systems: Theory and Practice." The Computer Journal 53.9 (2010): 1341-1343.
18. K. Rangan, A. Cooke, M. Dhruv, J. Post, N. Schindler, J. Fidacaro, W. Mohan, G. A. B. III, and J. Vleeschhouwer. The cloud wars: $100+ billion at stake. Technical report, Merrill Lynch, 2008.
19. Nguyen, V. (2011), ontologies and information systems: a literature survey, Australian government department of defiance.
20. Qin, J. & Paling, S. (2001). Converting a controlled vocabulary into an ontology: the case of GEM. Information Research 6(2). Retrieved August 24, 2013, from: http://informationr.net/ir/6-2/paper94.html.
21. Reidsma, D., Kuper, J., Declerck, T., Saggion, H., Cunningham, H. (2003). Cross document annotation for multimedia retrieval, in 10th Conference of the European Chapter of the Association for Computational Linguistics (EACL)
22. Sim, K, Wong, P. (2012), toward agency and ontology for web based information retrieval
23. Sheu, P, Wang, S, Hao,K. (2009), semantic computing, cloud computing, and semantic search engine, international conference of semantic computing.
24. Wongthongtham, P, Chang, E, Dillion, T. (2010), ontology based multi agent systems, springer, Berlin