SEMANTIC NESTED SEARCH ENGINE VIA INTEGRATION ONTOLOGY WITH MULTI-AGENT SYSTEM

Eman K. Elsayed1, AbdAllah A. ElHabshy2, Raghda M. Khaled1

1Department of mathematics and Computer Science, Faculty of science, Al-Azhar University (Girls), Cairo, Egypt. raghdakhaled90@gmail.com.
2Department of mathematics and Computer Science, Faculty of science, Al-Azhar University, Cairo, Egypt. abdAllah@azhar.edu.eg.

ABSTRACT

Global and private search engines retrieve a huge amount of information within thousands of web pages. That is based on user’s search criteria. Nevertheless, it is a challenge to analyze all the information retrieved to match her/his required information. Solving this challenge needs to combine information retrieval systems (IR) with text mining and artificial intelligent approaches. This paper proposes enhanced search engines to be semantic nested search which can be used in business domain. This is via integrating Ontology, Multi-agent technology, and human computer interaction (HCI) concepts. Where Ontology provides a semantic view for understanding the pages contents then Multi-agent collects synchronously the selected information from each link. The proposed semantic nested search engine (SNS) have three phases and use four agents. The SNS engine will be applied on three case studies; jobs search, ecommerce products search and the scientific conferences search. We use different accuracy measurement such as relevant results, recall, precision and F-score. Also, we compare the proposed SNS engine with Google, Yahoo and Bing general search engines. The proposed system consume time. But the run time is not bad relative to the manual searching time inside the links, i.e. overall, the proposed SNS saves users’ time and effort.

Keywords: Multi-agent, ontology, Human Computer Interaction (HCI), Information Retrieve (IR).

1. INTRODUCTION

In recent years, the Internet is used to publish (seek/pursue) information such as jobs offers and ecommerce activities. Internet search engines were created to make retrieving the required information easier. These search engines allow getting a set of Internet pages based on the requested keywords/queries defined by the user. Nowadays, all kind of information can be found on the web, but it is huge, unstructured, scattered, distributed and impossible to analyze manually. Private and global search engines allow retrieving information by combining keywords. Usually, the keywords specified by the user are common and abundant. Subsequently, the user gets too many results, and often some of them are paid search (advertisements). Thus, the user gets thousands of useless pages and sometimes fails to find the information she/he needs. The proposed solution in this paper integrates two main concepts multi-agent and Ontology. The main goals of Multi-agent systems are the cooperation of agents to solve the complex problems and to solve tasks that could not be solved by individual entities or at least to obtain an optimal solution [1]. Also, Multi-agent systems characterized by their speedup, robustness, reliability, low-cost and efficiency [2]. The uses of the multi-agent system technology are to give the system the ability of autonomous and dynamic, to maximize their perceived utility [3].

Ontology is artificial intelligent knowledge representation method. Ontology is represented by a hierarchical taxonomy of concepts. Every concept (class) is connected to a parent concept (super-class) and thus a class and all its ancestors define a class path. Every class contains a set of slots which represent important properties and characteristics of this specific class in the general ontology. Every class also inherits all slots. Ontology is the guideline in defining how data is transferred between the agents in multi-agent systems.

There are many attempts to enhance the search engine to be semantic. In 2017 Sayed and Al-Muqrishi [4] proposed a semantic
search engine (IBRI-CASONTO) for college of applied science (CAS), Sultanate of Oman. This engine is based on the Resource Description Framework (RDF) Ontology and supported both Arabic and English languages. Also, authors in reference [5] tried to support personalized ranking result of searching on the web depending on her/his browsing history IMSS-P (intelligent meta search personalized) tool to improve the performance of the search.

In references [4],[5] and many other research papers which enhanced search engines, the results are just Information Retrieve (IR); i.e. enhances retrieve the most appropriate hyperlinks. While the proposed SNS engine in this paper is nested search work deeply in the content of the links which is identical to the user’s predefined ontology. The multi-agent systems open the hyperlinks and collect the most appropriate contents to the user.

Constructing nested multi-agent search based on ontology handles the problems of distributed information. Also, generating nested multi-agent based on specific Ontology shall present only relevant information to users' request. This paper integrates multi-agent with Ontology to fast and accurate search. Moreover, Ontology helps Agents to get more automation. This paper also describes in detail three case study from human computer interaction (HCI) perspective; that improve web search engines using predefined ontology through nested multi-agent search; “jobs search”, “e-commerce product search” and “scientific conferences search”.

The reminder of this paper is organized as follows. Section 2 introduces related works. Section3 presents relation between Multi-agent and Ontology. Then the proposed method architecture of semantic nested search engine (SNS) presented in section 4. Section 5 applies the proposed method on three case studies. Analysis of the experimental results is presented in section 6. Finally, Section 7 presents the conclusion and future works.

2. Related Works

Recently, simple general information retrieval (IR) tools based on keywords matching algorithm is not enough. There are two main classes of methods for the information retrieval. The old one treats query as a document and applies feature extraction methods on it [6], while the new methods based on associate the queries with the results. Authors in reference [7] introduce approach that uses ontology derived for the Indian tourism domain for the efficient information retrieval. The performance of the search engine is improved by using Ontology. The relevant results satisfy through the semantics of the query.

Also, reference [8] proposed Multi-agent system named (SCKE) (Statistics and Collaborative Knowledge Exchange) that allows the retrieval and exchange of information between hospitals. The proposed system uses semantic search technique to provide results that are more relevant and a semantic retrieval system for disease information. Ontologies play an important role in achieving interoperability across organizations and on the Semantic Web.

In [9] authors introduced an approach called Secured Cognitive Agent based Multi-Strategic Intelligent Search System (CAbMsISS) to help the user in search process. It helps to reduce the contextual and emotional mismatch between the search engine and user. Cognitive Agent based Multi-strategic Intelligent Search System (CAbMsISS) framework improves Query Retrieval Time (QRT) and retrieving relevant results as compared to Present Search Engine (PSE). It can give suggestions to the user when he/ she accesses a resource previously tagged with negative emotions. The target of the system is to enhance the search experience. The suggestions of the system are introduced through the search log that tracks the queries searched, resources accessed, and emotions experienced during the search. The implemented framework also considered secure.

Reference [10] presented a search engine for performing semantic search in news. In addition to the use of Semantic Web techniques that enhance the efficiency of search operations. The Mean Average Precision (MAP) of their framework is higher than Google news and Yahoo News. Furthermore, the event Ontology increase the Average Precision by 0.8 percent.
compared to the other traditional search engines.

Most researches in Information Retrieval try to enhance the search runtime and relevant results, but still the user needs to open each link to get the specific result. So, this paper proposes a new semantic nested search engine to enhance the search results besides collecting the relevant data from different links and display it to users.

3. Agents and multi-agent systems

An agent is a software program that automatically performs tasks on behalf of the user. Also, an agent have the ability of exchange or communicate their own knowledge [1]. Agents can communicate with each other in many various ways including exchanging messages and working together to solve complex problems [2][3][4]. They can understand semantic data in the environment. They can do actions, which are described by the predefined ontology which establish agent’s knowledge base [4][5][6].

The Agent technology has the ability to decide what they need to do in order to achieve their objectives [10]. The main goals of Multi-agent systems are the cooperation of agents to solve the complex problems and to solve tasks that could not be solved by individual entities or at least to obtain an optimal solution. Also, Multi-agent systems characterized by their speed up, robustness, reliability, low-cost and efficiency.

In this paper, ontology and multi-agent working in a cooperative way to achieve the ultimate of their purposes in the same environment. Ontologies help Agents to have a higher degree of automation [2]. While the predefined ontology is a guideline which defines how the data is transferred between the agents.

4. The Proposed Semantic Nested Multi-Agent Search Engine

Semantic Nested Multi-Agent search provides a solution for search engines to save users’ time and effort. When user searches for a keyword in normal search engines the result will be a lot of pages and links [11]. To get the required information, the user must open each page then open each link, so she/he will lose a lot of time and effort. While with semantic nested search, the user sets an ontology which describes the search result she/he needs, then the agents search on the web, filtering the retrieved information according to the predefined ontology. Then, the required information will present to the user as a web page without any links or further information. The main purpose of using multi-agents is to search and analyze the retrieved data according to the predefined ontology to give the user the most appropriate search result. Moreover, the proposed implementation has the feature of the Human Computer Interaction (HCI) where the user can set her/his personal criteria to make the search result more accurate. The search process was implemented in JADE (Java Agent Development Framework) environment, four different agents were identified. When user defines the ontology and search keywords, the multi-agent shall be started and process the query. This task combines several stages: retrieving the web pages, save links, gets the contents of each link, and finally the user gets the search result. The result should be identical to reflect the predefined ontology. This technique has many benefits such as, reducing the number of pages, collecting information based on ontology classifications criteria, reducing time of normal search and one document for all search results. This section presents the architecture of the proposed Semantic nested search (SNS) system. The agents in SNS Semantic nested search are Four. Their names are search agent, saver agent, content agent and Ontology agent. The Semantic nested search has three phases as shown in Fig.1.

First Phase: This phase is called a parallel search phase. Once the predefined suitable Ontology for search field according to user profile and search keywords have been defined by the user, the multi-agent search process will be started. The search agents search queries and retrieving the web pages. So, search agent automatically creates agents in this phase according to number of web pages fig. 2.

Second Phase: this phase is called nested phase. In this phase SNS use automatically two agents; saver agent saves links and content agent gets the contents of each link in the first phase fig. 3. The agents (search, saver, and
content) start work and send information to each other.
Fig. 1 The architecture of Semantic Nested Search SNS three phases

Fig. 2. SNS create agents according to number of pages

Fig. 3. part of code of the content Agent
**Third Phase:** this phase is called semantic phase. In this phase Ontology agent merges information comes from contents agent with the predefined ontology using ontology agent.

As a result, for every search query/keyword a set of web pages are retrieved. The SNS GUI is shown in fig. 4.

Consequently, the saver agent stores the retrieved pages. Subsequently, the content agent processes each link to get its contents and pass it to ontology agent. Finally, Ontology agent processes the contents of each link using the predefined Ontology then pass the result to the user. That is according to user’s profile to respect human computer interaction (HCI) perspective. Where, the user can set his personal information in Ontology class to get result more appropriate and interact. For example, when the user search for a job, she/he can set his personal information such as experience years, age, education, etc. So SNS gives the result which is suitable for her/him only.

5. Implementation

The creation of Semantic Nested Search system (SNS) needs

- JADE “Java Agent Development framework” which is open source for developing multi-agent systems.
- Protégé as a general Ontology editor.
- Jena API.

The case studies present in this section need three different predefined OWL-Ontology as “Products Ontology. owl” for Ecommerce search, “Vacancies Ontology. owl” for Jobs search, and “Conference. owl” scientific conferences search.

In the following subsections the proposed SNS engine will be applied on three case studies.

5.1 Jobs Searching Process

When the user searches for a job in general search engines the result will be a lot of pages and links [11]. Although the user may use a jobs’ search engine, but still she/he should open links and collecting data. To get the required information, the user must open each page then open each link; i.e. she/he will lose a lot of time and effort. But with the proposed semantic nested search the user inserts her/his requests to Ontology and insert the keyword in the GUI (Graphical User Interface) then the agents search on the web, filtering the retrieved information according to the Ontology and give the user the result as web page without any links or further information.

The user writes the keyword in the GUI. The result will be presented as a document that contains all jobs that identical to the structure of the predefined Ontology as shown in Fig.5 and Fig.6.

5.2 Searching Process for Online Products

![Fig. 4 SNS graphical user interface](image)

![Fig. 5. Part from Lightweight Ontology “VacanciesOntology.owl”](image)
Dealing with E-commerce (Electronic commerce) means dealing with enormous types of products [12] and each product may have different prices on different web sites.

If the user wants to know specific information about a product such as its price, description and Country of Origin (without any further information), then the user just inserts her/his requests to “Products Ontology.owl” as shown in Fig. 7, which clarifies the structure of the ontology and search for certain product. Then, the SNS engine presents the required result in a web page according to the predefined ontology as shown in Fig.8.

5.3 Searching Process for Conferences

Around the world, the conferences are important scientific events. Scientific researchers need to know just certain information as the address of the conference and its important dates. When the user searches for conferences in normal search engines the result will be a lot of pages and many links with each page. To get what she/he is looking for, the user must open each page then open each link; i.e., she/he will lose a lot of time and effort.

Utilizing the semantic nested search (SNS) engine gives the user the ability to insert her/his request to ontology “Conference.owl” then the agents search on the web, filtering the retrieved information according to the Ontology class and give the user the result without any further information.

When a user searches for certain conference field, the SNS retrieves only the required information which suitable for him as shown in Fig.9.

6. Analysis of result

In this section we compare among our proposed search engine prototype SNS and the three famous general search engines Bing, Yahoo and Google. Where in the first step of SNS engine we use Google and then enhance automatically his result by nested and semantic phase in SNS engine.

Tables 1, 2 and 3 summarized the different between SNS and Google, Bing, Yahoo respectively.
Fig. 8. The SNS result for certain product for certain person

Fig. 9. SNS results for certain request from “conference. owl”
Although in SNS engine the search time is relatively long, but SNS engine retrieves relevant links and support nested data collection as shown in fig. 10. Where, one of the important aspects of the search engine evaluation is the relevance judgment of the results according to the queries. Consequently, the proposed SNS saves users’ time and efforts.

![Fig.10. The graph of the comparison between SNS, Bing, Google and Yahoo](image)

Also, we calculate the accuracy for our proposed search engine prototype SNS. That is by calculating (Recall, Precision and F-score) functions. The average number of true positive is 164, the average number of possible is 189, and the number of false positive is 26.

Where, recall function is the fraction of relevant retrieved result over the total possible ones.

\[
\text{Recall} = \frac{\text{No of true positive}}{\text{No. of possible}} \times 100\% \quad (1)
\]

While precision function is the fraction of relevant retrieved instances over the no. of relevant and irrelevant retrieved.

\[
\text{Precision} = \frac{\text{No. of true positive}}{\text{No. of true positive + No. of false positive}} \times 100\% \quad (2)
\]

And F-score or F-measure is a measure of test’s accuracy. It is harmonic average between Recall and Precision.

\[
F\text{-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)
\]

Accordingly, table illustrates the precision, recall and F-score for Google, Yahoo, Bing and SNS search engines.

As shown in Table 4, the proposed SNS engine is more efficient and accurate than well know search engine.
7. Conclusion and Future works

This paper proposed a new efficient Semantic Nested Search (SNS) engine, which is a prototype of semantic multi-agent search engine. SNS’ phases used four agents. It has permission to access to the retrieved results for searching inside links. SNS engine was implemented using Java programming language which offers access to JADE (Java Agent Development Environment) and Jena libraries. Also, Protégé editor was used to create and handle Ontology engineering part of the proposed system. The evaluation of the proposed SNS engine was examined using different factors and it showed a high accuracy rate. Although in SNS engine the search time is relatively long, but SNS engine retrieves relevant links and supports nested data collection, which saves the users’ time and efforts.

In the future, this semantic engine could be improved by using text and opinion mining techniques. In addition, a good interface and voice recognition mechanism should be implemented to this search engine. Moreover, a mobile application for this search engine should be created, which should be suitable and easy to use. Furthermore, using heavy universal Ontology which merged with user’ history leads to more relevant results.

Data Availability

Data used to support the finding of this study are available from the corresponding author upon request.

Conflict of Interest

Authors declare that they have no conflicts of interest.

ACKNOWLEDGMENTS

This research was supported by Faculty of Science, Alazhar University (girls branch).

REFERENCES

[1] Zou Y, Finin T, Ding L, Chen H, Pan R. Using Semantic Web Technology in Multi-Agent Systems: A Case Study in the TAGA Trading Agent Environment. Proc 5th Int Conf Electron Commer 2003:95–101. doi:10.1145/948005.948018.

[2] Benaissa Ezzeddine, Benaissa Mounir, Benabdelfatif Abdellatif. Multi-Agent Technique and Semantic Web Services for Extended. International Journal of Intelligent Information Processing, 2013;4:40–8. doi:10.4156/ijiip.vol4.issue1.5.

[3] García-Sánchez, F., Martínez-Béjar, R., Valencia-García, R., Fernández-Breis, J.T.: Knowledge Technologies-Based Multi-Agent System for Semantic Web Services Environments. In, Berlin, Heidelberg 2008. Artificial Intelligence and Soft Computing – ICAISC 2008, pp. 1222-1233. Springer Berlin Heidelberg”. Please, visit “https://link.springer.com/chapter/10.1007/978-3-540-69731-2_114.

[4] Sayed A, Al Muqrishi A. IBRI-CASONTO: Ontology-based semantic search engine. Egypt Informatics J 2017;18:181–92. doi:10.1016/j.eij.2017.01.001.

[5] Malhotra D, Rishi OP. IMSS-E: An intelligent approach to design of adaptive meta search system for E commerce website ranking. ACM Int. Conf. Proceeding Ser 2016;12–13‧Augu. doi:10.1145/2979779.2979782.

[6] Yin, Zhijun, Milad Shokouhi, and Nick Craswell. "Query expansion using external evi-dence." European Conference on Information Retrieval. Springer, Berlin, Heidelberg, 2009.
SEMANTIC NESTED SEARCH ENGINE VIA INTEGRATION ONTOLOGY..... 37

[7] Laddha SS, Koli NA, Jawandhiya PM. ScienceDirect Indian Tourism Information Retrieval System: An Onto-Semantic Approach. Procedia Comput Sci 2018;132:1363–74. doi:10.1016/j.procs.2018.05.051.

[8] Naseebah H. Alkahtani, Sarah Almohsen, Nasibah Mohammed Alkahtani, Ghadah Abdullah Almalki, Safaa Samir Meshref, Hebah Kurdı, A Semantic Multi-Agent system to Exchange Information between Hospitals, Procedia Computer Science, Volume 109, 2017, Pages 704-12

[9] Gulati N, Garg A. A Secured Cognitive Agent based Multi-strategic Intelligent Search System. J King Saud Univ-Comput Inf Sci 2018;30:206–22. doi:10.1016/j.jsuci.2016.06.005.

[10] Beheshti H, Poorahangaryan F. NewsSE: An Ontology-based Search Engine for News. NewsSE: An Ontology-based Search Engine for News 2017. doi:10.13189/csit.2017.050201.

[11] Zhao, Bo. (2017). Web Scraping. 10.1007/978-3-319-32001-4_483-1.

[12] Išoraitė, Margarita & Miniotienė, Neringa. (2018). Electronic Commerce: Theory and Practice. IJBE (Integrated Journal of Business and Economics). 2. 73. 10.33019/ijbe.v2i2.78.