Impact Factor Value: 4.029 ISSN: 2349-7084
International Journal of Computer Engineering In Research Trends
Volume 4, Issue 7, July-2017, pp. 296-300 www.ijcert.org

Searching of Web Data Using Ontological Matching

Shweta Shrivastava*[1], Dr. Sanjay Agrawal*[2]

[1]AISECT University , Bhopal Ph.D. scholar Computer Sc. Department,
[2] NITTTR , Bhopal Prof Computer sc. Department.
Email ID: shwetashri.26@gmail.com[1], sagrawal@nitttrbpl.ac.in[2]

Abstract: While phenomenally successful regarding size and number of users, today's World Wide Web is fundamentally a relatively straightforward artifact. Web content consists mainly of distributed hypertext and hypermedia and is accessed via a combination of keyword based search and link navigation. The explosion in both the range and quantity of web content has, however, highlighted some serious shortcomings in the hypertext paradigm. Every year, the number of documents on the Internet is increasing, presenting the correct information at the right time in the most appropriate form is important, and it results in better browsing experience for users. To deal with this issue, ontology's are proposed for knowledge representation, which is nowadays the backbone of semantic web applications. This is a challenging task as it requires complex queries to be answered with only a few keywords. Furthermore, it should allow the inferred knowledge to be retrieved easily and provide a ranking mechanism to reflect semantics and ontological importance. Proposed paper gives. A technique to improve the efficiency of matching web data with background knowledge.It finds correspondences between semantically related entities of ontology.

Keyword—Semantic Web, Ontology, Information Retrieval, inferred knowledge, ranking mechanism, Web Search.

I. Introduction

Semantic Web

The Semantic Web [14] is an extension of the existing World Wide Web. It provides a standardized way of expressing the relationships between web pages, to allow machines to understand the meaning of hyper linked information. The Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined meaning, better-enabling computers and people to work in cooperation." It is a source to retrieve information from the web.Today's search engines cannot search more precise that they do it now. May be the main reason that the structure and size of current Web do not allow to make search more precise and efficient. The second reason Web contains now a huge number of documents, and this number has a strong tendency to double each one or two years. The structure of documents and Web itself, probably, can be changed in "a better – machine process able way. "Semantic web is being to be developed to overcome the following problems for current web.-

- The web content lacks a proper structure regarding the representation of information.
- The ambiguity of information resulting from the poor interconnection of information.
- Automatic data transfer is lacking.
- Usability to deal with an enormous number of users and content ensuring trust at all levels.
- Incapability of machines to understand the provided information due to lack of a universal format.

The current studies on semantic query interfaces are carried in four categories, namely, keyword-based, form based, view-based and natural language-based systems. Out of these, keyword-based query interfaces are the most user-friendly ones, and people are already used to use such interfaces easily, This is a challenging task as it requires complex queries to be answered with only a few keywords. Furthermore, it should allow the inferred knowledge to be retrieved easily and provide a ranking mechanism to reflect semantics and ontological importance. Ontology typically provides a vocabulary that describes a domain of interest and a specification of the meaning of terms used in the vocabulary. Depending on the precision of this specification, the notion of ontology encompasses several data and

© 2017, IJCERT All Rights Reserved

conceptual models, including, sets of terms, classifications, thesauri, database schemas. These correspondences can be used for various tasks, such as ontology merging, query answering, or data translation. The remaining sections of the paper are as follows. Section II makes readers aware of Ontology and its importance in Semantic Web.Section III defines different query languages and advantages of our technique. Section IV gives what work is done in semantic web searching.Section V gives proposed methodology. Finally, Section VI concludes our work with its future scope.

II. Ontology

Ontology is an explicit specification of a conceptualization. Ontology is a body of knowledge describing some domain, typically common sense knowledge domain. In both computer science and information science, an ontology is a data model that represents a set of concepts within a domain and their relationships between those concepts. It is used to reason about the objects within that domain. Ontologies are used in artificial intelligence, the semantic web, software engineering, biomedical informatics and information architecture as a form of knowledge representation about the world or some part of it. Ontology describes:

- **Individuals:** the basic or "ground level" objects.
- **Classes:** sets, collections, or types of objects.
- **Attributes:** properties, features, characteristics, or parameters that objects can have and share.
- **Relations:** ways that objects can be related to one another.
- **Events:** the changing of attributes or relations.

Ontology and Semantic Web Mining

Ontology is used to model real word entities and relations among them in a taxonomic structure. They are nowadays the backbone for the Semantic Web applications. Several languages are developed for the formal representation of ontology. RDF Schema (RDFS) was the first attempt towards developing an ontology language. RDFS was built upon RDF. It extends the RDF [27] vocabulary with additional classes and properties such as rdfs:Class and rdfs:subclass Of. The latest W3C recommendation for ontology languages is the Web Ontology Language (OWL). OWL[16] further extends RDFS by providing additional features such as cardinality constraints, equality, disjoint classes, efficient reasoning support and much more. OWL language has three sublanguages, which are OWL-Lite, OWL-DL, and OWL-Full in the order of increasing impressibility. OWL-Lite and OWL-Full are not widely used because the former is too restricted and the latter does not guarantee efficient reasoning. OWL-DL[20] provides maximum expressibility with a complete and decidable reasoning support. Semantic Search [13] By representing the data in RDF or OWL format, the Semantic Web allows more intelligent search engines to be developed. These search engines can make use of the metadata associated with the entities to improve search quality. Semantic relations defined in ontologies allow very complex queries to be answered which are not possible otherwise. For example, semantic search engines can easily answer queries.

III. Query Retrieval Languages

Different type of Query Retrieval Languages

1. RDF The Resource Description Framework (RDF) [27] is a family of World Wide Web Consortium (W3C) specifications originally designed as a metadata data model. It has come to be used as a general method for conceptual description or modeling of information that is implemented in web resources. An RDF query language is a computer language, specifically a query language for databases, able to retrieve and manipulate data stored in Resource Description Framework format.

2. RDQL RDQL [8] was developed by Hewlett Packard and submitted to the W3C in January 2004. RDQL is a query language for RDF in Jena models. The idea is to provide a data-oriented query model so that there is a more declarative approach to complement the fine-grained, procedural Jena API. It has been implemented in a several RDF systems including Jena, RDFStore, Sesame, PHP XML Classes, 3 Store and RAP-RDF API for PHP. RDQL was derived mainly from SQUISH, an earlier language. The syntax of RDQL is similar to an SQL-like select pattern where the select clause allows the projection of the variables.

3. SPARQL [9] is a query language designed specifically to query RDF databases. SPARQL(Simple Protocol and RDF Query Language) queries sent from a client to a service known as an SPARQL end point using the HTTP protocol. The interaction between the client and the endpoint is defined in a machine-friendly protocol that is not intended to be interpreted by humans, so the use of SPARQL requires an interface that allows the user to enter the queries and to display the results in a meaningful way. As with traditional database languages such as SQL, those interfaces are commonly constructed so that the queries are built and launched through forms that do not require the human user to have any knowledge of RDF and SPARQL.

© 2017, IJCERT All Rights Reserved

4. SeRQL ("Sesame RDF Query Language," pronounced "circle") is a new RDF/RDFS query language that is currently being developed by Aduna as part of Sesame. It combines the best features of other (query) languages (RQL, RDQL, N-Triples, N3) and adds some of its own. This document briefly shows all of these features. After reading through this document, one should be able to write SQL queries.

IV. Background

The Web Search Engines methodologies, following recent years, is always less uncommon that the results provided by them are a greater burden of useless pages to the users. The third generation Web architecture the semantic web provides the three layered architecture possibly allowing overcoming the limitation and burdens that caused in existing searching methodology. There are several search engines have been proposed and adopted, their main contribution is increasing information retrieval accuracy by exploiting a key content of Semantic Web resources, mainly based on relations between the concepts, however, in order to rank or prioritise results, most of the existing search solutions need to work on the whole annotated knowledge database. Ontology typically provides a vocabulary that describes a domain of interest and specification of the meaning of terms used in the vocabulary. When several competing ontologies are used in different applications, most often these applications cannot immediately inter operate. OWL is succeeding to a great extent as a knowledge representation standard for instances used for building knowledge system.

Ontology matching can take advantage of linked data as an external source of information for ontology matching this is entirely relevant to the "Matching with background knowledge" challenges.OAEI companions gave some preliminary evidence of the scalability characteristics of the ontology matching technology. Various strategies have been used to deal with the lack of background knowledge.

[1] Using the web as background knowledge and specifically, explaining linked data as background knowledge or the work on search engine weighted approximate matching.

[2] Using domain specific ontologies e.g. in the field of anatomy upper-level Ontologies or all the ontologies available on the semantic web.

V. Proposed Work

In proposed methodology, present a complete ontology-based framework for the extraction and retrieval of semantic information in limited domains. We applied the framework in a specific domain and observed the improvements over classical keyword-based approaches. The system consists of an automated information extraction module, an ontology library module, a cluster module, and a keyword-based semantic query interface, a ranking module. Our main concern, in this study, is achieving a high retrieval performance while preserving the user-friendliness. We show that our system can handle even the very complex queries a user can ask in the domain. Furthermore, we evaluate and report the effects of information extraction and inference on the query performance. Semantic Web Services (SWSs) [18] represent the most recent and revolutionary technology developed for machine-to machine interaction.

As for the conventional web services, the problem of discovering and selecting the most suitable web service represents a challenge for SWSs to be widely used. In this method, we propose a mapping algorithm that facilitates the redefinition of the standard web services annotations (i.e., WSDL) using semantic annotations (i.e., OWL-S). This algorithm will be a part of a discovery mechanism that relies on the semantic annotations of the web services to perform its task. The proposed algorithm is implemented, and its components are validated using some test collections and real examples. An experimental test of the proposed techniques is reported, showing the impact of the proposed algorithm in decreasing the time and the effort of the mapping process.

In proposed work emphasis on some follows issues:

1. The efficiency of Matching Techniques:-

It is important in dynamic applications especially, when a user cannot wait too long for the system to respond or when memory is limited. Efficiency issue can be tackled through.

Concurrent matching of the task using cluster computing.

- Distribution of matching task with available computational resources.
- Modularization of ontologies, yielding smaller more targeted mat thing task.
- Optimization of existing and empirically proved to be a useful matching method.

Matching with Background knowledge :-

Matching can be performed by discovering a common context or background. Knowledge for the ontologies and use it to extract relations between ontology. The way background knowledge sources are identified to be useful, it there is enough entities in common for a particular matching task. The way background knowledge sources are selected i.e. given multiple sources such as domain specific ontologies and upper-level ontologies.

© 2017, IJCERT All Rights Reserved

Matcher Selection, Combination, and Tuning:-

For dynamic applications, it is necessary to perform matcher combination and self-tuning at run time. As the number of available marchers increases the problem of their selection will become more critical. There are many different constraints and requirements applied to the matching task e.g. correctness completeness, execution time, main memory thereby involving multi decision criteria. The main issue is the semantic web combination of matches by looking for complementarity, balancing the weakness and reinforcing the strength of the components.

Ranking using user profile or using query searching:-

The Web has significant amount of data and extracts correct information from that data is a big task if the system has the user background knowledge, it is easy to choose desired information form a large amount of data.Another method for ranking is query knowledge.

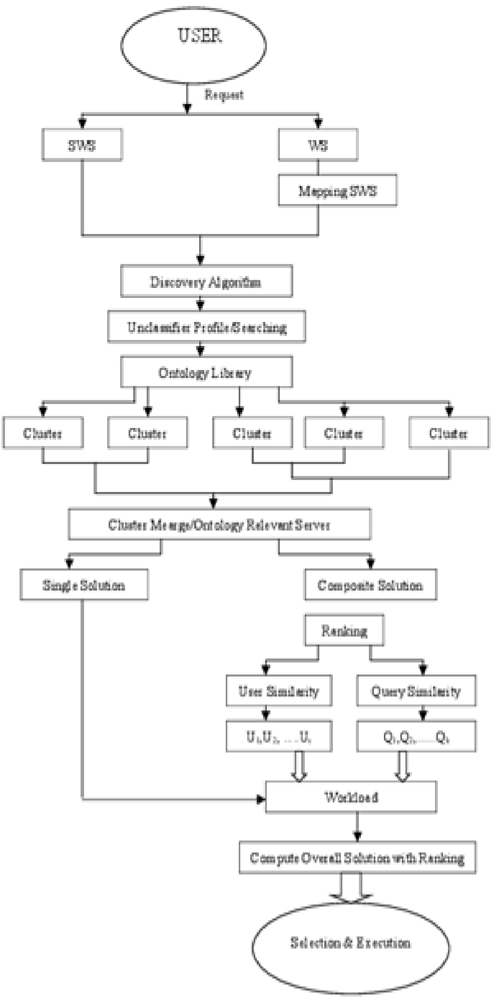

Figure 1. Flowchart of Ontology based searching technique

VI. Conclusion

The semantic web is capable of extracting millions of information from World Wide Web. Handling that information is a huge task for Semantic search engine developers. In this paper, we propose a new algorithm that must improve the efficiency of information retrieval. Using cluster-based technique and ranking mechanism we should increase the CPU processing speed.Also, this algorithm helps to eliminate the unwanted nodes that improve the search accuracy. Further work has been involved in improving the environment of the Micro semantic web with enhanced algorithms. Because of decentralized as well as the heterogeneous structure of web pages, it is impossible to use the same ontology for all domains hence brief study of semantic communication is needed.

Reference

[1] P Shvaiko, J Euzenat Ontology Matching: State of the Art and Future Challenges IEEE Transaction on Knowledge and Data Engineering, Vol. 25, No1. January 2013 Page(s): 158-176

[2] Farrag, Tamer Ahmed, Saleh, Ahmed Ibrahim; Ali, Hesham Arafat Toward SWSs Discovery: Mapping from WSDL to OWL-S Based on Ontology Search and Standardization Engine IEEE Transaction on Knowledge and Data Engineering, Vol. 25, No5. May 2013pages 1135-1147.

[3] A Telang, C Li, S Chakravarthy One Size Does Not Fit All: Toward User- and Query-Dependent Ranking for Web Databases IEEE transaction on knowledge and data engineering, vol 24, no9, September 2012 pages 1671 - 1685.

[4] P Kremen, Z Kouba Ontology-Driven Information System Design IEEE Transaction on System, Man and Cybernetics-PartC: Applications and Reviewes, Vol. 42, No. 3, May 2012 pages 334 – 344.

[5] V Milea, F Frasincar, U Kaymak town: A Temporal Web Ontology Language IEEE Transaction on System, Man, and Cybernetics-PartB: Cybernetics, Vol. 42, No. 1, February 2012 pages 268 - 281.

[6] P Panov, S Dzeroski, L Soldatova OntoDM: An Ontology of Data Mining 2008 IEEE International Conference on Data Mining Workshops Page(s): 752 - 760.

[7] Wong, W Liu, M Bennamoun Ontology Learning from Text: A Look-Back and into the Future ACM Computing Surveys, Vol. 44, No. 4, Article 20, Publication date: August 2012.

[8] Dimitrios A. Koutsomitropoulos, Ricardo Borillo Domenech, Georgia D. Solomou"A Structured Semantic Query Interface for Reasoning Based Search and Retrieval "The Semantic Web: Research and Applications Lecture Notes in Computer Sc. Volume 6643,2011,pp17-31.

[9] J Zhai, K Zhou - Information Science and Management "Semantic Retrieval for Sports

© 2017, IJCERT All Rights Reserved

Information Based on Ontology and SPARQL" Published in:Information Science and Management Engineering (ISME), 2010 International Conference of (Volume:1)Date of Conference:7-8 Aug. 2010 Page(s): 395 - 398 .

[10] Z Li, YL Zheng, SN Li, WW Liang "A Knowledge Sharing Convergence Platform Based on OWL-S and Semantic Relations" Published in Software Engineering (WCSE), 2010 Second World Congress on (Volume:1) 19-20 Dec. 2010 Page(s): 65 - 68 .

[11]G. Shiva Prasad , N.V. Subba Reddy, U,Dinesh Acharya "Knowledge Discovery from Web Usage Data: A Survey of Web Usage Pre processing Techniques" Information Processing and Management Communications in Computer and Information Science Volume 70, 2010, pp 505-507.

[12]J. Euzenat and P. Shvaiko, Ontology Matching. Springer, 2007.

[13]F. Giunchiglia, M. Yatskevich, and P. Shvaiko, "Semantic Matching: Algorithms and Implementation," J. on Data Semantics, vol. 9, pp. 1-38, 2007.

[14] J. Cardoso, Semantic Web Services: Theory, Tools, and Applications Idea Group, Inc., 2007.Web Services Description Language (WSDL), W3C Note, HTTP:// www.w3.org/TR/wsdl, 2001.

[15]Web Ontology Language for Services (OWL-S), W3C Member Submission, http://www.w3.org/Submission/OWL-S/, 2004.

[16]D. Martin, M. Burstein, D. Mcdermott, S.Mcilraith, M. Paolucci, K.Sycara, D.L. Mcguinness, E. Sirin, and N. Srinivasan, "Bringing Semantics to Web Services with OWL-S" World Wide Web, vol. 10, no. 3, pp. 243-277, Sept. 2007.

[17] T.A. Farrag and H.A. Ali, "A Cluster-Based Semantic Web Services Discovery and Classification," Proc. ACME Second Int'l Conf. Advanced Computer Theory and Eng., pp. 1825-1834, 2009.

[18]B. Di Martino, "Semantic Web Services Discovery Based on Structural Ontology Matching," Int'l J. Web and Grid Services, vol. 5, no. 1, pp. 46-65, 2009.

[19]OWL, http://www.w3.org/2004/OWL/, 2004.

[20]S Bhattacharjee, A Dwivedi, RR Prasad "Ontology based spatial clustering framework for implicit knowledge discovery" India Conference (INDICON), 2012 Annual IEEE 9 Dec. 2012 Page(s): 561 - 566

[21]S. Chaudhuri, G. Das, V. Hristidis, and G. Weikum, "Probabilistic Ranking of Database Query Results," Proc. 30th Int'l Conf. Very Large Data Bases (VLDB), pp. 888-899, 2004.

[22[C. Dwork, R. Kumar, M. Naor, and D. Sivakumar, "Rank Aggregation Methods for the Web," Proc. Int'l Conf. World Wide Web (WWW), pp. 613-622, 2001.

[23]Yiyao Lu, Hai He, Hongkun Zhao, Weiyi Meng, Member, IEEE, and Clement Yu, Senior Member, IEEE Annotating Search Results from Web Databases IEEE Transaction on Knowledge and Data Engineering , Vol 25, No 3, March 2013.

[24] G Singh, V Jain, M Singh "Ontology development using Hozo and Semantic analysis for information retrieval in Semantic Web " Image Information Processing (ICIIP), 2013 IEEE Second International Conference on 9-11 Dec. 2013 Page(s): 113 - 118 .

[25] Z Yun, S Huayou, Q Hengnian "A Semantic Web Services discovery mechanism design and implementation based on OWL Ontology " Educational and Network Technology (ICENT), 2010 International Conference on 25-27 June 2010 Page(s): 139 - 143.

[26] Myint Myint Thein , Soe Lai Phyue , Mie Mie Su Thw," Semantic Web Information Retrieval in XML by mapping to RDF schema " Published in Networking and Information Technology (ICNIT), 2010 International Conference on 11-12 June 2010 Page(s): 500 - 503.

[27]T Bhatia – IJCST " Link Analysis Algorithms For Web Mining" IJCT vol 2 issue 2 June 2011

[28]G Kumar, N Duhan, AK Sharma "Page ranking based on a number of visits of links of Web page " Computer and Communication Technology (ICCCT), 2011 2nd International Conference on 15-17 Sept. 2011 Page(s): 11 - 14 .

[29]J.R.G. Pulido a, M.A.G. Ruiz b, R. Herrera c, E. Cabello d, S. Legrand e, D. Elliman Ontology languages for the semantic web: A never completely updated review 0950-7051/$ - see front matter © 2006 Elsevier B.V.

[30]. B.ASWINI, B.RANJITH," Robust Model-Based Data Management", International Journal of Computer Engineering in Research Trends, Volume 1, December 2014, pp 453-460.

© 2017, IJCERT All Rights Reserved