ISSN (Online): 2456-0448
International Journal Of Innovative Research In Management, Engineering And Technology
Vol. 3, Issue 4, April 2018

Personalized Web Search using Ontology Matching and Ontology Mapping

[1] A.Sangeetha, [2] Dr.C.Nalini
[1] Phd. Scholor Department of Computer Science And Engineering, Bharath Institute of Higher Education & Technology,Chennai ,India
[2] Professor Department of Computer Science And Engineering Bharath Institute of Higher Education & Technology,Chennai ,India

Abstract: ***Web mining is the use of the information mining which is valuable to remove the learning. Most study on Web mining has been from information driven. Web utilization mining, structure mining and content mining are the sorts of Web mining. Web use mining is utilized as a part of mining the information from the web server log records. Web Personalization is one of the ranges of the Web use mining that can be characterized as conveyance of substance to a specific client or as personalization requires verifiably or unequivocally gathering data of the client. Utilizing that learning in your substance conveyance structure to control what data you present to your clients and how you introduce it. Personalized web search varies from bland web search, which returns indistinguishable research results to all clients for indistinguishable inquiries, paying little mind to shifted client interests and data needs. Web search engines have completed tremendous commitments to the web and the social order.***

I. INTRODUCTION

For all intents and purposes any application that includes various ontologies must set up semantic mappings among them, to guarantee interoperability. We decide comparability through tenets which have been encoded by philosophy specialists. These guidelines are then joined for one general outcome. We display an Ontology Matching calculation that incorporates a Word Sense Disambiguation procedure of OWL classes and gives semantic arrangements utilizing an augmentation of OWL constructors. The disambiguation procedure manufactures a system of words and connections, utilizing outside thesaurus as WordNet and Roget. We utilize the system like a portrayal of our setting for the determination of the significance. We utilize various methods in view of the utilization of particular guidelines and a blend of weighted terms, frequencies and kind of connections utilizing the system of information. In the assessment stage, we contrast our model and two benchmarks: one in regards to with the disambiguation of classes of ontologies and the other one utilizing Ontology Alignment Evaluation Initiative datasets. The quality and number of arrangements demonstrate a slight change. Most metaphysics coordinating strategies depend on insights. Conversely, this theory is dedicated to upgrading philosophy coordinating by rationale based thinking, both for mapping age and for mapping refinement. We adjust various coherent standards from existing work to recognize and evacuate mapping recommendations that are sensibly unintended. Diverse ways to deal with settling watched disjointedness, irregularity or infringement of standards are examined. Various heuristics are likewise considered to improve the mapping refinement process. Our trials yielded a promising gauge accuracy of 78% and recognized an arrangement of basic issues that should be considered to accomplish the maximum capacity of the worldview. Other than giving a decent execution as a remain solitary matcher, our worldview is corresponding to existing systems and along these lines could be utilized as a part of half breed instruments that would additionally propel the best in class in the philosophy coordinating field.

II. ONTOLOGY RANKING

Finding a significant relationship (ontology ranking) among this metadata is a fascinating and testing research point. Similarly as ranking of documents is a basic part of the present web crawlers, the ranking of complex connections will be a critical segment in tomorrow's Semantic Web examination motors. Just restricted work exists that tends to the issue of looking and ranking ontology in light of a given question term. In this work, the objective is to utilization of semantic connections for ranking records without depending on the presence of a particular structure in a report or connections between documents. Rather, genuine substances are recognized and the significance of records is resolved utilizing connections that are known to exist between the elements in a populated ontology. We present a measure of importance that depends on traversal and the semantics of connections those connection substances in ontology. We expect that the

IJIRMET ISSN (Online): 2456-0448
International Journal Of Innovative Research In Management, Engineering And Technology
Vol. 3, Issue 4, April 2018

semantic relationship-based ranking methodology will be either an option or a supplement to generally sent report scan for finding very pertinent records that customary syntactic and factual procedures can't discover. In this paper we present Graph based Semantic Similarity Ranking (GSSR), a two tire ranking calculation for semantic ranking. We expect that the semantic relationship-based ranking methodology will be either an option or a supplement to broadly conveyed archive look for finding very applicable records that customary syntactic and factual systems can't discover. We contrast GSSR and best in class ontology ranking models and conventional data recovery calculations. This assessment demonstrates that GSSR fundamentally beats the best ranking models on a benchmark ontology accumulation for most of the example queries characterized in the benchmark.

III. PROPOSED WORK

To give personalized search results to users, personalized web search maintains a user profile for every person. A user profile stores approximations of user tastes, interests and preferences. It is created and refreshed by misusing user-related information. Such information may include: 1. Statistic and land information, including age, sexual orientation, instruction, language, nation, address, interest areas, and other information; 2. Search history, including previous queries and clicked documents. User browsing conduct when seeing a page, such as staying time, mouse click, mouse development, scrolling, printing, and bookmarking, is another imperative component of user interest. 3. Other user documents, such as bookmarks, favorite web sites, visited pages, and emails. User information can be specified by the user (expressly gathering) or can be automatically learnt from a user‟s historical activities (verifiably gathering). Gathered user information is processed and sorted out as a user profile in a specific structure, contingent upon the need of personalization algorithm. This can be finished by creating vectors of URLs/domains, keywords, topic categories, tensors, or something like that [90].

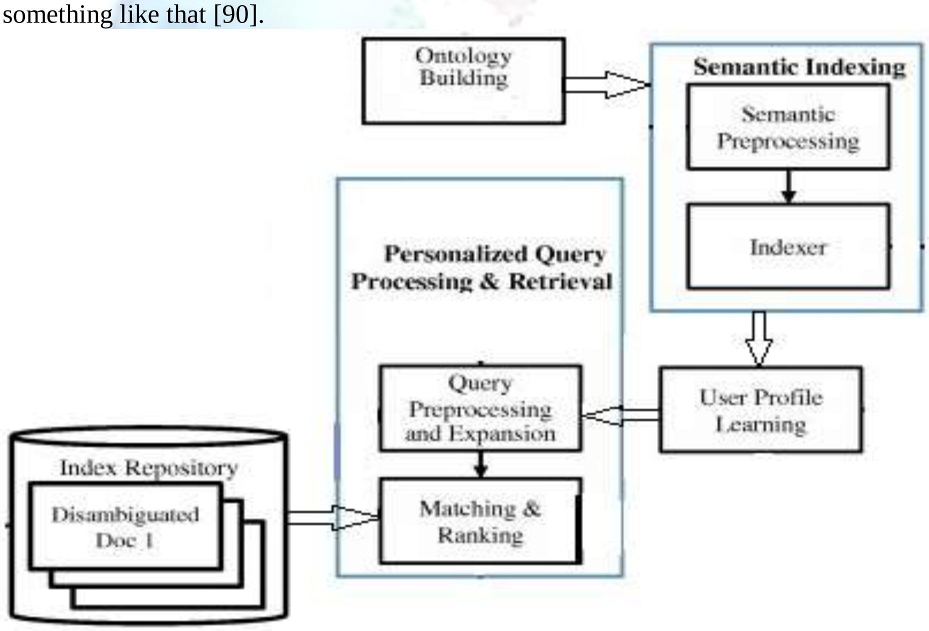

Figure 3.1 Architecture

The architecture of the proposed research work focuses to enhance the following components,
1. User Profile and Ontology Construction
2. Inquiry mapping
3. Substance and keyword extraction
4. Ranking
5. Search Results.

When a user enters his/her search word in our ontology search, a set of Ontologies are created. Ontology is the terms, concepts, relations related to the user question word. Each inquiry is processed and the results are produced by the search engine. These results are produced by the search engine based on ontology for each inquiry. The search engine also gets the user profile information for creating search results. The user question is mapped with user personal information for breaking down every user conduct, interests to identify the correct requirement of every individual user and results are delivered. Question mapping maps a user inquiry to a set of categories, which represent the user's search goal. This set of categories can serve as a setting to disambiguate the words in the user's question.

Based on the user inquiry, search engine results are retrieved. Each result is independently examined based on keywords and substance. User Query is pre-processed to identify the root words. Data preprocessing is a data mining system that involves transforming crude data into an understandable format. Real-world data is often deficient, inconsistent, as well as ailing in specific behaviors or trends, and is probably going to contain numerous errors. Data preprocessing is a demonstrated technique for resolving such issues. Data preprocessing prepares crude data for further processing. Each root words are considered for Dictionary construction and Dictionary is worked with synonyms for the user inquiry

Each result page keywords and substance words are pre-processed and compared against the lexicon. If a match is found then specific weight is granted to each word. Finally, the aggregate relevancy of the specific connection against user request is processed by summarizing every one of the weights of the keyword and substance words and term frequencies are ascertained to determine the aggregate relevancy .The term frequency is a numeric statistic that defines how essential a term to a document or to a gathering of corpus. The page which contains add up to relevancy esteem nearest to 1 are ranked as first page and 0 are ranked as last page. The times the terms in the question show up in the document can help determine how relevant the document is. The ranked results are automatically produced by the search engine to upgrade the efficiency of the search quality

IV. RESULTS AND EVALUATION

We basically take two ontologies and create mappings between the entities based on a given strategy. These mappings are validated against the correct mappings which had been created in beforehand. Our goal was to reach the best number of mappings, which is quantified in the f-measure`. As the absolute quality of mappings is highly dependent of the complexity of the ontologies themselves, we focus on the relative performance of different mapping strategies. All the tests were run on a standard notebook. To allow for comparability not only between our own test series, but also with existent literature we will focus on using standard information retrieval metrics. The definitions of precision and recall is adapted by us to fit the given evaluation scenario.

4.1 Data Sets

DBLP is a bibliographic database for computer field. The principle issue in DBLP is the assignment of papers to creator entities. This dataset1 provides bibliographical information about software engineering journals and proceedings. It includes 50,000 records. The second dataset 2 has been constructed from parts of DBLP that was automatically cleaned (using fine-tuned heuristics) or physically cleaned (because of creator requests), where different aliases for a person are known or ambiguous names have been resolved. The data set consists of paper reference pairs that can be assigned to the following categories: Two papers from the same creator, Two papers from the same creator with different name aliases (e. g., with/without center beginning), Two papers from different authors with different names, Two papers from different authors with the same name For each paper combine, the matching task is to choose whether the two papers were composed by the same creator. The data set contains 2,500 paper pairs for every class (10,000 altogether). This does not represent the first distribution of ambiguous or alias names in DBLP (where around 99.2 % of the creator names are non-ambiguous), yet makes the matching task more difficult and interesting.

4.2 Results

We will focus on the averaged table for the discussion, which already covers the complete results. In figure 4 we present the results of the first data set with ISOM, CURE and SAT solver with each new mapping we recalculate the other measures. The graphs show the respective precision, recall, and f-measure values.

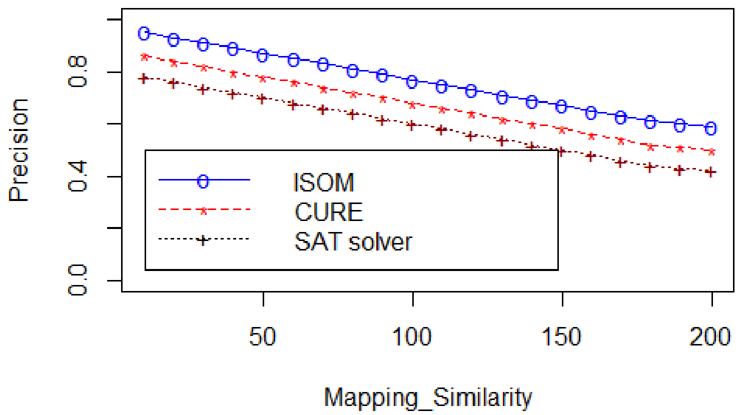

Figure 4.3 Precision with Mapping Similarity

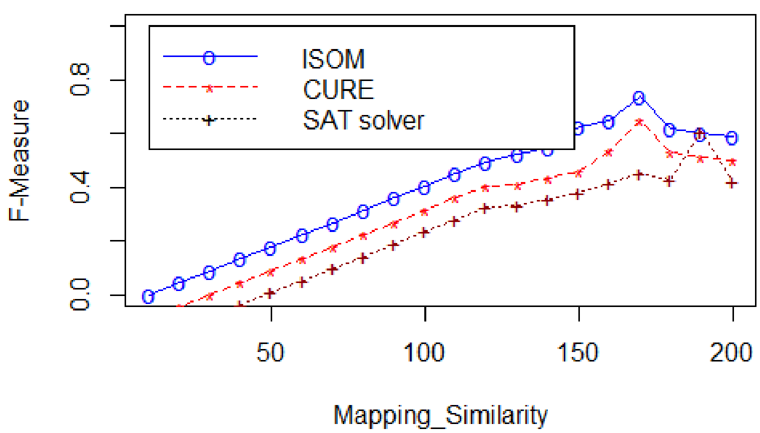

Figure 4.4 F-Measure with Mapping Similarity

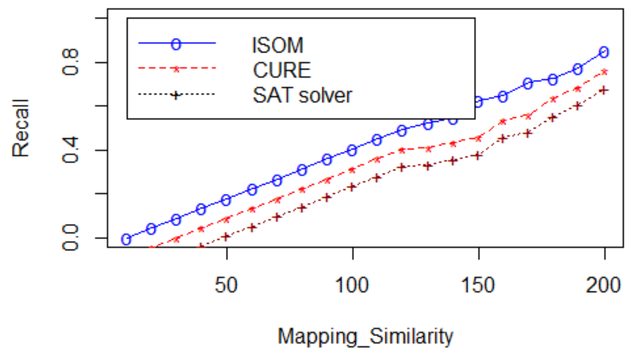

Figure 4.5 Recall with Mapping Similarity

We will now compare the two strategies with respect to the defined evaluation measures. The highest mappings are all correct, what one can see from the precision value of 1 for all the strategies. But what one can also see is that ISOM keeps the precision value high for many more mappings than CURE [64] and SAT solver [72]. Recall only reaches a medium level; the final level is much higher for ISOM. A consequence of these two measures is that the f-measure is also higher for the advanced approach in comparison to the naive approach. They often missed the highest f-measure value considerably. A general comparison graph is plotted in figure 4.3, 4.4. and 4.5 respectively. This comparison graph shows the average results over all four data sets, each with the different strategies. Precision, recall, and f-measure reach their highest values ISOM. In general one can say that there is an increase with the rise of strategy complexity.

V. CONCLUSIONS

We described the use of ISOM, and in particular, of multi-strategy learning, for computing concept similarities. This learning technique makes our approach easily extensible to additional learners, and hence to exploiting additional kinds of knowledge about instances.

1. A.Sangeetha, C.Nalini , Department of Computer Science and Engineering, Bharath University, Tamil Nadu, India "Web Personalization Using Web Usage Mining" International Journal of Innovative Research in Computer and Communication Engineering Vol. 3, Issue 6, June 2015

2. A.Sangeetha, C.Nalini , Department of Computer Science and Engineering, Bharath University, Tamil Nadu, India "Personalized Web Search Based On User Profile And Ontology" International Journal Of Applied Engineering Research, Issn 0973-4562 Vol. 10 No.60 (2015) © Research India Publications; Httpwww.Ripublication. Comijaer.Htm Vol. 10 No.60 (Sep 2015) AnnexureII

3. J. Lai and B. Soh, "Personalized web search results with profile comparisons," in Information Technology and Applications, 2005. ICITA 2005. Third International Conference on, vol. 1. IEEE, 2005, pp. 573–576.
4. Z. Dou, R. Song, J.-R. Wen, and X. Yuan, "Evaluating the effectiveness of personalized web search," IEEE Transactions on Knowledge and Data Engineering, vol. 21, no. 8, pp. 1178–1190, 2009.
5. C. Chen, K. Lin, H. Li, and S. Dong, "Personalized search based on learning user click history," in Cognitive Informatics (ICCI), 2010 9th IEEE International Conference on. IEEE, 2010, pp. 490–495.
6. R. Hu, W. Dou, X. F. Liu, and J. Liu, "Personalized searching for web service using user interests," in Dependable, Autonomic and Secure Computing (DASC),2011 IEEE Ninth International Conference on. IEEE, 2011, pp. 156–163.
7. R. M. Kaingade and H. A. Tirmare, "Personalization of web search based on privacy protected and auto-constructed user profile," in Advances in Computing, Communications and Informatics (ICACCI), 2015 International Conference on. IEEE, 2015, pp. 818–823.

8. [6] H.-j. Kim, S. Lee, B. Lee, and S. Kang, “Building concept network-based user profile for personalized web search,” in Computer and Information Science (ICIS), 2010 IEEE/ACIS 9th International Conference on. IEEE, 2010, pp. 567–572.
9. K. Makvana, P. Shah, and P. Shah, “A novel approach to personalize web search through user profiling and query reformulation,” in Data Mining and Intelligent Computing (ICDMIC), 2014 International Conference on. IEEE, 2014, pp. 1–10.
10. Xavier Amatriain, Mining large streams of user data for personalized recommendations, ACM SIGKDD Explorations Newsletter, v.14 n.2, December 2012 [doi>10.1145/2481244.2481250]

11. Z. Dou, R. Song, J.-R. Wen, and X. Yuan, “Evaluating the effectiveness of personalized web search,” IEEE Transactions on Knowledge and Data Engineering, vol. 21, no. 8, pp. 1178–1190, 2009.
12. A.Sangeetha, C.Nalini , Department of Computer Science and Engineering, Bharath University, Tamil Nadu, India “Enhancing web mining approach with ontology” Journal of Chemical and Pharmaceutical Sciences April - June 2016 E -301 JCPS Volume 9 Issue 2 ISSN: 0974-2115

13. H.-j. Kim, S. Lee, B. Lee, and S. Kang, “Building concept network-based user profile for personalized web search,” in Computer and Information Science (ICIS), 2010 IEEE/ACIS 9th International Conference on. IEEE, 2010, pp. 567–572.
14. .A.Sangeetha, C.Nalini , Department of Computer Science and Engineering, Bharath University, Tamil Nadu, India “An Effective approach for increasing the efficiency of Web searching with Feedback sessions” International Journal of Innovative research in Science, Engineering and Technology –IJIRSET Vol 3 Special Issue Mar 2014 ISSN (Print)2347-6710 ISSN(Online)2319-8753pp 1837-1842

15. A.Sangeetha, C.Nalini , Department of Computer Science and Engineering, Bharath University, Tamil Nadu, India “Web Search By Inferring User Interest Based On Personalized Search And Implicit Indicators “ International Journal Of Engineering Associates(Issn: 2320-0804) # 9 / Volume 4 Issue May 2015