An Ontology Based Context Aware Protocol in Healthcare Services

Anirban Chakrabarty and Sudipta Roy

Abstract Interpretation of medical document requires descriptors to define semantically meaningful relations but due to the ever changing demands in healthcare environment such information sources can be highly dynamic. In these situations the most challenging problem is frequent ontology search keeping with user's interest. To manage this problem efficiently the paper suggests an ontology model using context aware properties of the system to facilitate the search process and allow dynamic ontology modification. The proposed method has been evaluated on Cancer datasets collected from publicly accessible sites and the results confirm its superiority over well known semantic similarity measures.

Keywords Ontology mapping · Context awareness · Search personalization · Healthcare

1 Introduction

Most Healthcare systems contain a large collection of related documents which necessitates a semantic search system to swiftly and accurately identify documents which satisfy user's needs. It has been found that keyword based search engines most often fail to give the expected results relevant to the query context. One of the key factors for personalized information access is the user context [1]. Context can be defined as the user's objective for seeking information [2]. In this paper, context is defined through the notion of ontological profiles which are updated over time to reflect changes in user interests. Ontology consists of a formal conceptualization in a particular field of interest that can be easily visualized or can be considered as a

Anirban Chakrabarty (✉) · Sudipta Roy
Assam Central University, Silchar, India
e-mail: chakrabarty.anirban@redffmail.com

Sudipta Roy
e-mail: Sudipta.it@gmail.com

J.K. Mandal et al. (eds.), *Proceedings of the First International Conference on Intelligent Computing and Communication*, Advances in Intelligent Systems and Computing 458, DOI 10.1007/978-981-10-2035-3_15

description of the elements it contains and when aligned with other ontologies it can help in forming a community of such elements or concepts [3, 4].

The major contributions suggested in this paper is mentioned as follows: (1) It provides an ontology for cancer patients containing both medical and socio economic conditions (2) to search for an optimal treatment plan suited to the patient (3) The ontology can be extended by healthcare experts based on changing treatment options, different disease symptoms and missing information with minimum involvement from external sources. (4) An ontology mapping strategy has been suggested that exploits contextual information of source ontology and supplements it in linguistic and structural similarity to enhance the strength of match with the target ontology.

2 Related Work

Most retrieval systems suffer from keyword barrier phenomenon which refers to the inability of information retrieval systems to convey semantic context of documents [5]. Related work developed a framework for Formal Concept Analysis which extended the Tf-Idf weighting model by introducing ontology dependent concepts [6]. Most search engines do not include user preferences and search context but provide users with a generalized search facility [7]. A related study shows that little work has been done on contextual retrieval to combine search technologies and ontology alignment using context in a single framework to provide the most accurate response for a user's information requirement [8].

A popular method to facilitate information access is through the use of ontology. Researchers have attempted to utilize ontology for improving navigation effectiveness as well as personalized Web search and browsing for generating user profiles [9]. An innovative approach for Ontology based on Concept merging was suggested were first the horizontal technique checks all relationships between concepts at the same level of two ontologies and merges them as defined by WordNet, then the vertical approach completes the merging of concepts at different levels but placed on same branch of the tree [10].

The notion of concept similarity calculation using ontology is that two concepts have a semantic correlation, and there exists a path in class hierarchy diagram. In a pioneering work, Resnik measures the semantic similarity of two words according to the maximum amount of information of their common ancestor node [11]. Leacock and Chodorow proposed a semantic similarity model based on distance which is an easy and natural approach but is heavily dependent on the ontology hierarchy already established [12]. Study of related work shows the use of RIMOM: a dynamic multi strategy ontology alignment framework which uses both linguistic and structural similarity to map source and target ontologies but it is suitable for only for 1:1 mapping [13]. The ontologies can be expressed in various ways. The most common language used to model ontologies is the Web Ontology Language (OWL), which is a form of RDF and is written using a subset of XML [14].

Unfortunately, most of these systems can provide little support in support for knowledge sharing and context reasoning because of their lack of ontology. Recent study focus on the use ontology based personalization system for assisting healthcare industry in patient diagnosis and intervention plans [15].

3 Methodology

In this work ontology has been developed to organize the terms of cancer disease, its symptoms, diagnosis, treatment options, intervention plans, patient condition and other social issues related to disease. On top of this a search algorithm based on context information stored in Ontology has been suggested which extracts all correlated words of the search string and can map from one ontology to another (Figs. 1 and 2).

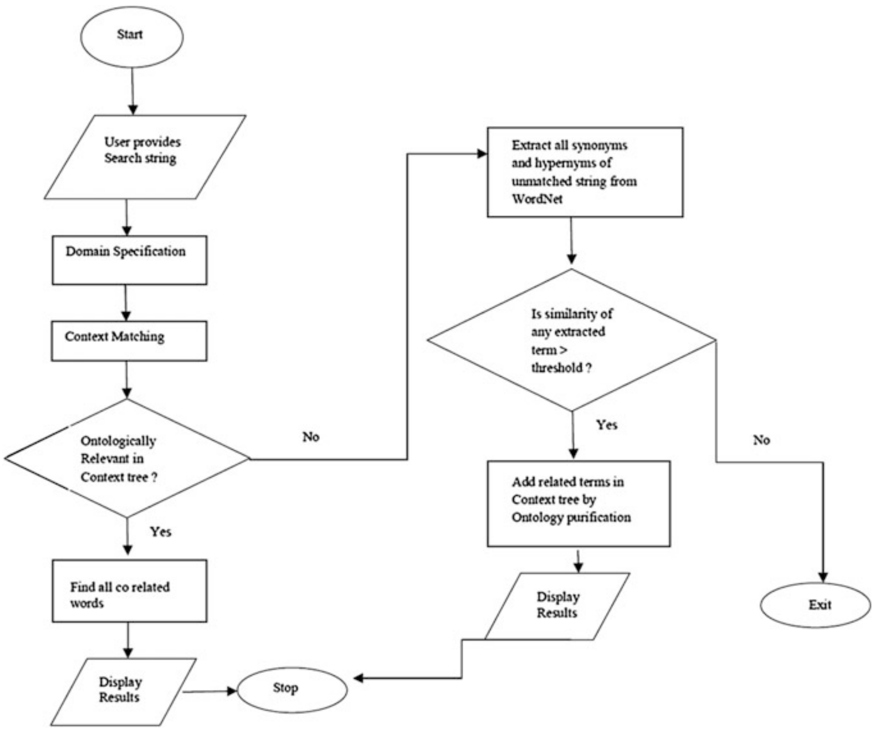

Fig. 1 Flowchart of context matching using an ontology framework

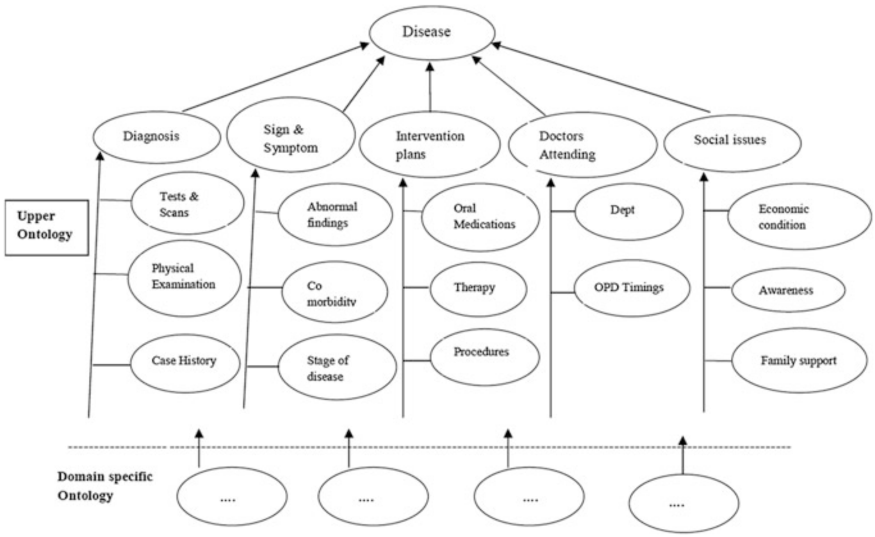

Fig. 2 Ontology hierarchy for cancer patients

4 Ontology Used

There are many existing ontology, but since critically ill patients need accurate and urgent medical intervention we have developed Cancer related Ontology containing medical and social concepts to assist in searching information and decision making.

5 Explanation of Proposed Algorithms

5.1 Context Search Technique

Step 1 Perform linguistic match between search string and the one present in target ontology based on Word Net. If direct match is found then display output and exit.
If no direct match found then based on the similarity results the pairs are sorted into three buckets: (a) above the upper threshold—provisionally similar (b) between the upper and lower threshold—uncertain bucket and (c) below the lower threshold—no match. Add all composite matches to the "uncertain bucket" along with the concepts between the upper and lower threshold (selected based on experimentation).

Step 2 Structural matching: For concepts whose similarity values are in uncertain bucket, perform the following:

(a) for each pair of concept, compare their parent nodes (when matching is between two ontology) to determine the total parental similarity for updating the original pairs similarity value. If match is found increment the score by 0.1.
(b) for each pair of concepts, compare their grandparent nodes to determine the total similarity for updating the original pairs similarity value. If match is found increment the score by 0.05.

Step 3 Use Word Sense disambiguation on the set of terms in ‘Uncertain’ and ‘Provisionally similar’ buckets to select the words actually representing the given context. The WSD method used here takes average of similarity values for each concept searched to show its relatedness to the ontology.

5.2 Ontology Modification

This procedure is called when the context is not found in the ontology tree but the searched context has a comparable relation then the context is dynamically added as a node in the ontology tree in appropriate location. Each ontology node is associated with an Id for comparison.

Pseudo code:

```
Modify (Root,Id,mi,N,M,Left,Right,Par)
Set Ptr = Root, Id= TRUE.
 Apply Algorithm-2 to find the similarity or relevance of searched  area.
 [If Node found] Set Ptr=Node.
Else if [node was not found but has to be added based on relevance threshold of the
domain set then check immediate ancestor. ]
{ Id[NodeO1]->Par=Id [NodeO2]->Par

Add new node.
If{ Id [NodeO1]<Id [NodeO1]->Par
Set Left[Par]=New
Else Set Right[Par]=New.
Set Id of new node.}
Print Updation Successful.
}}
[Else Updation Not Possible. ]
}
```

6 Experiment

6.1 Data Set

The experimental data used in this work has been collected from different cancer related websites which allows open access and made publicly available for research and study purposes [16–21]. The training data set comprising of 2721 documents was used for the representation of the cancer ontology indexing 204 concepts in the hierarchy.

6.2 Experimental Metrics

To perform a comparison of the improvement of proposed method with other established semantic search algorithms used in contemporary literature like k-NN, Resnik similarity, Rocchio based methods [22, 23].

The similarity measure suggested by Resnik, gives the information content (IC) of the Least Common Subsequence (LCS) for two concepts:

$$\mathrm{SimRes} = \mathrm{IC(LCS)} \tag{6}$$

where IC is defined as:

$$\mathrm{IC(c)} = -\log \mathrm{p(c)} \tag{7}$$

here p(c) is the probability of encountering a context c in a corpus.

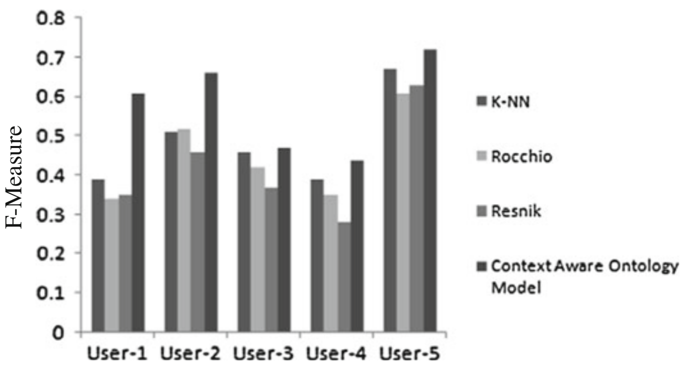

Fig. 3 Comparison of F-measure with other similarity measures for five users

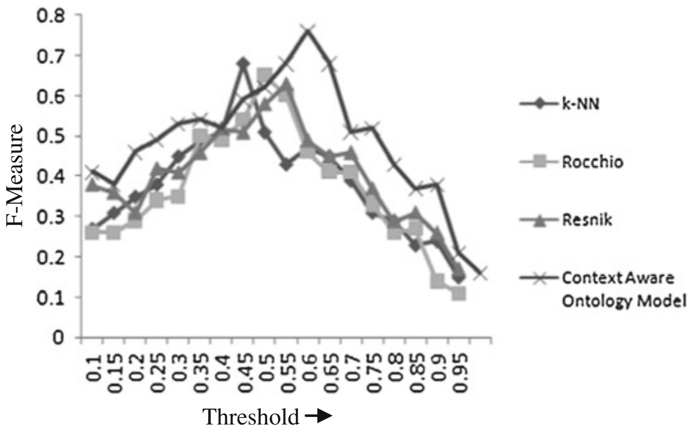

Fig. 4 Comparison of F-measure with varying threshold

The effectiveness of search was measured in terms of Top-n Recall and Top-n Precision [24]. The F-measure defined as F = 2 * P * R/(P + R) is a balanced mean between precision and recall metrics and was used for comparing searches made by five different users.

6.3 Results

See Figs. 3, 4, 5 and 6.

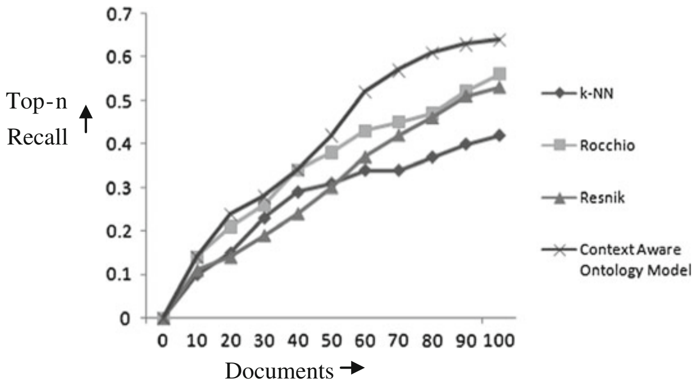

Fig. 5 Variation of average top-n recall in top-n documents using overlap queries

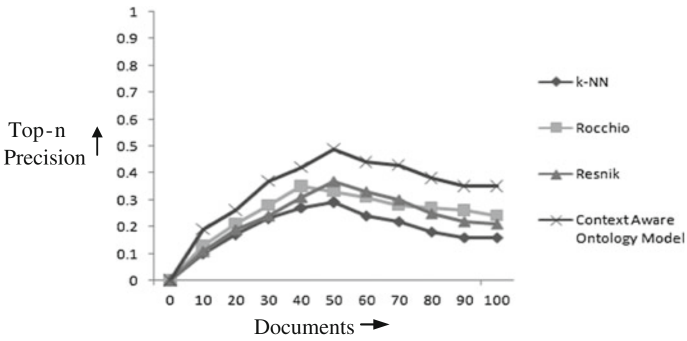

Fig. 6 Variation of average top-n precision in top-n documents using overlap queries

6.4 Discussion of Experimental Results

The comparison of F-measure with other semantic search techniques like k-NN, Rocchio and Resnik methods is depicted in Fig. 3. For all such users there has been improvement in F-measure values in case of our Context Aware Ontology model. Since the value of similarity depends to some extent on the threshold considered in the function so a variation of threshold with F-measure for different similarity techniques have been studied for same set of five users. The highest F-measure value was reached for a threshold of 0.6 in our context aware ontology model as shown in Fig. 4. Figures 5 and 6 gives the comparison of precision and recall values for top-n search results for all similarity techniques as shown above.

7 Conclusion

The work presents a framework for contextual information access using ontologies and demonstrated that the semantic knowledge embedded in cancer ontology can efficiently assist in search process and facilitate dynamic ontology modification. A comparison with other semantic search techniques shows that if contextual information present in ontology is retrieved effectively it can improve the search results based on user's requirements. This search technique can be extended to compare two separate ontologies as well.

It may be noted that the experimental results reported here is based on usage of randomly selected users, a few hundred queries, and a limited number of relevant documents. Future research in this area will consists of much larger scale of experiments and optimization parameters.

References

1. Pei-Min Chen, Fong-Chou Kuo , An information retrieval system based on a user profile, *The Journal of Systems and Software*, vol. 54, pp. 3–8, 2000.
2. Dey, A.K., Salber, D. Abowd, G.D., "A Conceptual Framework and a Toolkit for Supporting the Rapid Prototyping of Context-Aware Applications", *Human-Computer Interaction Journal*, Vol. 16(2–4), pp. 97–166, 2001.
3. Cimiano, P., Ontology Learning and Population from Text: Algorithms, Evaluation and Applications. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.
4. Gargouri, Y., Lefebvre, B., Meunier, J. Ontology Maintenance using Textual Analysis. Proceedings of the Seventh World Multi-Conference on Systemics, Cybernetics and Informatics (SCI). Orlando, USA (2003).
5. M-Y. Chen, H-C. Chu, Y-M. Chen, Developing a semantic enable information retrieval mechanism. *Expert Systems with Applications*, 37:322–340, 2010.
6. Rohana K. Rajapakse, Michael. Denham, Text retrieval with more realistic concept matching and reinforcement learning. *Information Processing and Management.* 42(5), 1260–127, September 2006.
7. Ashraf, J., Khadeer Hussain, O., Khadeer Hussain, F.: A Framework for Measuring Ontology Usage on the Web. In: The Computer Journal. Oxford University Press, 2012.
8. J. Allan, et al. Challenges in information retrieval and language modeling. ACM SIGIR Forum, 37(1):31–47, 2003.
9. Fang Liu, Clement Yu, Personalized Web Search for Improving Retrieval Effectiveness, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 16, No. 1, January 2004.
10. Miyoung Cho, Hanil Kim, and Pankoo Kim. A new method for ontology merging based on concept using wordnet. In Advanced Communication Technology, 2006. ICACT 2006. The 8th International Conference, volume 3, pages 1573/1576, February 2006.
11. Philip Resnik, "Using Information Content to Evaluate Semantic Similarity in a taxonomy", In Proceedings of the *14th International Joint Conference on Artificial Intelligence*. Monte real, Canada, pp. 448–453, 1995.
12. C. Leacock, M. Chodorow. "Combining Local Context and WordNet Similarity for Word Sense Identification", *Computational Linguistics*, vol. 24, no. 1, pp. 147–165, 1998.
13. Juanzi Li, Jie Tang, Yi Li, and Qiong LuoRiMOM: A Dynamic Multistrategy Ontology Alignment Framework, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 21, No. 8, August 2009.
14. Ashraf, J., Khadeer Hussain, O., Khadeer Hussain, F.: A Framework for Measuring Ontology Usage on the Web. In: The Computer Journal. Oxford University Press, 2012.
15. David Riano; Francis Real; Joan Albert Lopez; Fabio Campana; Sara Ercolani; Patrizia Mecocci; Roberta Annicchiarico & Carlo Caltagirone. An Ontology based personalization of Healthcare knowledge to support clinical decisions for chronically ill patients, Journal of Biomedical Informatics, Elsevier, 429–446, 45 (2012).
16. http://www.cancer.org/research/index [Last accessed on 12th October, 2015].
17. http://www.cancercare.org/accessengagementreport [Last accessed on 12th October 2015].
18. http://www.oncolink.org/resources [Last accessed on 7th October, 2015].
19. http://med.stanford.edu/cancer.html [Last accessed on 25th September, 2015].
20. http://www.ncbi.nlm.nih.gov/pubmed [Last accessed on 21st September, 2015].
21. http://www.cdc.gov/cancer/ [Last accessed on 16th September, 2015].
22. Oh-Woog Kwon, Jong Hyeok Lee. Text categorization based on *k*-nearest neighbor approach for Web site classification, *Information Processing and Management* Volume 39, Issue 1, Pages 25–44, January 2003.

23. Guanyu Gao, Shengxiao Guan. Text categorization based on improved Rocchio algorithm. In Proc. of Systems and Informatics (ICSAI) IEEE International Conference, pp. 2247–2250, 2012.
24. Paulo Cremonesi, Yehuda Koren, Roberto Turrin. Performance of recommender algorithms on top-n recommendation tasks. In Proc. of the fourth ACM conference on Recommender Systems, pp. 39–46, ACM New York, USA, 2010.