Comparative Study of Different Ranking Algorithms

Neera Chaudhary[1], Suresh Kumar[2]
[1]*Ambedkar Institute of Advanced Communication Technologies & Research, Govt. of NCT, Delhi*
[2]*Ambedkar Institute of Advanced Communication Technologies & Research, Govt. of NCT, Delhi*
E-mail: neerachaudhar18@gmail.com, sureshpoonia@aiactr.ac.in

***Abstract*—**The backbone of Semantic Web (SW) depends upon the knowledge representation where knowledge is represented by ontologies. SW requires proper mapping, merging and reusability inside the ontology and outside the ontology to form proper knowledge representation scheme. The proper merging and mapping of ontology enables the increase demand of ontology and produces an increasing number of operational ontology on the web. Ontology provides us a strong and common understanding among large information and specify a specific domain. This enables the easier access to environment. Moreover, ontologies are required to be ranked for search in a better and efficient way. Ranking methods of ontologies are available in literature which increases the probability of useful knowledge extraction by using ontology based searches. In this paper, we analyzes the popular ranking algorithm which are based on ontology so that it may be helpful to researchers in proceeding the further research.

Keywords*—*Ranking; Ontology; Semantic web

I. INTRODUCTION

With continuous growth of information over web, users have many options related to their queries inside the search box. But this increases many challenges in terms of time, skill and domain for the users to find desire and get relevant result. The key objective of Semantic Web is to improve the search mechanism using ontologies. Ontologies are playing an important role in representing domain knowledge [1]. Semantic Web (SW) offers several mechanisms that can be used to define and categorize the information and its context retrieving on Web. According to Tom Gruber a researcher in AI, an ontology is defined as "the explicit specification of conceptualizations, used to help programs and humans share knowledge" [2]. This means that Ontology defines the concepts and relationships used in describing and representing an area of concern. One mechanism that is the backbone of Semantic Web and helps to develop a knowledge base to represent the knowledge over web is ontology engineering. Ontology in semantic web helps to achieve the reusability and interoperability of data and minimize the challenges. It focuses to the study or concern about what kinds of things exist and what entities are there in the whole world.

Ontologies are sometimes also called vocabularies and help to classify the terms that can be further used in a particular application, classify and making possible relationships and defining probable constraints by using those terms. This helps to organize and making rich knowledge by integrating the data. To describe and define different forms of ontologies, Web consortium has recommended three formats to construct an ontology i.e. Resource Description Framework (RDF), Resource Description Framework Schema (RDFS), web ontology (OWL). Ontology helps to understand and integrate the knowledge in better way while we fetch the relevant information from search engines and interchange the data on web. A resource may be anything such as a Web page, a Web service, a thing, a concept, a property, etc.

Semantic Web based search engine uses concept/theme/domain based ranking schemes to provide effective and relevant information to users [3]. Ranking algorithms extract the most relevant and desired information based on the query given by user in semantic search engine environment. The basic concept of ranking is to make relationship between a set of items. If two documents got same rank, then it is considered a tie and gives it to equal probability. In the literature, we have found that ranking algorithm based on ontology provide useful results and also unique criteria such as internal structure, content similarities, Interlinks, semantic analysis, keyword and entity etc. In this paper, we have done detailed survey on various ontologies based ranking approaches based on three criteria's-functional process, advantages and disadvantages. This paper focuses on the analysis of most important ontological rank algorithms. By studying various ontology based ranking algorithm, a comparative study is done by focusing the criteria such relative strengths, benefits and limitations. This turns to a valuable research path in ontological based ranking in semantic web. This paper is classified into different section. Section 2 describes about the ontology Ranking, section 3 is related to ontological based ranking algorithms, section 4 is about the comparative analysis of ontology based ranking algorithm and section 5 is related to the conclusion. Lastly, the paper is summed up with references.

II. ONTOLOGY RANKING

The main purpose of ontology is to represent the domain knowledge and make useful for further share and reusable which is currently used by several search engines (Swoogle search engine, Ontokhoj etc.). Reusing ontology has the inherit capability which makes it cost effective, accurate and high quality ontologies. It also decreases the recursive nature of development costs by rebuilding existing ontologies. This helps to users to find user's needs by ranking mechanism which is based on semantic similarity matching. The ranking mechanism applies on different concepts (terms) making relational mapping (links) or matching taxonomy. Based on the given weight, ranking algorithm measures the rank of each selection in the given context. By doing ranking we can retrieve suitable

ontologies of a particular domain [3]. Various Semantic Search Engines like Swoogle, Ontokhoj and Falconetc are using ontology based ranking approach are performing better results with respect to precision and recall.

III. ONTOLOGICAL BASED RANKING ALGORITHMS

The ranking mechanisms that are applied on different aspects of ontology such as its structure, interlinks, contents, relations etc. are called ontological based ranking algorithm. There are various ranking algorithm available in literature which are effectively and efficiently used to the management of ontology.

- AKTiveRank Scheme [3].
- Content-based ontology rank Scheme [3]
- Onto Rank Scheme [4]
- Ontology Structure Rank Scheme [4]
- SIF Rank Algorithm [3]
- SHOE Rank Approach [3]
- Sem Rank Approach [5]
- Page Rank Approach [6]
- Relational Based Page Rank Approach [5]
- Rare Rank Approach [5]

A. AKTiveRank Algorithm

AKTiveRank is a concept based algorithm which uses different analytical measures to retrieve the result in response to user query. In this ranking scheme an ontology is evaluated based on the depth of coverage and identify the concepts based on the query submitted by user. It is based on four parameters viz Class Match Measure, Density Measure, Semantic Similarity Measure and Between-ness Measure [1][4] which computes all the measures of ontology and the evaluated values are merged and produced the total rank for the ontology [3].

B. Content Based Ontology Rank Algorithm

Content of an ontology is a vital component. The Content Based Ontology Ranking Algorithm pays attention on content similarities of ontology. In this scheme each ontology is ranked by matching the new items from class labels [14]. The items are specified by knowledge engineer and these are matched within a list of ontologies. In this process the item is matched from the list of ontologies and the output of matching is then ranked. Here List of ontologies are contained in corpus such as word net. It focuses on various content labels in different ontologies of the domain, matches and extracts a set of terms [1] [3] [4].

C. Onto Rank Algorithm

Onto rank algorithm deals with the concepts of Google API. The Google API acquires the initial 100 ontologies [15] to allow users to specify different types of criteria. Onto Rank search is used by two search strategies i.e. searching for structure and searching for classes. It focuses on "if and only if" association which exists among the classes in a relation set and detects directional or transitive reference relations. Based on the reference relation, ontology is evaluated [1] [3] [4].

D. Ontology Structure Rank Algorithm

The main focus of this algorithm is the semantic relation and the structure of Ontology where ranking criteria depends on the three ranking scores:

- Name of class.
- Semantic based relation.
- Structure of ontology.

The above measures helps to users for fetching the results (ontology) in response to the users query and then ranking is performed. The ranking measures depends on the user's needs and importance of application [1] [3] [4].

E. SIF Rank Algorithm

SIF stands for Semantic-aware importance flooding focuses on Ontology Web Language ontology, their retrievals and transforms them into directed graph. Each graph uses a repetitive but fix point computation to calculate the importance of nodes. Different patterns used in the algorithm are treated semantically. The nodes which are not semantically linked are neglected in computation and it assumes that it is reached as maximum number of iterations [1] [3] [4].

F. SHOE Rank Approach

The SHOE Rank Approach is a domain specific ontological approach. It is based on the semantic markup language which improves efficiency of results by describing the context of resources in web pages properly. It uses the context and annotation so that user can build query in a context and add markup to web pages using annotation.

It relates to the document types (web pages) to ontological concepts [5]. This means that different context on web pages are related themselves to ontological concepts. It uses different properties which are recognized only to semantic based search engine not by browsers. Thus, user of semantic based search, select the concept and mention the features from ontology and then the system returns results.

G. SEM Rank Approach

This approach ranks the search results by using various factors such as importance of results, results the user expects over complex text etc. In SEM rank, prediction of results is one of the important factor the user expects. Based on the prediction, ranked in results is formulated. The deepness in information affects the results of user query [5]. The SEM Rank method allocates the top rank value to unpredictable result and deepest or lower most rank to conventional results.

H. Page Rank Approach

Two scholars Brin and Page [13] at Stanford University developed this algorithm. The main concept in the algorithm is link structure which uses random surfer model to govern the

significance of web page where a user randomly preserves clicking of links on a page and if users are not interested in a particular page then it randomly switches to another page. Thus, a user under this model shows no bias towards any page or link. Therefore, each in-link to a page has its significance. The recursive nature of this algorithm maps each pages by linking. In addition to it, algorithm uses hyper link model which works on a particular document or website based on the number of times, the link of the document is accessed by the user. Higher the access rate of the link, Higher the ranking [6].

I. Relation Based Page Rank Approach

This ranking approach focuses on keywords and concepts given by the user to make search. The relationship is made between these two factor- keywords and concepts which is intended by user that acts as an extension of traditional search. Keywords and concepts exist in the page. If the page involves adequate keyword-concept, the associations are made to the expected user search and the likelihood increment of frequent page depends on the number of linked between concepts within the query linked with other concept is larger [5].

J. Rare Rank Approach

This approach provides a more effective search retrieval using link analysis model. For this, it uses large caches of data (WWW) and link structure. This approach uses the large cache as a dataset where algorithm finds the relation between the relevance and the quality score. The fundamental principle is that entity such as citations, publishers, authors, journals in combination with the topic in a terminological ontology will be able to simulate an environment suitable for the researches to conduct research or for any person to be able to carry out search operations [5].

IV. COMPARATIVE ANALYSIS OF RANKING ALGORITHM

We have made comparative analysis of methods of ranking based on their properties, their pros and cons which is mentioned in Table 4.1 below.

V. CONCLUSION AND FUTURE SCOPE

By comparing different ontology based algorithm, it is found that the semantic web- a future of current web describes the information on web over structured data and provides a knowledge representation framework for further use. This capability makes it more powerful than simple web. Since semantic web uses ontology, plays an important role to frame the semantic web. Ranking algorithm over ontology keeps the appropriate ontology in response to user's query keeps the rank on top list. This allows search engines to encounter the user's need before spreading over long list of retrieved items without wasting user's time.

In addition to it, we have observed that time complexity and ontologies are poorly inter- referenced are the major issues in ranking schemes. This gives us motivation to overcome such kind of issues by developing better algorithm and designing an ontology on different areas such as-education, Govt. organizations, commercial domain, etc.

REFERENCES

[1] R.Subhashini, Dr. J. Akilandeswari, A novel approach for ranking ontologies based on the structure and semantics, Journal of Theoretical and Applied Information Technology, 10th july2014.vol.65No.1.

[2] Thomas R. Gruber .1993. Towards a principles for the Design of Ontologies used for Knowledge sharing.

[3] Poonam Chahal, Manjeet singh, Suresh Kumar, ranking of web documents using semantic similarity, IEEE 2013 international conference on information system and computer networks.

[4] S.Sudha, Lakshmi.S, Comparison of Ontology Ranking Algorithms on Semantic Web, International Journal of Advanced Research in Computer and Communication Engineering Vol.5, April 2016.

[5] R.Subhashini, Dr.J.Akilandeswari, V.Sinthuja, A Review on Ontology Ranking Algorithms, International Journal of Computer Application (0975-8887) Volume 33-No.4, November 2011.

[6] Darshan Bhansali, Harsh Desai, Khushali Deulkar, A Study of Different Ranking Approaches for Semantic Search, International Journal of Computer Application (0975-8887) Volume 129-No.5, November 2015.

[7] Fabrizio Lamberti, Andera Sanna, and Claudio Demartini, A Relation – Based Page Rank Algorithm for Semantic Web Search Engines, IEEE Transactions on Knowledge And Data Engineering, Vol. 21, No. 1, January 2009.

[8] Harith Alani, Christopher Brewster, and Nigel Shadbolt, Ranking Ontologies with Aktiverank, http://swoogle.umbc.edu/.

[9] Shashank Gugnani, Tushar Bihnay, Rajendra Kumar Roul, A Complete Survey On Web Document Ranking, International Journal of Computer Application (0975 8887) Volume ICACEA – No.2, 2014.

[10] Ali I. El-Dousky, Hesham A. Ali, Rabab S. Rashsed, Ranking Documents Based on the Semantic Relation 878673433341 6Using Analytical Hierarchy Process (IJACSA) International Journal Of advance computer science and application Vol.7, No. 2, 2016.

[11] K. SRIDEVI and Dr. R. UMARANI, Ontology Ranking Algorithms on Semantic Web: A Review, International Journal of Advance Research in Computer and Communication Engineering Vol. 2, Issue 9, September 2013.

[12] Zhiguo Ding, Zhengjie Duan, Improved Ontology Ranking Algorithm Based on semantic Web, 978-1-4244-6709-9/10$26.00©2010 IEEE.

[13] S. Brin, and Page L., 'The Anatomy of a Large Scale Hypertextual Web Search Engine', Computer Network and ISDN Systems, Vol. 30, Issue 1-7, pp. 107-117, 1998.

[14] M. Jones and H. Alani, "Content-based Ontology Ranking", The 9th Int. Proteige Conf. Stanford, California, July 2006.

[15] Y. Zhang, W. Vasconcelos, and D. Sleeman, "Ontosearch: An Ontology Search Engine." The 24th SGAI Int. Conf. On Innovative Techniques and Applications of AI, Cambridge, UK, Sep. 2004.

Neera Chaudhary is postgraduate research student in the Department of Computer Science and Engineering at Ambedkar Institute of Advanced Communication Technologies & Research, New Delhi. She received B. Tech degree in Computer Science and Engineering (2017) from Raj Kumar Goel Institute of Technology for Women. Her field of interest is Semantic Web.

Dr. Suresh Kumar is an Associate Professor of Computer Science and Engineering at Ambedkar Institute of Advanced Communication Technologies & Research. He had more than 15 years of experience in teaching as well as in research. He received M. Tech and Ph. D. degree in Computer Science and Engineering. His research interests include Information Retrieval, Semantic Web, Data Mining etc. He guided more than 28 M. Tech. students and three Ph.D. research scholars are under his supervision.

TABLE 4.1. COMPARATIVE ANALYSIS OF ALGORITHM

S. No.	Approach Applied	Method Applied	Merits	Demerits
1	AKTiveRank [3]	It is based on four parameters viz class match, density, Semantic similarity and between-ness measures.	Whole idea of this scheme cover internal structure.	1. Less efficient 2. Time complexity is high 3. Low CEM value [7] 4. Slow processing.
2	Content-based Ontology Rank [3]	• Internal structure and content similarity in ontology in a corpus are basic components.	The ontology is ranked high which has more labels matches with the storage/database.	Difficult in retrieving suitable record if search term is specific.
3	Onto Rank [4]	• Based on google API • Based on link analyze method	Very effective in if and only if association which applies between classes in a relation set [4].	1. Ranking undergoes the problem of poorly inter-referenced among ontologies. 2. Quality of ontologies are degraded.
4	OS Rank [4]	• Focus on class name, semantic relation and ontology structure [4].	1. Effective ranking procedure in both native ontology repository and associated with ontology search engine [4]. 2. Effective methodology on both ontology structure and semantic analysis [4].	1. Time consuming. 2. Highly complex.
5	SIF Rank [3]	• Focus on semantics of concept or relation or ontology structure.	Concepts strengthens with iterative manner.	1. Concept importance retrieved is complex. 2. No two users can give the same priority for a large ontology.
6	SHOE Rank [5]	• Relation based ranking	It make Use of semantic mark-up language for remarks to improve efficiency.	Stand alone
7	SEM Rank [5]	• Relation based ranking	Effective only for small set.	Still to be tested on large set
8	Page Rank [6]	• Hyperlink model based ranking	1. Less time consuming. 2. Feasible 3. Less susceptibility to localized links.	1. Rank sink spider trap problem. 2. Relevance of result is very less.
9	Relational-based Page Rank [5]	• keywords and concepts based relationship	Effective as it used to infers hidden concepts behind keywords.	Lack of automated assignment of rank values to categorical partial tuples.
10	RaRe Rank [5]	Entity based ranking	Very effective when compared to page Rank and HITS algorithm.	May be ambiguous in some scenario.