Ranking of Semantic Search Paths using Personalized Weights for Aiding Web Search

Saravanakumar C, Sendhilkumar S, Geetha T.V

Department of Computer Science and Engineering
Anna University, Chennai- 600 025, India
saravanacsk@gmail.com, {thamaraikumar, tvgeedir}@cs.annauniv.edu

Abstract. The problem associated with web search is the difficulty in identifying/choosing the best page from the huge set of result pages. Any personalized search system will help the user to identify the page(s) relevant to the query by analyzing the pages visited by them. But such systems will not recommend pages that don't occur in user's browsing history. However such pages might be the direct answers to the user query. In order to avoid such problems and to recommend relevant unvisited pages, this paper utilizes the concept of conceptual graphs. This paper focuses towards the automatic construction of a conceptual graph called Personalized Page-View (PPV) graph which intend to provide a conceptual relation between search queries and result pages returned by any existing search engine. The search paths that exist in the PPV graph are identified and ranked for improving the search through the WWW.

Key words. Semantic Web, Web Search, Personalization, Conceptual Graph, Personalized Page-View Graph

1. Introduction

The web creates new challenges for information retrieval. The amount of information of the web as well as the number of new users is growing rapidly. Because of this phenomenal growth, searching for particular information consumes more time and it becomes a daunting task. Traditional search engines obtain the results by matching the words in the query with the document content and finally they give-out thousands of result pages, of which only a handful of them are relevant. Personalization is a popular remedy for the above problem. Personalization improves search by considering the user's interests and hence provides a context-oriented search results, which are the direct answers to the user's information need. Semantic searches on the other hand attempts to augment and improve traditional search results by using semantic information from resources like concept graphs and thesaurus. The individual search behaviors like, the pages visited, the order of visit and the actions performed on a visited page can be used to confirm the context of search derived from the search query. This way, an effective personalization system could decide autonomously whether or not a user is interested in a specific webpage and, in the

Advances in Computer Science and Engineering.
Research in Computing Science 34, 2008, pp. 237-248

Received 23/03/08
Accepted 26/04/08
Final version 04/05/08

negative case, prevent it from being displayed or, the system could navigate through the Web on its own and notify the user if it found a page or site of presumed interest.

Normally users browse through only a few pages among the top twenty pages. However, the pages that go unvisited sometimes prove to be important/ relevant to the user's context of search. Hence personalized search systems that heavily depend on user's browsing history might miss such pages. In order to identify the relevant pages from the unvisited page category this paper focuses towards the development of automatically identified user profile called Personalized Page-View (PPV) graph that provides a conceptual link between visited and unvisited page categories. The PPV graph provides an optimal (shortest) search path that connects the pages that are direct answers to user's information need. A set of conceptual paths are identified and assigned weights. The paths are then ranked according to their weights and then the top ranked paths are recommended to the user. Based on the ranked paths and their weights a conceptual link needs to be provided between concepts in the visited pages and concepts in the unvisited pages. Such conceptual links help a personalized search system to identify relevant pages from the unvisited category. This paper concentrates more on the construction of PPV graph using search queries and visited pages. However, providing a conceptual link between visited and unvisited pages is out of scope of this paper.

2. Related Work

Several attempts for using Conceptual Graphs (CGs) in the semantic web exist. A conceptual graph is a connected graph of various concepts that are semantically inter-related in any particular domain. Corese an ontology-based search engine [1] is one of such kind. Corese retrieves web resources annotated in RDF(S). A query is translated into RDF then into CG. The RDF annotations of web pages are also translated into CGs. Corese provide a user with approximate answers to queries. CGs can be successfully applied for mining transformation rules for semantic matching [2]. The knowledge is represented as CGs and the underlying algorithm discovers transformation rules between CGs describing semantically close assertions. With the perspective of WWW as a labeled graph, SUBDUE is a data mining tool that discovers repetitive substructures in graph-based data [3]. In particular, this graph-based data-mining tool can discover structure formed by a user query within a graph representation of the WWW.

One increasingly popular way to structure information is through the use of ontologies. OntoSeek [4] is an ontology based tool, which is designed for content-based information retrieval from online yellow pages and product catalogs. OntoSeek uses simple conceptual graphs to represent queries and resource descriptions. The system uses the Sensus ontology [5], which comprises a simple taxonomic structure of about 50,000 nodes. The system developed by Labrou, Y. and Finin T. [6] semantically annotates Web pages via the use of Yahoo! categories as descriptors of their content. The system uses Telltale [7] as its classifier. Telltale computes the similarity between documents using n-grams as index terms. The ontologies used in the above examples use simple structured links between concepts. Analysing the

semantic relationships of named-entities [8] can improve relevance in search and ranking of documents. Work of Sheth et. al., [9] shows simple and complex relationships by using predefined multi-ontology relationships for query processing. Anyanwu K and Shetac A. [10] in their work provide querying using ontological concepts, inferencing as part of query answering, and ability to specify incomplete queries through the use of path expressions.

A new search aiding index called User Conceptual Index (UCI) [11] provides a conceptual link between search queries and relevant pages visited by the user. The UCI takes into account the implicit factors of personalization like the page-view time, query as well as page hit counts and query usage time. All the above-mentioned concept based methods does not utilize the powerful concept of personalization. The work explained in this paper is an extension of the work done by S. Sendhilkumar and T.V. Geetha [11]. In this paper, we attempt to modify the UCI by utilizing an important factor of personalization, i.e. the actions performed on a page. Using this, we construct a Personalized Page-View (PPV) graph which recommends to the user, the shortest path to their information need.

3. The Personalized Page-View (PPV) Graph Based System

The PPV Graph based system as shown in figure 1 comprises of the following modules: 1) User Behavior Tracking, 2) Pre-processing, 3) PPV graph construction, 4) Path Weight computation and ranking, and 5) Page Recommendation.

3.1. User Behavior Tracking

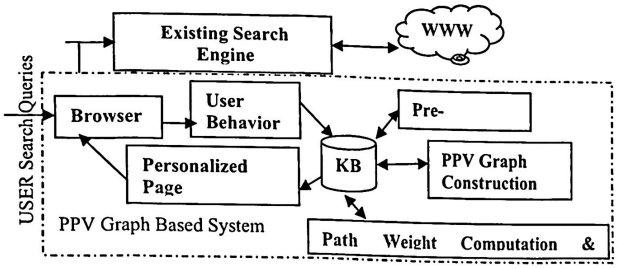

Figure 1. Personalized Page View (PPV) Graph Construction System

The users submit their search queries through a specially designed browser. The browser keeps track of all the user data like the search queries, pages visited, time spent on a page and actions performed (save, copy, print and bookmark) by the users during their search sessions. Thus the user data is collected implicitly from the user end. The user queries are submitted to an existing search engine. The results given by

the existing search engine are analyzed and re-ranked based on user interests and path weights. Finally the personalized search results are recommended to the users in the browser.

3.1.1. Experimental Setup

The number of users involved was six. Among the six four of them were post graduate students in computer science & engineering and the other two were research scholars in the same. Hence they all had at least five years experience of working with computers. All the six users performed regular searches using the new browser developed for this work. Their behavior on every page they accessed was recorded by the browser. The users were also asked to explicitly rate the relevancy of each page which was minimal disruption to their regular work, but necessary for the experiment. The users were asked to select one among the following six options: 0 – No Idea, 1- Not Relevant, 2 – Leads to Useful Link, 3 - Partially Relevant, and 4 – Exactly Relevant. When a user issues a new search query or modifies the previous search query he/she is asked to rate the search as a whole for the given search query using one among the following four options: 0 – No Idea, 1 – Not Useful, 2 – Partially Useful and 3 – Very Useful. A week search information was collected which comprises of 76 search queries, 405 visited pages and top twenty result pages (Total pages = 1520 pages, of which 108 pages were visited and the remaining 1312 pages were left unvisited) returned by existing search engine for each search query.

3.1.2. Pre-Processing

The pre-processing module depends on the data collected and deposited in the knowledge base (KB) by the User Behavior Tracking module. The pre-processing module does the following activities: 1) HTML to text conversion, 2) POS tagging, 3) Noun extraction, 4) Term-Frequency (TF), Inverse Document Frequency (IDF) and weight computation and 5) Feature word selection.

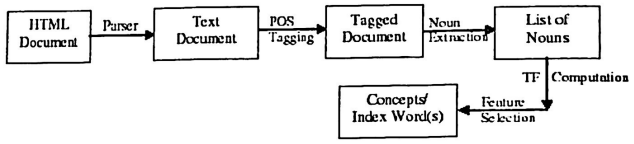

Figure 2. Steps involved in Pre-Processing Module

Every HTML document/page visited by the user is given to a parser, which removes all the HTML tags and then converts the HTML document into a text document. The text document now contains only the textual content of the page. The text document is then tagged with a POS tagger, which gives out a tagged text document. Nouns are the indicators of concept(s) spoken in a page. The nouns are extracted from the POS tagged text document, whose Term-Frequencies (TF) and Inverse Document Frequencies (IDF) are computed. Then the weight for each concept in a page is computed based on the TF and IDFs. The three concepts that have the

highest weight in a page are selected as features/index words/concepts. These top three concepts are used to index the pages. The complete pre-processing steps are shown in figure 2 and the whole pre-processing process is repeated for all the pages that are visited by the user. Thus every page is semantically tagged with the concepts that represent the content of the page.

3.2. PPV Graph Construction

The Index Words/Concepts thus extracted are used to build the user profile which is represented as a PPV graph. User profiles store approximations of the interests of a given user. User profile creation and updating is very essential for a personalized web search because personalization will be effective only when the user's interests are incorporated into the search. User's interests are identified automatically from the collection of various search queries and the relevant pages visited by the user. The terms in user's search queries, Index Words of the visited pages, frequency of usage of similar search terms are implicit indicators of user's interest on a particular concept. Also the user profile contains the user's favorite pages, which are identified by means of the frequency of visits to a particular page. The methodology proposed in this paper for the generation of user profiles differs from the majority of other approaches in that the profiles are: 1) generated automatically, without explicit user feedback, 2) dynamic, i.e., not based on a fixed period of time, but it evolves over time and 3) they are represented by Personalized Page-View (PPV) graph. The content relation between the various pages visited by the user is another important factor that can be utilized for personalization. Deeper semantics on the page content can be inferred with the help of concept graphs that represent the visited pages.

The PPV graph is an incremental graph and it is constructed based on the set of pages visited by the user and actions (like save, print, copy, and bookmark) performed on those pages. The page link weights based on user's actions are computed and updated in the PPV graph. The PPV graph is mainly generated to exploit the semantic relation between the various pages visited by the user during a search session. Also every path in the PPV graph is assigned with a link weight. The link weight is computed based on the following factors: 1) The Concept (Ci) that a page speaks about and 2) User actions on a page.

3.3. Automatic Creation of Personalized Page-View (PPV) Graph

The PPV graph is constructed automatically from the set of pages visited by the user. The processes involved in the construction of PPV graph is shown in figure 3. The Concepts/ Index Words extracted in the preprocessing module are the input to the PPV graph construction module. WordNet is used for extracting the relations between the various concepts that represent various pages visited by the users. The extracted concepts and relations are given to Protégé Tool for building the conceptual graphs.

The PPV graph thus constructed is an incremental graph since it is updated with the new concepts visited by the user in the future.

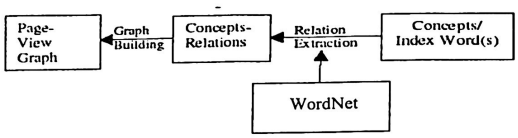

Figure 3. Steps in Personalized Page-view (*PPV*) graph Construction

The Concepts/Index Words extracted in the preprocessing layer and the relationships among the Concepts/Index Words which were extracted using WordNet are given to the Protégé tool for constructing the conceptual graphs. For sample queries like "Heart Attack Causes", "Heart Failure Causes", "Drugs for Heart Attack" and "Blood Pressure" issued in a search session, pages that were visited by the user are collected. Few of the sample Index Words/Concepts that are extracted like {PRESSURE, BLOOD, HEART, STRESS, DRUGS, ASPIRIN, BLOCKER, HYDROCHLORIDE HEART_FAILURE, STROKE, KIDNEY_FAILURE, HEART_ATTACK, and BLOOD_PRESSURE} represent the pages visited by the user during his/her search session. These index words and the extracted relations/properties like {"Leads to", "Drugs for"} are given to the Protégé tool.

```
<owl:ObjectProperty rdf:ID="leads_to">
  <rdf:type
rdf:resource="http://www.w3.org/2002/07/owl#FunctionalPro
perty"/>
  <rdfs:domain
rdf:resource="#HIGH_BLOODPRESSURE"/>
  <rdfs:range>
   <owl:Class>
    <owl:unionOf rdf:parseType="Collection">
     <owl:Class rdf:about="#HEART_ATTACK"/>
     <owl:Class rdf:about="#HEART_FAILURE"/>
     <owl:Class rdf:about="#KIDNEY_FAILURE"/>
     <owl:Class rdf:about="#STROKE"/>
```

Figure 4. Code for the property "HIGH_BLOODPRESSURE"

```
<owl:Class rdf:ID="HEART_FAILURE">
        <rdfs:subClassOf>
<owl:Class rdf:ID="MAKES"/>
<owl:Class rdf:ID="STROKE">
        <rdfs:subClassOf rdf:resource="#MAKES"/>
</owl:Class>
<owl:Class rdf:ID="KIDNEY_FAILURE">
        <rdfs:subClassOf rdf:resource="#MAKES"/>
</owl:Class>
<owl:Class rdf:ID="HEART_ATTACK">
        <rdfs:subClassOf rdf:resource="#MAKES"/>
</owl:Class>
```

Figure 5. Code for the subclass "MAKES"

The sample code that was generated by the Protégé tool is given in figure 4 and 5 respectively. The sample PPV graph constructed for the query "Blood Pressure" is given in figure 6. The code in figure 4 shows that HIGH_BLOODPRESSURE property "leads to" HEART_FAILURE, STROKE, KIDNEY_FAILURE and HEART_ATTACK. The code in figure 5 shows that HEART_FAILURE, STROKE, KIDNEY_FAILURE and HEART_ATTACK are the subclass of MAKES. The subclasses of MAKES are grouped by *UnionOf* tag. Thus for the extracted concepts various weights associated with path-weight computation is explained in the following sections.

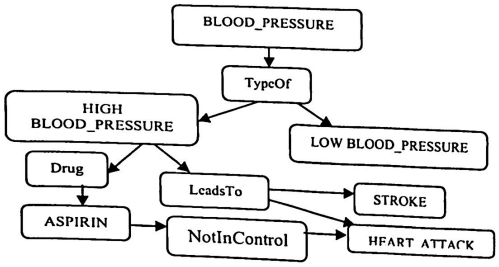

Figure 6. Sample Page-View Graph for the query "Blood Pressure"

3.4. Path Weight Computation and Ranking

A typical search can result in many semantic paths semantically linking the entities of interest. Because of the expected high number of paths, it is likely that many of them would be regarded as irrelevant with respect to the user's domain of interest. Thus, the semantic associations need to be filtered according to their perceived relevance. Ranking approach defines a path rank as a function of various intermediate weights. The weights involved in the path-weight computation are: 1) Concept Weight (C_i), 2) Path Length Weight (PL_i) and 3) Personalized Context Weight (C_P).

3.4.1. Concept Weight

When considering concepts in graph, those that are lower in the hierarchy can be considered to be more specialized instances of those further up in the hierarchy. For the sample blood pressure hierarchy given in figure 6, the HEART_ATTACK conveys more meaning than BLOOD_PRESSURE, HIGH_BLOOD_ PRESSURE. Higher weights are assigned to more "specific" semantic associations because they convey more meaning then "general" associations. The weight for a concept in the hierarchy is computed as in equation 2.

$$\text{Concept Weight, } C_i = H_{ci} / H \qquad (2)$$

where, H_{ci} - level of i^{th} concept in the PPV hierarchy; H - total height of the hierarchy.

$$\text{BLOOD\_PRESSURE } C_1 = H_{c1} / H = 1/6 = 0.17$$
$$\text{HIGH\_BLOOD\_PRESSURE } C_2 = H_{c2} / H = 3/6 = 0.50$$
$$\text{LOW\_BLOOD\_PRESSURE } C_3 = H_{c3} / H = 3/6 = 0.50$$
$$\text{ASPIRIN } C_4 = H_{c4} / H = 5/6 = 0.83$$
$$\text{HEART\_ATTACK } C_5 = H_{c5} / H = 5/6 = 0.83$$
$$\text{STROKE } C_6 = H_{c6} / H = 5/6 = 0.83$$

3.4.2. Path Length Weights

The path weights are context specific. For some queries, a user may be interested in the most direct paths (i.e., the shortest path – a link from one page leading to the most other relevant page(s)). This may infer a stronger relationship between the concepts in two different pages. Hence, path length must be determined and should be used. The path length weight PL is computed using equation 3.

$$\text{Path length weight } PL_i = 1/|C| \qquad (3)$$

where, |C| is the number of components in the path P (excluding the start and end concepts). A component in a path refer to both the concept and relation.

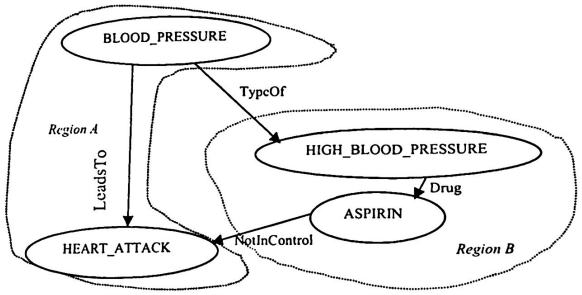

Figure 7. Path length examples

The sample paths given in figure 7 show that there exists two paths, 1) a direct path PL_1 in region A(from the page that contains the concept BLOOD_PRESSURE the user has clicked a link leading to another page containing the concept HEART_ATTACK which is the direct answer to the user's information need) and 2) a longer path PL_2 in region B (from the page that contains the concept BLOOD_PRESSURE the user visited another page containing the concept HIGH_BLOOD_PRESSURE, which in turn linked the user to a new page about a drug ASPIRIN and finally the user moved from ASPIRIN page to the relevant page that speaks about HEART_ATTACK which is the direct answer to the user's information need) between BLOOD_PRESSURE and HEART_ATTACK. The path length weights are computed as given below,

For the direct path $PL_1 = 1/C = 1/1 = 1$

For the longer path $PL_2 = 1/C = 1/5 = 0.20$

Thus it can be observed that $PL_1 > PL_2$, hence the longer paths are ranked lower than the shortest paths.

3.4.3. Personalized Context Weight

The user can perform any of the actions mentioned in table 1 on any page which, he/she visits. According to the action performed weights are assigned to the respective pages. The user has performed the action "Save" on the page represented by the concept/index word HEART_ATTACT, hence, a weight 1 is assigned to that page. Similarly the user has performed "copy" part of the pages represented by the concepts HIGH_BLOOD_PRESSURE and ASPIRIN, hence, a weight 0.25 and 0.5 is assigned respectively.

Table1. User Actions and respective weight

Actions	Page Relevancy to user's context of search	Weights Assigned
Save	Highly relevant	1
Print	Highly relevant	1
Copy	Partially relevant	0.25 to 0.75 depending upon the amount of the content copied
Book Marking	High/Partial according to the page usage	1 – used in same session 0.5 - used in another session

The weight assignment highlights that the user is more interested in the page represented by the concept/index word HEART_ATTACK and also wants to consider the associations between BLOOD_PRESSURE and DRUGS, but with lesser priority. Now the personalized context weight C_P for a path traversed by the user while searching for his/her information need can be calculated by equation 3,

$$C_P = \frac{1}{|C|}\left[\sum_{i=1}^{m} r_i * Ci \in R\right] \qquad (3)$$

where, r_i is the weight assigned according to the action performed and the page-view time on a page P_i (if no action performed ignore ri and just use the concept weight alone), C_i is the concept weight in the path P, m is the maximum number of concepts along a path P and |C| is number of components in the path (again excluding the start and end entities). That is, for each concept that P passes through, sum the total number of components in P that are in the region R_i and multiply it by the weight attributed to that region, r_i.

The Overall Path Weight of a path P denoting a semantic association between two pages P_i and P_j will be a linear function as in equation 4.

$$W_{Pi \to Pj} = C_P + P_L \qquad (4)$$

Some of the sample relations that can be extracted from the figure 6 are:

Relation 1: BLOOD_PRESSURE<LeadsTo>HEART_ATTACK

Relation 2: BLOOD_PRESSURE<TypeOf>HIGH_BLOOD_PRESSURE <Drugs>ASPIRIN<NotInControl>HEART_ATTACK

Relation 3: BLOOD_PRESSURE <TypeOf> HIGH_BLOOD_PRESSURE<LeadsTo> HEART_ATTACK

These semantic relations are then ranked according to the overall path weight $W_{Pi \rightarrow Pj}$ and table 2 depicts the weight computation for the extracted semantic relations and the ranks assigned according to the computed weights. Consider the above semantic relation 1 and 2 in region A and region B respectively from the figure 7. Let us assume that the user in interested in region A. While computing the personalized weight the page view time is added with the action weight. The page view time (in seconds) is normalized by dividing it by 1530 seconds (25.5min) which is the average session out time. From table 2 it can be observed that the overall weight of Relation 1 is greater than that of relation 2 and according to their weights, ranks has been assigned.

Such kinds of relations between the various concepts from the collected web pages provide semantic relations between the pages. Also the constructed graph is very useful to analyze the relationships between the collected most frequently occurring patterns and the links available in a web page. Thus the extracted semantic relations from the domain ontology can be ranked according to the user actions and can be used in the page re-ranking process. The PPV graph that was developed for the personalized web search is thus different from the others by the above mentioned concept weights. The various semantic paths and the respective path ranks are updated into the KB which will be used for further page recommendation process.

Table 2. Weight Assignments and Ranking of Semantic Relation

Semantic Relation No.	Path Length Weight (PL_i)	Personalized Context Weight (C_P)	Overall Path Weight $W_{Pi \rightarrow Pj} = PL_i + C_P$	Rank
	= 1/1	= 1/1 *(0.17+1*0.83)	= 1+ 0.1411	1
1	= 1	=0.1411	= 1.1411	
		[No action performed in page represented by C1 ; hence ignore r1]		
2	= 1/5	=1/5 *	=0.20+0.318	2
	= 0.20	(0.17+0.25*0.50+0.5*0.83+1*0.83)	=0.518	
		= 0.318		

For the constructed conceptual graph the following metrics 1) Total number of paths (TNOP) 2) Total number of relations (TNOR) 3) Information content for the concepts "HIGH_BLOODPRESSURE" and "HEART_ATTACK" are calculated and visualized in a graph. TNOR is the sum of relations of each concept and is equal to 8 for "HIGH_BLOODPRESSURE" concept and 7 for "HEART_ATTACK" concept in

WordNet which is shown in Table 3. TNOP is the sum of paths of each concept and is equal to 20 for "HIGH_BLOODPRESSURE" concept and 31 for "HEART_ATTACK" concept in the PPV graph constructed that is shown in Table 3.

Table 3. TNOP and TNOR metric

Concept	WordNet		PPV graph	
	TNOP	TNOR	TNOP	TNOR
HIGH_BLOODPRESSURE	20	8	10	6
HEART_ATTACK	31	7	11	10

Information Content denotes the level of information content a concept conveys and can be calculated for a concept c using the formula below

$$IC(c) = \log [(hypo(c) +1)/max.]/ \log(1/ max_{wn})$$

$$= 1- \log(hypo(c)+1)/\log(max_{wn})$$

where hypo(c) is the number of hyponyms of a concept and max_{wn} is the maximum number of concepts. IC of "HIGH_BLOODPRESSURE" concept is found to be 0.386 and "HEART_ATTACK" concept is 0.3007 in WordNet and IC for the concepts "HIGH_BLOODPRESSURE", "HEART_ATTACK" in PPV graph are found to be 0.3701 and 0.3472.These IC values show the content level of concepts in the whole PPV graph.

4. Issues and Limitations

The PPV graph is incremental in nature, hence some powerful inference mechanisms and managing tools are essential to manage the large conceptual graphs. The goodness of the graphs constructed needs to be evaluated before using them for further analysis. Any search system that eliminates the above issues can emerge as a powerful personalized search system. Also the page recommendations according to the personalized rankings will be our future work.

5. Conclusion

The aim of this work is to perform personalized web search by recording user profile for users from their browsing pattern and to retrieve more relevant and related documents that are semantically related to the given search query. To achieve this, it is essential to know the meaning and context of search query as well as pages visited. To understand the semantics of the search query and the pages visited, conceptual graph is developed. The construction of Personalized Page-View graph is automatic and hence does not require any human interference. Personalization using such

conceptual graphs can produce better results as compared to keyword-based searching by providing conceptual link between visited and unvisited pages and thus pages that are unvisited but relevant can also be recommended to the users.

References

1. Corby O., Dieng-Kuntz R. and Catherine Faron-Zucker: Querying the Semantic Web with Corese Search Engine. Proceedings of the 16th European Conference on Artificial Intelligence (2004) 705-709
2. Yeh, P., Porter, B., and Ken Bakker: Mining Transformation Rules for Semantic Matching. Proceedings of the Workshop on Mining Graphs, Trees and Sequences (MGTS'04). Pisa, (2004) 83-94
3. Cook, D. J. and Holder, L. B.: Graph-Based Data Mining, IEEE Intelligent Systems, 15(2), (2000) 32-41
4. Guarino N., Masolo C., and Vetere G., OntoSeek: Content-Based Access to the Web, IEEE Intelligent Systems 14(3), (1999) 70-80
5. Knight, K. and Luk, S.: Building a Large Knowledge Base for Machine Translation. Proceedings of the twelfth national conference on Artificial intelligence, vol. 1 (1994) 773-778
6. Labrou, Y. and Finin T.: Yahoo! As an Ontology – Using Yahoo! Categories to Describe Documents. Proceedings of the 8th International Conference on Information and Knowledge Management (1999) 180-187
7. Pearce, C. and Miller, E.: The Telltale Dynamic Hypertext Environment: Approaches to Scalability. Intelligent Hypertext: Advanced Techniques for the World Wide Web, Springer-Verlag Vol. 1326, (1997) 109 – 130
8. Boanerges Aleman-Meza, Chrisian Halaschek-Wiener, I.Budak Arpinar, Cartic Ramakrishnan, and Amit.P.Shetch: Ranking Complex Relationships on Semantic Web, IEEE Internet and Computing, Volume 9, issue 3, (2005) 37-44
9. Sheth .A, Budak Arpinar and Kashyap .V: Enhancing the power of Internet: Relationships at the heart of the semantic web: Modeling, discovery an exploiting complex semantic relationships, (2004) 63-94
10. Anyanwu .K and Shetac .A: ρ – Queries : Enabling Querying for Semantic Associations on the Semantic web, In the Proceedings of the 12th International WWW conference, ACM press (2003) 690-699
11. S. Sendhilkumar and T.V. Geetha: Personalized Web Search Using Enhanced Probabilistic User Conceptual Index, Journal of Intelligent Systems, Vol. 17, No. 1-3, (2007) 199-213
12. Catledge L. and J. Pitkow: Characterizing Browsing Behaviors on the World Wide Web, Computer Networks and ISDN Systems, 27(6) (1995) 1065-1073