

WEB MINING BASED FRAMEWORK FOR ONTOLOGY LEARNING

C.Ramesh¹, K.V.Chalapati Rao¹, A.Govardhan²

¹Department of Computer Science and Engineering, CVR College of Engineering,
Ibrahimpatnam, R.R.District, Telangana, India.

hmcr.ramesh@gmail.com, chalapatiraokv@gmail.com

²School of IT, JNT University Hyderabad, Hyderabad,
Telangana, India.

govardhan_cse@yahoo.co.in

ABSTRACT

Today, the notion of Semantic Web has emerged as a prominent solution to the problem of organizing the immense information provided by World Wide Web, and its focus on supporting a better co-operation between humans and machines is noteworthy. Ontology forms the major component of Semantic Web in its realization. However, manual method of ontology construction is time-consuming, costly, error-prone and inflexible to change and in addition, it requires a complete participation of knowledge engineer or domain expert. To address this issue, researchers hoped that a semi-automatic or automatic process would result in faster and better ontology construction and enrichment. Ontology learning has become recently a major area of research, whose goal is to facilitate construction of ontologies, which reduces the effort in developing ontology for a new domain. However, there are few research studies that attempt to construct ontology from semi-structured Web pages. In this paper, we present a complete framework for ontology learning that facilitates the semi-automation of constructing and enriching web site ontology from semi structured Web pages. The proposed framework employs Web Content Mining and Web Usage mining in extracting conceptual relationship from Web. The main idea behind this concept was to incorporate the web author's ideas as well as web users' intentions in the ontology development and its evolution.

KEYWORDS

Ontology Learning, Web Mining, Web Content Mining, Web Usage Mining, Ontology Evaluation

1. INTRODUCTION

World Wide Web, since its conceptual inception, has contributed greatly for the knowledge era, in which we are living today. As conceptualized by Sir Tim Berners-Lee, the introduction of World Wide Web (WWW) has given rise to enormous amount of information that can be accessed in digital form and most of these data are in the form of documents. The exponential growth of these documents has raised many challenges. Considering the structure of these

documents, we find that they are not descriptive enough to express themselves, overloaded with information and distributed all over the Web. Therefore, it has become a difficult task for the Web Users to search and retrieve the relevant information needed for them.

Semantic Web, as envisioned by Sir Tim Berners-Lee, addresses this problem by giving information a well-defined meaning, better enabling computers and people to work in co-operation. Semantic Web is implemented using W3C recommended Semantic Web Technologies and Standards and expresses the Web data in a machine-understandable and machine processable form, thereby supporting information exchange and sharing between applications. Ontologies play a significant role in building Semantic Web and provide a platform for promoting Semantic interoperability on the Web. However, constructing ontology's for the many and varied domains on the Web is a time-consuming process and their construction is a bottleneck to the wider deployment and use of Semantic Information on the Web. Since manual construction of ontology is costly, time-consuming, error-prone and inflexible to change, it is hoped that an automated or semi-automated process will result in better ontology construction and create ontologies that better match a specific application [1].

There have been several research attempts to automate ontology construction and update process by exploiting the content of Web pages. Most of the Web documents that exist today are in semi-structured format. However, there are few references to research attempts that focus on these semi-structured data on Web [2] [3] [4]. Further most of these research attempts use text mining and Natural Language Processing techniques to extract the semantics from Web documents, neglecting the embedded information in the semi-structured nature. Also most of the current approaches are dealing with some specific tasks or a part of the ontology learning process rather than providing complete support to users. There are few research attempts that use Web mining techniques such as Web Content Mining and Web Usage Mining in ontology development.

The benefits of analyzing the usage behavior analysis have been the driving forces for continuous research in the realm of Web Usage Mining, which aims at discovering navigational patterns from the logs of HTTP requests for Web resources [5]. Further Web Content Mining aims to extracts/mine useful information or knowledge from Web page contents. The benefits offered by these two techniques in Web Mining applications are noteworthy.

In this paper, we present a framework for Ontology Learning from Semi structured Web pages using the combined techniques of Web Mining namely, Web Content Mining and Web Usage Mining. We have employed the Web Content Mining to extracts the concepts and further discover the Conceptual relationships from Web pages. We applied the text mining techniques and extended Apriori Algorithm, which is most widely used for frequent mining, for extracting the concepts. The Semantics extracted from Web Usage Mining process, helps in refining the conceptual relationships extracted from Web Content Mining. Further the refined conceptual relationships are also used in enriching the Web site Ontology. Ontology Pruning and Ontology evaluation are other stages of Ontology Learning process.

The remainder of this paper is organized as follows. In section II, we present a survey of current research efforts on Ontology Learning and Web Mining Methods. In section III, we present our Ontology Learning framework and its main architectural components. In section IV, implementation and experimental results are discussed. In section V, enriched Ontology is evaluated. Finally, in conclusion, some plans for future work are presented.

2. RELATED WORK

“Ontology is an explicit, formal specification of a shared conceptualization of domain of interest [6], where formal implies that the ontology should be machine readable and the domain can be conceptual thing that is shared by a group or community”. During the last decade, several research attempts on ontology learning and systems have been proposed. These research efforts tried to build ontology in either of two ways. One way is using ontology development tools [8] like protégé and Onto-Edit. Knowledge engineers and Domain experts use these tools to build the ontology. Another one is semi-automatic way of constructing the ontology by learning it from different information sources [9] [10] with little human intervention.

Ontology learning refers to a process of applying various knowledge discovery techniques in constructing ontology by extracting concepts and relations using different input sources. It aims at building ontologies semi automatically or automatically from a given text corpus with a limited human exert. Ontology learning can also be defined as a set of methods and techniques used for building ontology from scratch, enriching or adapting an existing ontology in a semi-automatic fashion using several sources [9]. Ontology learning has recently been studied as an effective approach to facilitate the semi- automatic development of ontologies. Ontology learning use techniques and methods from diverse spectrum of fields such as machine learning, knowledge acquisition, natural language processing, information retrieval, artificial intelligence, reasoning and database management systems[11][9].

Manual construction of ontologies is costly, time-consuming, error-prone and inflexible to change. Ontology learning systems can be categorized according to the type of data from which they are learned. Unstructured, fully structured and Semi-structured types of data especially form the input sources to ontology learning systems. In literature, there are several research attempts, focusing on constructing ontology for semi-structured Web Pages using various techniques. Research attempts that focus on unstructured Web pages [12][13][14][1] with free text, mostly use Natural Language Processing techniques and simple text mining in the ontology development. The research attempts that focus on fully structured Web Pages, such as Wikipedia, move beyond simple text mining and take into account the structure of the Web pages [15][16]. However, there are only few research efforts that focus on extracting Semantics from semi-structured Web pages.

The work presented in [3] was the first attempt to discuss the synergy between Semantic Web and Web Mining. They discussed the role of Web Mining techniques in facilitating ontology development. They claimed that the synergy between Semantic Web and Web Mining will give rise to a form of close loop learning, by allowing the usage of Web Mining to extract Semantics and building the Semantic Web and then using the Semantic structures in increasing the performance of Web Mining results. The work presented in [4] draw attention of researchers to use the mark up tags of HTML pages to be used in Web Content Mining to facilitate Ontology development. Descriptions of various techniques provided by Web Usage Mining in improving site Semantics and supporting the users in their navigation is well presented in [2].

A framework for Web Usage driven adaptation of the Semantic Web is well presented in work [17]. The adaptation process employed in the framework, exploited the Web access logs of the users, together with the semantic aspect of the Web to facilitate the Web browsing. Based on the usage of Web, they performed evolution of Web site ontology and topology. However in their

work, mining the content of the Web pages was not considered to full extent in extracting concepts needed for facilitating the ontology development.

In another approach [18], similar to our work, has presented a framework that combines Web Content Mining with Web Usage Mining to extract conceptual relationships hidden in semi-structured Web pages and used in ontology development. The main idea behind this concept was to incorporate the Web author's ideas as well as Web Users' intentions on Web site in the ontology development. The above research attempts to use Natural Language Processing and Association rule mining to extract the conceptual relationships. However, a complete ontology learning process was not presented and much focus was given only to ontology creation.

A semi-automatic method for domain terminology extraction from e-learning resources and their organization as ontology is well described in [19]. However, the work is limited only to e-learning domain and used mostly the Natural Language Processing techniques. Few research works that try to use the semi structured nature of the Web pages in ontology development have become specific to special type of Web sites such as template driven Web sites [20].

Research work [21] made use of the structure of phrases appearing in the HTML documents' headings, in identifying new concepts and taxonomical relationships. However, in most of the current research works, plain text is extracted from the semi structured Web pages as part of preprocessing phase and simple text mining techniques are applied on the extracted free text to construct ontology. Here the ontology development process has not considered the users' intentions and aspirations on Web site.

3. ARCHITECTURE

The main aim of the paper is to investigate and develop a framework that enables ontology learning by partially automating the process of extracting conceptual relationships from semi structured Web pages using Web Mining techniques. In this section, we present the overall architecture of our Ontology Learning framework. Figure.1 shows the architecture of our proposed Ontology Learning framework, consisting of four stages.

They are :

- i. Mining the Web Page Contents to extract the Concepts and Conceptual relationships,
- ii. Mining the Web Usage data to extract hidden Conceptual knowledge and refine the Conceptual relationships discovered in step one,
- iii. Ontology construction and
- iv. refining the Web site ontology. The input for the proposed Ontology Learning framework consists of site Web pages, server's access logs, the initial domain ontology of the Web site.

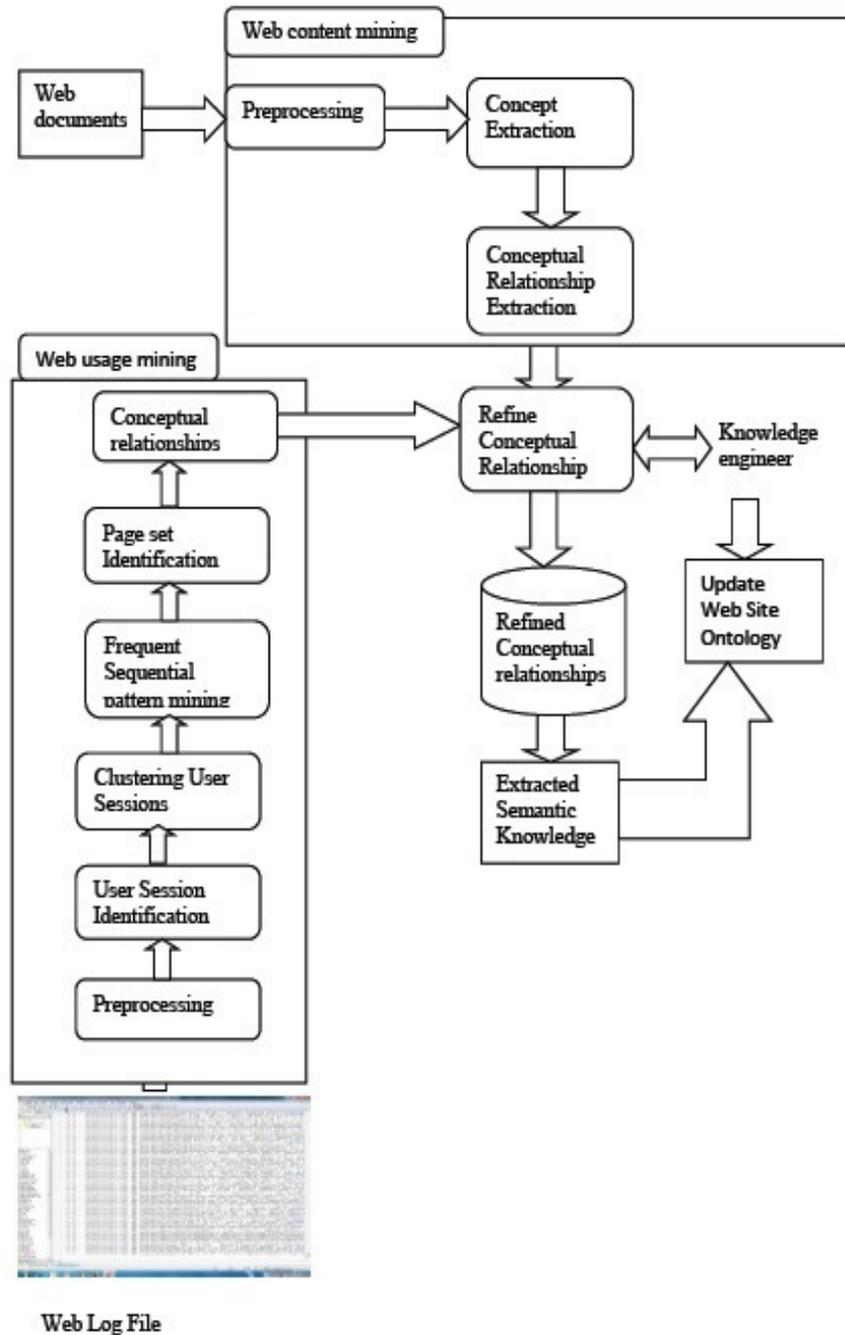


Figure.1. Architecture of Ontology Learning Framework

3.1 Mining the Web Page Content to Extract the Knowledge

Usually the Web page contents are organized from Web designer / Web author perspective. Mining the web page contents can reveal the conceptual relationships as seen by Web author. However extracting information directly from Web pages is a difficult task since most of the Web pages do not conform to HTML syntax. The ill-formed HTML pages need to be preprocessed and parsed before applying the concept extraction process. A Java based SAX parser is used to parse the Web pages and convert them into a formal structure. The Web pages are annotated with parts of speech tags. Weighted Frequency of the concepts is determined by considering the frequency of the concepts as well as the frequency of HTML markers containing those concepts. Here different HTML tags are given different weights to match their importance.

The concepts that have a weighted frequency higher than externally specified threshold values are considered as most significant set of concepts of that Web page. One or more keywords of sentence in Web page may define a concept. An extended Apriori algorithm [22] was used to determine the significant concept sets, while pruning the word sequences from the candidate word sequences that have no probability of selecting as a concept. Concept sets are generated using the above process iteratively.

Associations between the concepts are identified as the concepts that together occur frequently. These associations between the concepts hint the existence of conceptual relationships. The identified associations mostly represent the conceptual relationships that exist in Web author or Web designer mind. These extracted set of conceptual relationships are presented to the ontology developer for further refinement where he/she can include any new conceptual relationships or remove irrelevant ones from the extracted conceptual relationships and refine the existing Web site ontology. Association rule discovery techniques are used in extracting the frequent concept sets. The most widely used, most popular CloSpan algorithm [23] was employed for extracting frequent concept sets. Conceptual relationships are determined from the generated association rules.

3.2 Mining the Usage Patterns to Extract Conceptual Relationships

Web Usage Mining refers to the process of extracting users' navigational patterns by applying data mining techniques on Web access log files. Users' Web Browsing activity is heavily dependent on his needs, knowledge and interests. Users' view on Web pages could be different from Web site author views. Mining the Usage patterns could reveal the conceptual relationships that reside in the web pages as per Web users' perspective. Web Usage patterns are used in applications like Web Personalization, Web caching, Web perfecting, Web site restructuring and intelligent online advertisements [24].

Web Users browsing preferences could be learned and adopted in the Web adaptation process to suit the users' needs. The Proposed framework uses Web Usage Mining to extract conceptual relationships that could be learnt about the Web pages according to the discovered usage patterns. The extracted Semantics is used in the conceptual relationships' refinement stage along with the conceptual relationships extracted by mining the Web content of the Web pages. Web Usage Mining alone cannot be used in extracting the Semantic Knowledge from Web access logs as the users' navigational patterns would be insufficient in case of dynamic Web sites where the content of the Web pages changes frequently.

Web Usage Mining process mainly includes steps like preprocessing the web log files, User Session Identification, discovery of frequent Sequential Patterns, Analysis of the Usage patterns and uses the discovered patterns in various applications.

3.2.1 Preprocessing the Access Log Files

The irrelevant information that exists in the raw Web access log files has to be removed before applying the Mining techniques. Here various preprocessing tasks are done. Removing duplicate records and irrelevant requests such as request with response status code greater than 200 and removing records related to image requests are done as part of the preprocessing task.

3.2.2 User Session Identification

After preprocessing phase, user sessions are identified. We used a heuristic measure in performing sessionization. An idle time of 30 minutes is considered in constructing user sessions. The identified user sessions are then mapped into multidimensional vector space model of URL references (bit vector). We represented each Web page visited as '1' and each Web page not visited as '0' while mapping the user sessions into a vector space model. Table.1 illustrates the user sessions mapping into multi dimensional vector space model.

Table 1. Example of User Sessions Mapping to Multi dimensional Vector space

User Session	Web Transaction Set
S1 = < p1,p2,p4,p5>	W1 = <1,1,0,1,1>
S2 = <p2,p3,p5>	W2 = <0,1,1,0,1>
S3 = <p1,p3,p5>	W3 = <1,0,1,0,1>

The constructed vector space is clustered using K-means clustering algorithms. Each cluster represents a group of Web transactions that are similar based on the co-occurrences of the URL references. The results are presented to the ontology developer who specifies the number of clusters to be generated. Sequential association rule mining is applied on the cluster sessions. Table.2 shows an example of the cluster details.

Table 2. Example of a Cluster details

Property	Value
1	{(1,0,0,0)(1,1,0,0)}
2	{(1,1,1,1) (0,0,0,1)}
3	{(1,0,0,1)}

3.2.3 Sequential Frequent Pattern Mining

Page sets are extracted using association rules. Based on the extracted page sets, conceptual relationships are identified and then presented to the ontology developer for suggestions. The ontology developer extracts useful conceptual relationships, which refine the Web site ontology.

Then the extracted information has to be converted into machine understandable format. Owl is used to represent the Semantic information.

4. EXPERIMENT AND RESULTS

Experiments are conducted on an anonymous University Web site. We have collected the Web access log file over a period of one month from University Server. For performing experiments, we used domain ontology of the same anonymous University Web site. Figure.2 shows the snapshot of initial domain ontology of the University Web site. The size of raw Web log file collected was nearly around 25540 page views. After preprocessing the log, the Number of page views, are reduced to 6892. The Number of User Sessions obtained were 600 on an average basis of 80 sessions per day.

The ontology was edited and visualized using tool Protégé'4.3 [25]. OWL language was used in representing the updated ontology. After preprocessing task, User sessions were identified. K-means Clustering algorithm is employed to generate clusters over generated User Sessions. CloSpan algorithm was implemented on the usage clusters to generate frequent sequential concept sets.

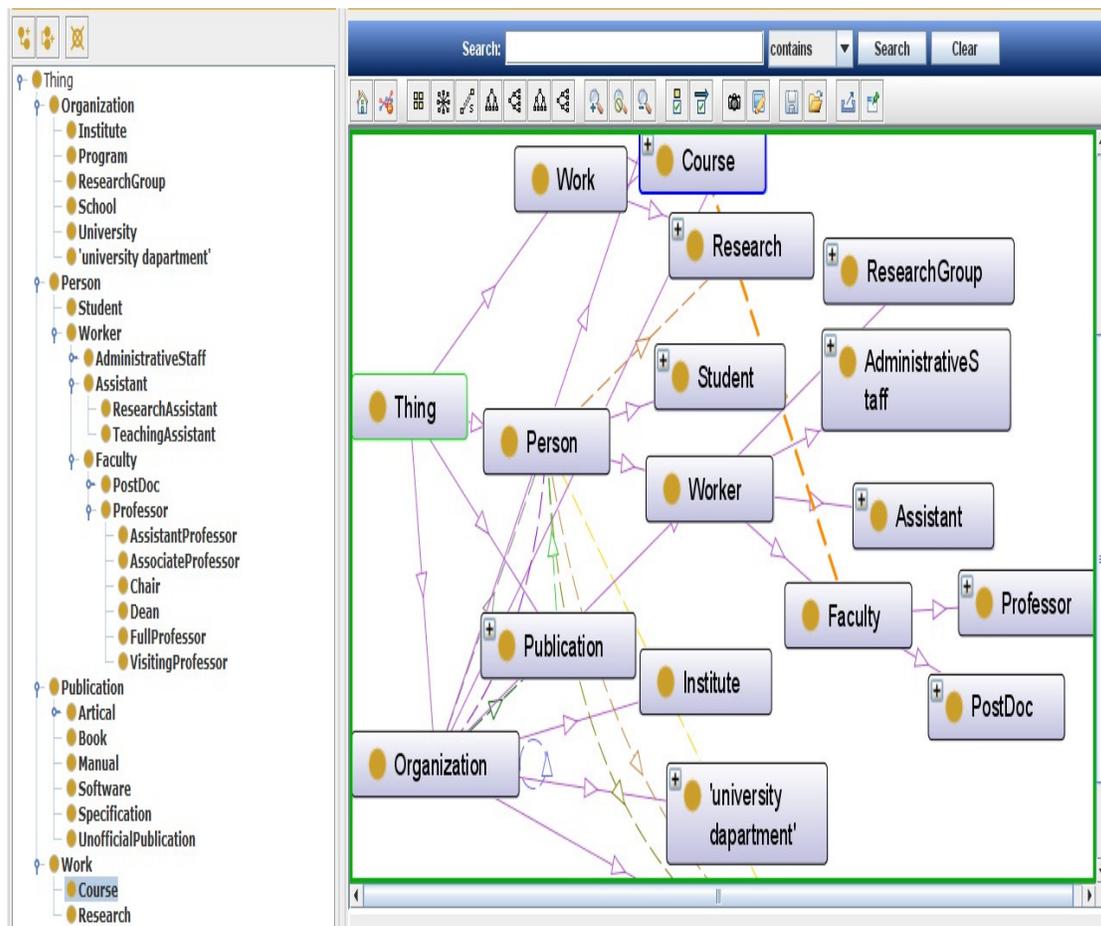


Figure 2. Domain Ontology of an Anonymous University Web site

We report in this section, some of the Sequential Association rules extracted in the Web Usage Mining process.

Pattern 1:

```
/person/student.html -> /person/worker/faculty/professor.html
-> /person/worker/teachingasst.html
```

Support: 0.02537

This behavior has a support of 2.53%. It corresponds to approximately 147 users of the web site. These users are likely to be interested in looking for teaching assistant. The support count of the above Association rules, hints that the existence of strong relationship between 'student' concept and 'teachingassistant' concept.

Pattern 2:

```
/person/worker/faculty/lecturer.html-> /person/worker/assistant/researchasst.html
-> /organization/researchgroup/course.html
```

Support: 0.02245

This behavior has a support of 2.20%. It corresponds to approximately 120 users of the web site. These users are likely to be interested in browsing research pages after viewing lecturer pages. It hints the existence of association between lecturer and research. Figure.3 shows the snippet code of the updated ontology in OWL representation.

```

3  xmlns = "http://www.jntuh.ac.in/~cse/2015/0401/univ-bench.owl#"
4  xml:base = "http://www.jntuh.ac.in/~cse/2015/0401/univ-bench.owl"
5  xmlns:rdf = "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
6  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
7  xmlns:owl="http://www.w3.org/2002/07/owl#">
8  <owl:Ontology rdf:about="">
9  <rdfs:comment>An university ontology for benchmark tests</rdfs:comment>
10 <rdfs:label>Univ-bench Ontology</rdfs:label>
11 <owl:versionInfo>univ-bench-ontology-owl, ver April 1, 2015</owl:versionInfo>
12 </owl:Ontology>
13 <owl:Class rdf:ID="Employee">
14 <rdfs:label>Employee</rdfs:label>
15 <owl:intersectionOf rdf:parseType="Collection">
16 <owl:Class rdf:about="#Person" />
17 <owl:Restriction>
18 <owl:onProperty rdf:resource="#worksFor" />
19 <owl:someValuesFrom>

```

Figure 3. Snapshot of an updated ontology in OWL representation

Reorganization of the concepts hierarchy was performed. For instance, the sub-concept “university research assistant” was previously not under the concept “Student “. However, the high frequency with which these two concepts were occurring together hints the existence of conceptual relationship between them and reflects the interest of the users. The result ontology is represented in OWL language. The resulting ontology after the application of changes as suggested by the system is shown in Figure 4.

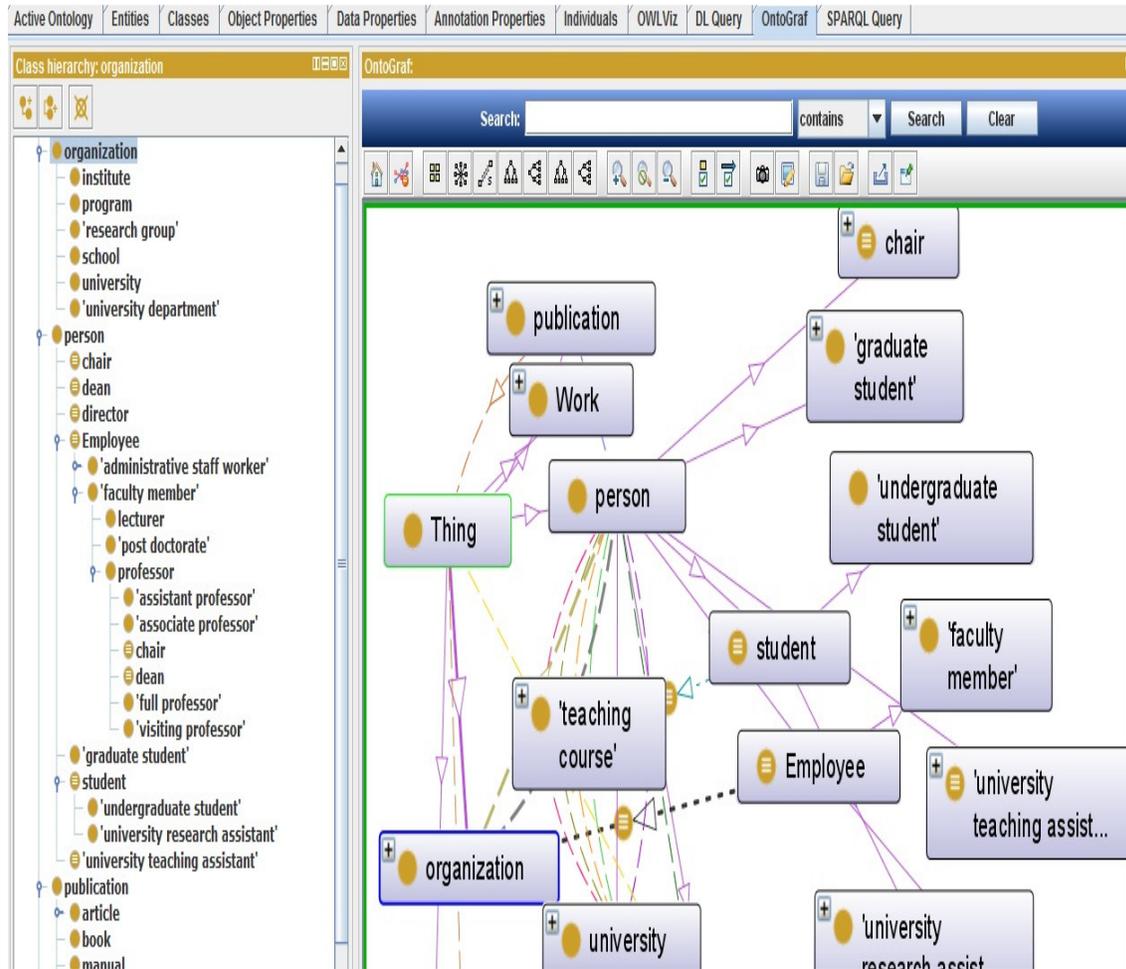


Figure 4. Updated University Ontology after incorporating the extracted relevant conceptual relationships proposed by the framework

4.1 Ontology Evaluation

Ontology evaluation is critically important if the ontologies are to be widely adopted in the Semantic Web and other Semantic aware applications. Evaluation of ontology refers to a process of assessing whether a given ontology represents the conceptual relationships in a given domain or selecting the best ontology among the candidate ontologies. One can assess the quality of the constructed ontology. In this paper we have assessed the quality of semi-automatically developed ontology. A comprehensive survey of existing ontology evaluation techniques is presented in research work [26]. Four types of evaluation approaches were discussed in the above work.

- Gold Standard Evaluation: Comparing the ontology based on a gold standard.
- Application based evaluation: based on the results obtained by using the ontology in an application.
- Data driven evaluation: based on the comparisons done on the source data.
- User evaluation: based on the evaluation done by humans.

Based on the nature of our research, user evaluation is well suited to assess the quality of the extracted Semantics because it is the user (knowledge engineer) who finally decides whether to add a particular conceptual relationship to web site ontology or not, and he/she is the best person to judge the quality of the extracted semantics. The report of the conceptual relationships is presented to ontology developer who decides for updating the ontology. He assesses the quality of generated ontology or enriched ontology.

5. CONCLUSIONS AND FUTURE WORK

Ontology is regarded as a backbone for Semantic Web. Manual acquisition of ontologies is tedious and cumbersome task. Constructing ontology for new domain is time-consuming and costly affair. In this paper we have presented a framework for semi-automatic construction of ontology using knowledge discovery techniques with an aim to reduce the effort required to produce ontology for a new domain.

The main contribution of this research paper is the concept of using both Web Content Mining and Web Usage Mining for semi-automatically developing the ontologies. The main idea behind this concept was to incorporate the Web author's ideas as well as Web users' intentions in the ontology development process. However, reliability of using only Web Usage Mining information was seen as not a viable solution because of the rapid changing nature seen in some of the web sites and hence, identifying users' navigational patterns is difficult and it may not reveal adequate information. Therefore, the proposed methodology extracts concepts and conceptual relationships using Web Content Mining and the information revealed by Web Usage Mining is incorporated to refine the ontologies. The quality of the constructed Ontology is assessed using the User Evaluation method. In future research, we plan to combine Web Structure Mining with the proposed approach in developing ontologies, which could be expected to give the profitable results. Another direction in which, the proposed approach can be extended is, by applying Web Mining techniques to domain specific multiple Web sites to develop domain ontology.

REFERENCES

- [1] Maedche and S. Staab, (2001). "Ontology Learning for the Semantic Web ", IEEE Intelligent Systems, *Special Issue on the Semantic Web*, Vol.16, No.2, pp. 72 - 79.
- [2] G.Stumme , A. Hotho and B. Berendt (2006). "Semantic Web Mining: State of the art and future directions", *Journal of Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 4, No. 2, pp. 124-143.
- [3] Bettina Berendt, Andreas Hotho, and Gerd Stumme (2002). "Towards semantic web mining ", In *International Semantic Web Conference (ISWC, Springer)*, pp. 264-278.

- [4] G.Stumme, B. Berendt and A Hotho, (2004). “Usage Mining for and on the Semantic Web ”, *Data Mining: Next Generation Challenges and Future Directions*, AAAI/MIT Press, pp. 461-480.
- [5] Julia Hoxha, Martin Junghans, Sudhir Agarwal. “Enabling semantic analysis of user browsing patterns in the Web of Data ”, In *Proceedings of the 2nd International Workshop on Usage Analysis and the Web of Data (USEWOD), 21st International World Wide Web Conference (WWW 2012)*, Arxiv abs/1204.2713, Lyon, France, April, 2012.
- [6] T. Gruber (1994).“Towards principles for the design of ontologies used for knowledge sharing ”, *International Journal of Human and Computer Studies*, Vol. 43, pp. 907–928.
- [7] Noy, N. F., Sintek, M., Decker, S., Crubezy, M., Fergerson,R.W., and Musen, M.A. (2001). “Creating Semantic Web Contents with Protege-2000 ”, In *IEEE Intelligent Systems*, Vol. 16, No. 2, pp. 60-71.
- [8] Sure, Y., Erdmann, M., Angele, J., Staab, S., Studer, R., and Wenke, D. (2002). “OntoEdit: Collaborative ontology development for the Semantic Web ”, In *International Semantic Web Conference (ISWC 2002)*, Sardinia, Italy.
- [9] Gomez-Perez, A., Manzano-Macho, D. (2003). “OntoWeb Deliverable 1.5: A Survey of Ontology Learning Methods and Techniques ”, *Universidad Politecnica de Madrid*.
- [10] Shamsfard, M. and Barforoush, A. A. (2003). “The state of the art in ontology learning: A framework for comparison ”, *The Knowledge Engineering Review*, Vol. 18, No.4, pp. 293-316.
- [11] Sabou, M., Wroe, C., Goble, C., and Mishne, G.(2005). “Learning Domain Ontologies for Web ServiceDescriptions: an Experiment in Bioinformatics”, In *Proceedings of the 14th International World Wide Web Conference (WWW2005)*, Chiba, Japan.
- [12] Maedche and S. Staab (2000). “Semi-automatic engineering of ontologies from text”, In *Proceedings of the 12th Internal Conference on Software and Knowledge Engineering*, pp. 231-239.
- [13] Sean P. Igo, E. Desmontils, and Ellen Rilo, (2009). “Corpus-based semantic lexicon induction with web-based collaboration ”, In *NAACL-09 Workshop on Unsupervised and Minimally Supervised Learning of Lexical Semantics*.
- [14] Eduard Hovy, Zornitsa Kozareva, and Ellen Rilo, (2009). “Learning and evaluating the content and structure of a term taxonomy ”, In *Proceedings of AAAI-09 Spring Symposium*.
- [15] D. P. T. Nguyen, Y. Matsuo, and M. Ishizuka,(2007). “Exploiting Syntactic and Semantic Information for Relation Extraction from Wikipedia ”, *IJCAI Workshop on Text-Mining Link-Analysis*.
- [16] Changqing Li and Tok Wang Ling, (2005). “From XML to Semantic Web”, In *10th International Conference on Database Systems for Advanced Applications*, pp. 582-587.
- [17] Mikroyannidis and B. Theodoulidis, (2005). "Web Usage Driven Adaptation of the Semantic Web ", In *Proc. End User Aspects of the Semantic Web Workshop, 2nd European Semantic Web Conference (ESWC 2005)*, Heraklion, Greece, pp. 137-147.
- [18] Jayatilaka A.D.S and Wimalarathne G.D.S.P, (2011). “Knowledge Extraction for Semantic Web Using Web Mining ”, *International conference on Advances in ICT for emerging regions (ICTER 2011)*, pp. 89-94.
- [19] Tatyana Ivanova, (2010). “A Semi-Automatic Ontology Learning Method for E-Learning Resources Terminology Extraction”, *International Conference on ICL*, pp. 1030-1034.
- [20] Marko Brunzel and Myra Spiliopoulou, (2006). “Discovering multi terms and co-hyponymy from XHTML documents with xtream ”, In *Proceedings of PAKDD Workshop on Knowledge Discovery from XML Documents (KDXD 2006)*, LNCS 3915.
- [21] Hazman, M., El-Beltagy, S. R., and Rafea, A. (2009). “Ontology Learning from Domain Specific Web Documents ”, In *International Journal of Metadata, Semantics and Ontologies*, Vol. 4, No. 1-2, pp. 24 – 33.
- [22] R. Agrawal and R. Srikant. (1995). “Mining Sequential Patterns ”, *Proceedings. of the 11th Int’l Conference on Data Engineering (ICDE- 95)*, pp. 3-14.
- [23] Yan, X., Han, J., Afshar, R. (2003). “CloSpan: Mining Closed Sequential patterns in large datasets ”, In *Proceedings of SIAM International Conference on Data Mining*, pp.166–177.
- [24] J. Srivastava, R. Cooley, M. Deshpande and P. Tan, (2000). “Web usage mining: Discovery and applications of usage patterns from Web data ”, *SIGKDD Explorations*, Vol. 1, No. 2, pp. 12-23.
- [25] <http://protege.stanford.edu>

- [26] Brank, J., Grobelnik, M., and Mladenic (2005). A survey of ontology evaluation techniques ", In *Proceedings of the 8th International Multi-Conference Information Society IS-2005*.

AUTHORS

Mr. C. Ramesh is working as Associate professor in Computer Science and Engineering Department at CVR College of Engineering. He is pursuing his PhD from Jawaharlal Nehru Technological University, Hyderabad. He received his B.E in Computer Science and Engineering from Osmania University and M.Tech from Jawaharlal Nehru Technological University, Hyderabad. He has guided 16 M.Tech projects and published 3 papers in International and National journals.



His areas of interest include Databases, Data Mining, Semantic Web Mining, Web Usage Mining and Social Networks. He has 14 years of teaching experience.

K.V. Chalapati Rao is a Professor of Computer Science & Engineering and Dean, Academics at CVR College of Engineering. Prior to joining the CVR, he served Osmania University as a Professor and Head, Department of CSE and Dean of Engineering. After obtaining his PhD, he joined Electronics Corporation of India Limited and worked in various capacities for 16 years, before joining the Osmania University.



He guided number of PhD scholars in areas of Real time systems, Operating Systems, Software Engineer. He has published a number of papers in International/National journals/conferences including IEEE, ACM, Springer and Elsevier.

A.Govardhan is presently a Professor of Computer Science and Engineering, Director at School of Information Technology and Executive Council Member, Jawaharlal Nehru Technological University Hyderabad (JNTUH), India. He served and held several Academic and Administrative positions including Director of Evaluation, Principal, Head of the Department, Chairman and Member of Boards of Studies and Students' Advisor. He received Ph.D. degree in Computer Science and Engineering from Jawaharlal Nehru Technological University in 2003, M.Tech from Jawaharlal Nehru University in 1994 and B.E from Osmania University in 1992.



He is the recipient of 24 International and National Awards. He has guided 48 Ph.D theses, 1 M.Phil, 135 M.Tech projects and he has published 350 research papers at International/National Journals/Conferences including IEEE, ACM, Springer and Elsevier. He is a Member on Editorial Board for 11 International Journals and PC Member and Reviewer for several International/National Conferences. He is a Life Member/Member in several Professional and Service Oriented Bodies. His research of interest includes Databases, Data Warehousing & Mining, Information Retrieval, Computer Networks, Image Processing, Software Engineering and Object Oriented Technologies.