IJRSET November 2015 Volume 2, Issue 12 Pages: 26-31
ISSN 2394-739X
International Journal for Research in Science Engineering and Technology

SEARCH ENGINE OPTIMIZATION USING FUZZY ONTOLOGY GENERATION FRAMEWORK

[1] J. Mala, [2] J.Meena
[1] Assistant Professor/IT, [2] Assistant Professor/ECE
[1] Sri Ramakrishna Institute of Technology, [2] P.S.R Engineering College,
[1] Coimbatore, [2] Sivakasi.

Abstract

The main objective of this project is to optimize the search engine so that it retrieves only the relevant documents requested by the user using FOGA.The system is wholly based on machine learning rather than human interpretation for gathering relevant information. This is a semantic web-based information retrieval system to support scholarly activities performed in the semantic web environment. This system does machine learning when the particular keyword is specified by the user. Rather than retrieval of information or documents based on the occurrence of the particular keywords, relationships are formed between those set of keywords. And the machine learns what is the need of the user and retrieves only those information that are relevant. This system uses the Fuzzy Ontology Generation frAmework (FOGA) would be useful to construct ontology from insecurity data as it can represent insecurity information and construct a concept hierarchy automatically. Through this system the user fatigue in extracting the necessary information from the web is reduced and computers can automatically harness the enormous network of information and services on the web.

Keywords: - Web Mining, Knowledge Discovery, Knowledge Engineering, Ontology design, Semantic Web

1. INTRODUCTION

Text mining is a growing new field that attempts to collect meaningful information from natural language text. It is the process of analyzing text to extract useful information that is useful for particular purposes.Text is unstructured, amorphous, and difficult to deal with algorithmically. Text mining is used to denote any system that analyzes large amount of natural language text and identifies lexical or semantic usage patterns in an attempt to extract useful information.Text mining is also called "nontraditional information retrieval (IR) strategies."The goal of the IR strategies is to minimize the effort required of users to obtain useful information from large computerized text data sources.

A. Ontology

Ontology is a data model that represents a set of ideas within a domain and the relationships between those concepts. It is a conceptualization of a domain into machine-readable, human understandable format consisting of attributes, relationships, entities and axioms. It is used as a standard knowledge representation for the Semantic Web. The conceptualization supported by ontology may not be sufficient to represent uncertainty information which is found in many application domains due to the lack of specific boundaries between concepts of the domains. For example, a document can be relevant or irrelevant to a research field. In addition, keywords are extracted from scientific publications can be used to gather the corresponding research fields. Though, it is unsuitable to treat all keywords equally as some keywords may be more important than others. To tackle this type of problems, one solution is to combine fuzzy logic into ontology to handle uncertainty data. Fuzzy ontology is generated & used in text retrieval and search engines, in which membership values are used to evaluate the similarities between the ideas in a concept hierarchy. Manual generation of fuzzy ontology from a predefined concept hierarchy is a difficult task that often requires expert interpretation. So, automatic generation of concept hierarchy and fuzzy ontology from insecurity data of a domain is highly desirable.

B. Fuzzy Logic

Fuzzy logic can be integrated to ontology to represent uncertainty information. Typically, fuzzy ontology can be used for text retrieval and search engines.

C. Semantic Web

The Semantic Web plans to create a universal medium for information exchange by putting documents with computer processable on World Wide Web in a standardized way. A huge amount of data is interpretable by humans and limited machine support.Now it is almost impossible to retrieve information with a keyword search when the information is spread over several web pages. e.g., the query for Web mining experts in a company, where the clear information are stored and relationships between people and courses are attended on one side, and between courses and the topics are attended on other side. In that case, the use of a rule maintaining that people who attended a course which contains a knowledge about certain topic that might improve the results.

D. FOGA

Fuzzy Ontology Generation framework is an algorithm that automatically generates a fuzzy ontology from uncertainty data based on Formal Concept Analysis (FCA) concept.

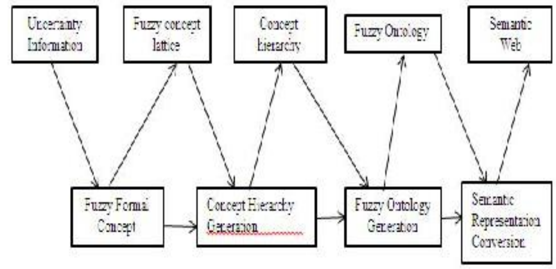


Figure1 FOGA Process

E. FCA

FCA is a proper technique for analysing of data and knowledge presentation. It defines proper contexts to represent relationships between objects and attributes in a domain.

2. OPERATIONS

A. Training:

1. Selection of documents and attributes
2. Cross table creation
3. Refining cross table
4. Lattice formation
5. Conceptual clustering
6. Concept hierarchy formation
7. Ontology generation

B. Online search:

1. Document mapping
2. Web page retrieval

Selection of documents and attributes:

Creating and maintaining the documents which are related to the attributes.The attributes are based on clustering, data mining and fuzzy logic. Each attributes are having documents which cover information about the above mentioned attributes So selection of documents are based on the attributes which I already maintained. Cross table creation & Refining cross table Fuzzy formal context can be denoted as a cross table. The context has three objects representing three documents like D1, D2, D3 etc. Moreover it has three attributes, "Data Mining", "Clustering "and "Fuzzy Logic" representing three research topics. The link between an object and an attribute is denoted by a membership value. Usually, the attributes of a formal concept are considered as the concept description. Hence, the relationships between the object and the concept should be the same of the relationships between the objects and the attributes of the concept. As each relationship between the object and an attribute is represented as a membership value in fuzzy formal context, the crossing of these membership values should be the smallest of these membership values as in fuzzy sets intersection.

Lattice formation:

The fuzzy formal concept in fuzzy set can be considered a special case of a many-valued context Fuzzy Formal Concept Similarity. Though, our fuzzy-based modification of FCA as presented in Fuzzy Formal Context and Fuzzy Formal Concept reserves differently continuous values of objects' memberships are calculated by similarities of the concepts. There are many super concepts and sub concepts are available in a formal context. But, the similarities of a concept to its super concepts and sub concepts are dissimilar. Such information cannot be shown in a traditional concept lattice. Using fuzzy concept lattice with fuzzy set theory to calculate the similarities between a concept and its subconcepts. The fuzzy concept lattice is easier than the L-fuzzy lattice among the number of formal concepts. It can provide additional information related to membership values of objects in each fuzzy formal concept and resemblances of fuzzy formal concepts, which are essential for the construction of concept hierarchy. Concept hierarchy formation Concept Hierarchy Generation groups the fuzzy concept lattice which is generated by Fuzzy Formal Concept to construct a concept hierarchy using conceptual clustering and Hierarchical Relation Generation. Conceptual clustering In concept lattice, the fuzzy concept lattice generated using FFCA is complicated sometimes due to generation of large number of fuzzy formal concepts. Since the formal concepts are generated mathematically, objects which have small changes in terms of attribute values are classified into distinct formal concepts. Such objects should belong to the same idea or concept when they are interpreted by human. Therefore, we cluster formal concepts into conceptual clusters using fuzzy conceptual clustering. Compared to outdated clusters, the conceptual clusters generated have the following properties:

1.Each conceptual cluster is measured as a human interpretable concept in the domain of the fuzzy concept lattice.

2. Each conceptual cluster is a sublattice mined from the fuzzy concept lattice.

3.A formal concept must belong to atleast one conceptual cluster. Conceptual clusters generated from the fuzzy concept lattice with similarity confidence thresholds and the conceptual clustering algorithm. A conceptual cluster can be measured as a set of fuzzy formal concept.

4.Each concept is related with a set of objects and attributes. Each conceptual cluster can also be denoted as sets of objects and attributes.

5.Furthermore, each object in each conceptual cluster should have a membership value.Hence, we define some formal definitions as follows: Fuzzy Representation of Conceptual Cluster, Object Set of a Conceptual Cluster, Attribute Set of a Conceptual Cluster, Membership Value of Objects in Conceptual Cluster. Hierarchical Relation Generation Set of conceptual clusters are generated from the Fuzzy conceptual clustering. To construct a concept hierarchy from the conceptual clusters which is used to find the hierarchy relations from the clusters. We need to define a concept hierarchy as follows: Concept Hierarchy, Subconcept and Superconcept on a Concept Hierarchy.

6.The concept Hierarchy assures that each conceptual cluster has at least one superconcept, except it corresponds to the source node of the concept hierarchy generated. However, we must prove that the relation given in Subconcept and Superconcept on a Concept Hierarchy is a partial order.

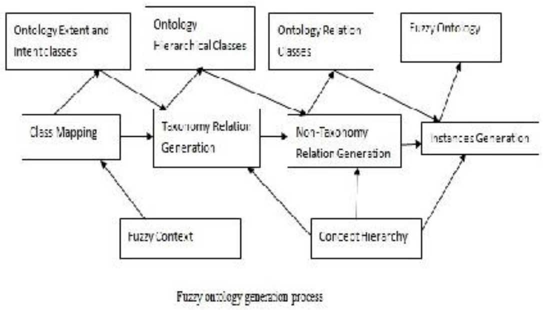


Figure 2Fuzzy Ontology Generation

This step creates fuzzy ontology from a fuzzy context using the concept hierarchy formed by fuzzy conceptual clustering. This is done based on the characteristics of FCA and ontology support formal definitions of concepts. Though, a concept defined in FCA has both additional and objective information, where a concept in ontology highlights on its objective part. To construct the fuzzy ontology, we need to convert both additional and objective information of FCA concepts into the consistent classes and relations of the ontology.

Ontology consists of two concepts or classes. They are document and research area. The attributes of class document are the major things of a scientific publications such as title, author, etc. A Research Area is indicated by a set of suitable keywords. Non taxonomy relation set consists of relations between Document and Research Area indicating that a document can fit to some research areas which consist of several scientific documents. Taxonomy relations set RT define hierarchical relations between research area which may be a super area or subarea of the others. The axiom sets X contains several basic rules that suggest the inversed relations of defined relations.

The attributes and relations of examples of the Ontology are associated with membership values that indicate their degree of uncertainty are discussed earlier. Fuzzy Ontology Generation has some steps. They are as follows, Class mapping, Taxonomy relation generation, Instances generation,Non-Taxonomy relation generation.

3. EXISTING SYSTEM

Currently user keywords are searched and documents are retrieved based on its occurrence in the documents. There is a lack of forming relationships between the set of keywords given by the user.It is inappropriate to treat all documents equally in terms of significance of keyword occurrences. The existing system uses methods like FCA methods like text processing, ontology merging, web-navigation, etc., L-Fuzzy FCA, Ontology Editing Tools.

4. PROPOSED SYSTEM

The proposed system uses FOGA which would be useful to construct ontology from uncertainty data and construct a concept hierarchy automatically. It is a semantic Web-based information retrieval system to support intellectual activities in the Semantic Web.

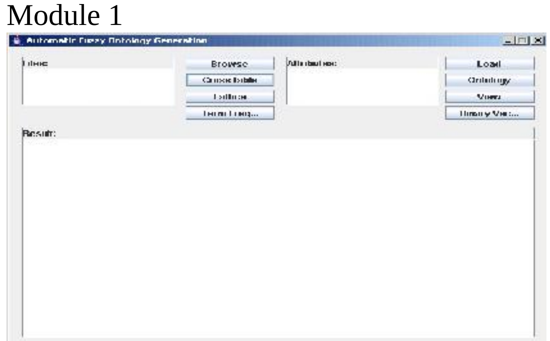
5. EXPERIMENTAL RESULTS

Module Design

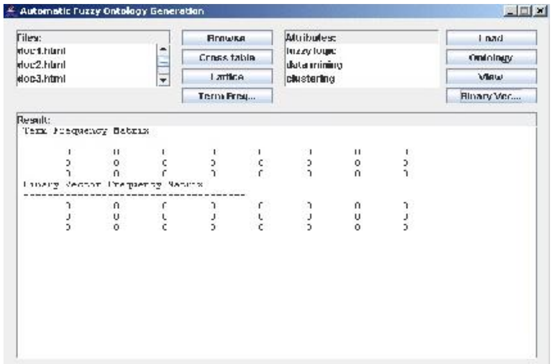
The methods used in the FOGA Framework involves,

1. Data collection and basic GUI design
2. .FFCA implementation
3. Fuzzyconceptual clustering
4. Fuzzy ontology generation
5. Online search

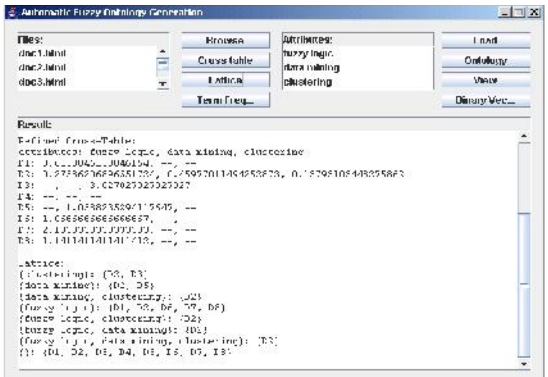
Module 1



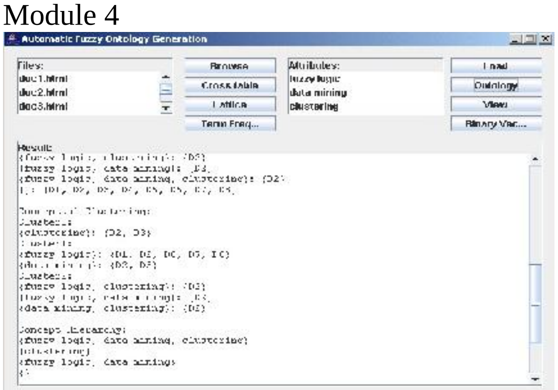
Module 2



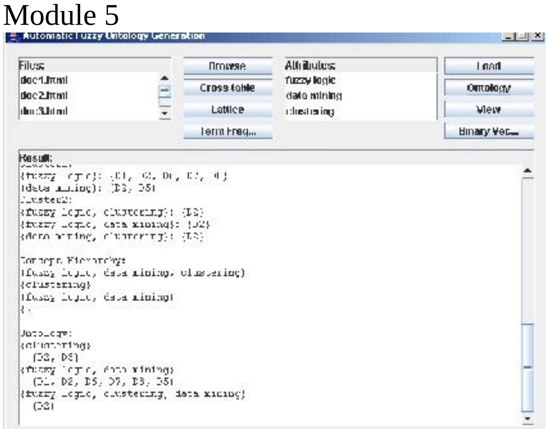
Module 3



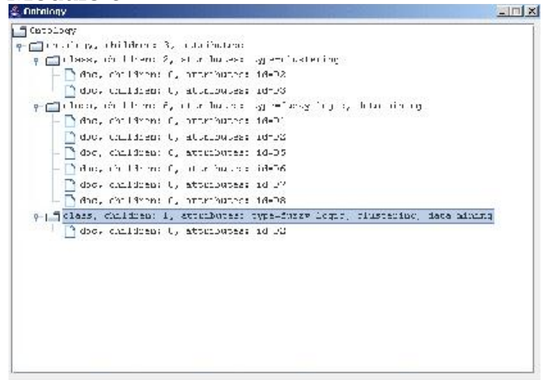
Module 4



Module 5



Module 6



CONCLUSION

In this paper, we have proposed the fuzzy ontology generation framework (FOGA) on uncertainty information. FOGA consists of Fuzzy Ontology Generation, Fuzzy Formal Concept Analysis, Fuzzy Conceptual Clustering and Semantic Representation Conversion. Furthermore, we have also proposed an approximate technique that allows the generated fuzzy ontology to be incrementally equipped with new instances. We have also proposed a technique to add extra attributes in a database to the ontology. The proposed work of FOGA framework would be useful to construct ontology from uncertainty data.

REFERENCES

[1] N. Guarino and P. Giaretta, Ontologies and Knowledge Bases: Towards a terminological Clarification. Toward Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing. Amsterdam: IOSPress, 1995.

[2] T. Berners-Lee, J. Hendler, and O. Lassila, "The Semantic Web," Scientific Am., http://www.sciam.com/2001/0501issue/0501 ber ners-lee.html, 2001.

[3] L.A. Zadeh, "Fuzzy Logic and Approximate Reasoning," Synthese, vol. 30, pp. 407-428, 1975.

[4] M.Z. Islam and L. Brankovic,"A Framework for Privacy Preserving Data Mining," Proc. Australasian Workshop Data Mining and Web Intelligence (DMWI '04), pp. 163-168, 2004.

[5]D.H. Widyantoro and J. Yen, “A Fuzzy User Studies,” Proc. 10th IEEE Int’l Conf. Fuzzy Systems, pp. 1291-1294, 2001.
[6] B. Ganter and R. Wille, Formal Concept Analysis: MathematicalFoundations. Springer, 1999.
[7]W3C,“WebOntologyLanguageOverview, ”http://www.w3.org/TR/owl-features/,2006.
[8]OntoWebdeveliverable1.3,http://www.on towe b.org/, 2006.
[9]A. Gomez-Perez, O. Corcho, and M. Fernandez-Lopez, Ontological Engineering: With Examples from the Areas of Knowledge Management,e-Commerce and the Semantic Web (Advanced Information and Knowledge Processing). Springer, 2004.
[10]B. Bachimont, A. Isaac, and R. Troncy, “Semantic Commitment forDesigning Ontologies: A Proposal,” Proc. Int’l Conf. KnowledgeEng. and Knowledge Management, pp. 114-121, 2002.
[11]D.I. Moldovan and R.C. Girju, “An Interactive Tool for the Rapid Development of Knowledge Bases,” Int’l J. Artificial Intelligence Tools (IJAIT), vol. 10, nos. 1-2, 2001.
[12]A. Maedche and S. Staab, “Ontology Learning for the Semantic Web,” IEEE Intelligent Systems, vol. 16, no. 2, 2001.
[13]R. Navigli, P. Velardi, and A. Gangemi, “Ontology Learning and Its Application to Automated Terminology Translation,” IEEE Intelligent Systems, vol. 18, no. 1, 2003.
[14]F. Xu, D. Kurz, J. Piskorski, and S. Schmeier, “A Domain Adaptive Approach to Automatic Acquisition of Domain Relevant Terms and Their Relations with bootstrapping,” Proc. Third Int’l Conf. Language Resources and Evaluation (LREC 2002), 2002.
[15]L. Khan and F. Luo, “Ontology Construction for Information Selection,” Proc. 14th IEEE Int’l Conf. Tools with Artificial Intelligence, pp. 122-127, 2002.