

An Ontology-based Web-Mining Technique for Supply-Chain Management

Chang Jung University

Chung-Hong Lee* and Hsin-Chang Yang

Abstract:

This paper describes an ongoing research project investigating a platform for discovering knowledge associated with industrial supply-chains using an ontology-based web-mining technique. In addition, we propose a novel approach to establish industrial domain-specific ontologies by means of integrating text mining and XML document techniques. We apply a text-mining approach to automating the construction and maintenance of a concept-hierarchy, in order to establish a XML document database based on the extracted metadata and ontology. The algorithm for knowledge extraction in this work is mainly using a Web-content mining method. Thus, the existing WWW pages can be analyzed to generate a set of metadata to describe their contents and produce ontologies for organizing the XML document database through a text-mining technique, with a neural-net machine learning method for implementation.

1. Introduction

With the explosive growth of information on the World Wide Web, many attempts have been made to take advantage of the Web resource to cope with the difficulties in acquisition of marketing information and knowledge. For knowledge acquisition of supply-chains, a simplified scenario can be established as follows. In the buyer's marketplace, the procurement department needs to define the scope of products to buy. Through the Web, specifications of the products and their prices are normally accessible by the employees of the buying company.

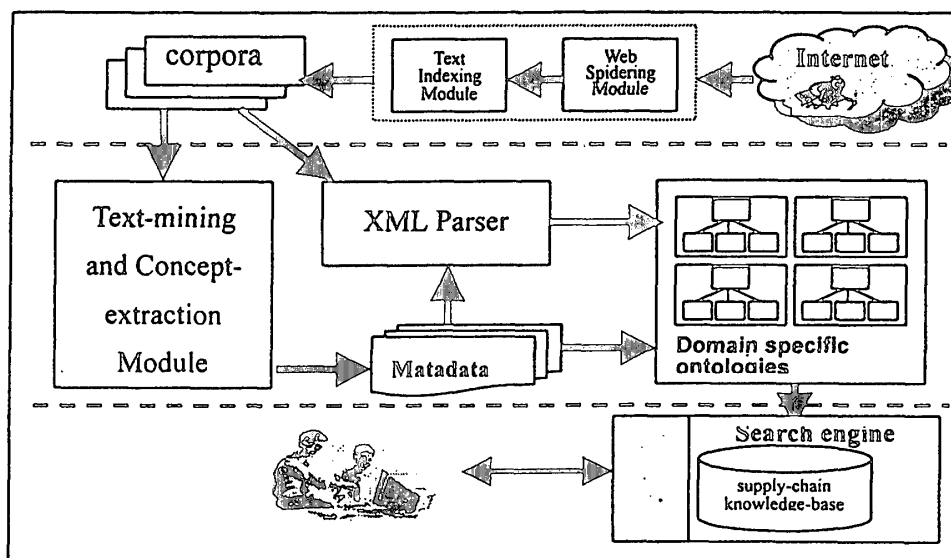


Figure 1 System architecture

The final buyer can very effectively compare the available alternatives in the electronic catalogs, and an organizational purchasing decision can be made according to the collected marketing information at hand. The knowledge components mentioned in the above scenario and their concerns established in this work including:

- **Supply chain of specific products** - with procurement management perspective
- **Competitor's supply chain** - with marketing management perspective

2. System Framework

According to the above scenario, the central part of the system framework includes a knowledge base comprised of collections of supplier profiles and product information. The developed platform acted as an information gateway for guiding the users to search the supplier or product information desired, as shown in Figure 1. It contains user interfaces, a search engine, a text database storing the documents related to supply-chain knowledge, and a web-mining module in which documents associated with supply-chain knowledge can be extracted from original text collections. In this work we also developed a novel approach to establish industrial domain-specific ontologies by means of integrating text mining and XML document techniques. By applying a web mining method, we extract concepts and knowledge from the contents of a huge HTML-document collection using the developed multilingual text-mining algorithm [3,4]. The extracted concepts and knowledge can be used to produce metadata and ontology to describe the information in the original web documents. Therefore, the original web documents can be transformed into XML documents and stored in the XML document database with the developed ontology based on a web-mining algorithm. As such, the existing WWW pages can be analyzed to generate a set of metadata to describe their content and produce an ontology for the XML document database through a text-mining technique, with a neural-net machine learning method [1,2] for implementation.

3. References

- [1] Kohonen, T.: "Self-Organizing Formation of Topologically Correct Feature Maps", *Biological Cybernetics*, Vol. 43, 1982, pp. 59-69.
- [2] Kohonen, T. *Self-Organizing Maps*, Springer Verlag, Berlin, 1995.
- [3] Lee, C.H. and Yang, H.C. (1999): "A Web Text Mining Approach Based on Self-Organizing Map", In Proc. ACM CIKM'99 2nd Workshop on Web Information and Data Management (WIDM'99), Kansas City, Missouri, USA, November 5-6, 1999, pp. 59-62.
- [4] Lee, C.H. and Yang, H.C. (2000): "Towards Multilingual Information Discovery through a SOM based Text Mining Approach.", In Proc. International Workshop on Text and Web Mining, The Sixth Pacific Rim International Conference on Artificial Intelligence (PRICAI 2000), Melbourne, Australia, August 28-September 1, 2000. pp.81-87.