Retrieving product information of collaborative enterprises based on Bayesian network

Junxiang Tu*

College of Mech. Eng., Fuzhou University, Fuzhou, P. R. China,

Fujian Haiyuan Automatic Equipments Co., Ltd., Minhou Fujian, P. R. China

Received 1 March, 2014, www.tsi.lv

Abstract

There exist many differences in nomenclature and descriptions of products and parts in collaborative enterprises, which greatly hinder the retrieval and sharing of web-based product information. In this paper, we present an extended Bayesian network for retrieving and integrating the product information of collaborative enterprises based on product ontology. This approach not only reduces the complexity of existing ontology mapping methods, but also increases the efficiency of product information integration.

Keywords: Product Information Retrieval, Bayesian Network, Ontology, Collaborative Enterprises

1 Introduction

With economic globalization and the development of network technology, the demand for collaborative manufacturing is increasing rapidly. Web-based product information sharing is the cornerstone for the implementation of collaborative manufacturing. However, due to the different enterprise cultures, there are many differences in nomenclature and descriptions of products and parts in different enterprises. The differences greatly hinder retrieving and sharing of enterprise products information and become a bottleneck in collaborative manufacturing development.

The methods of retrieving and integrating web-based product information of different enterprises can be summarized as name-based matching methods [1] and rule-based matching methods [2]. Name-based matching methods have been the most widely used solutions to retrieving product information of collaborative enterprises. Li [3] proposes the use of “term matching” method to explore a database for publishing content on the Web. Wang and Zhang [4] introduce a similar approach that allows querying databases through keywords. However, these methods are difficult to identify synonyms or semantically similar terms, so that there is the problem of poor accuracy with name-based matching methods. Differently, the rule-based matching methods aim to analyse the semantics of keywords based on language grammar rules, which are more targeted and accurate. Greiff [5] proposes an inference network model that allows for structured queries via a rich set of probabilistic operators. Models based on language analysis have been applied to information retrieval [6, 7]. However, the rule-based matching methods have not been widely used in practice because of complexity of algorithms and lack of scalability.

Bayesian network is used to dealing with uncertainty in artificial intelligence [8, 9]. This paper combines it with the product ontology to construct a product information retrieval model. The model utilizes fuzzy matching of ontology elements to automate product information retrieval and integration without building complex matching rules for semantic analysis.

2 Product Ontology and Product Information Retrieval

2.1 PRODUCT ONTOLOGY

For retrieving and integrating product information between collaborative enterprises, the respective product ontology of different enterprises should be built. According to Gruber [10], ontology is the specification of conceptualizations, used to help programs and humans share knowledge. The product ontology can be expressed as the following four-tuple of the form: $O=(C, R, I, A)$, where O represents the ontology to be defined, C is the set of concepts in the ontology, R is the relationship between concepts, I represents the set of instances and A is the axiom asserted.

The function mapping from product ontology O_1 to product ontology O_2 can be defined as follows: $F_m=(\{n_1\},\{a_1\},\{r_1\})\rightarrow(\{n_2\},\{a_2\},\{r_2\})$, where

$\{n_1\}$: the collection of product object names of O_1

$\{a_1\}$: the collection of product object attributes of O_1

* ***Corresponding author*** e-mail: junxtu@163.com

$\{r_1\}$: the collection of relationships between product objects in product ontology O_1

$\{n_2\}$: the collection of product object names of O_2

$\{a_2\}$: the collection of product object attributes of O_2

$\{r_2\}$: the collection of relationships between product objects in product ontology O_2

2.2 Product Information Retrieval and Sharing

As shown in Fig. 1, the process of collaborative enterprises product information retrieval and sharing include three steps: importing target product ontology of collaborative enterprises, ontology mapping and product information integration.

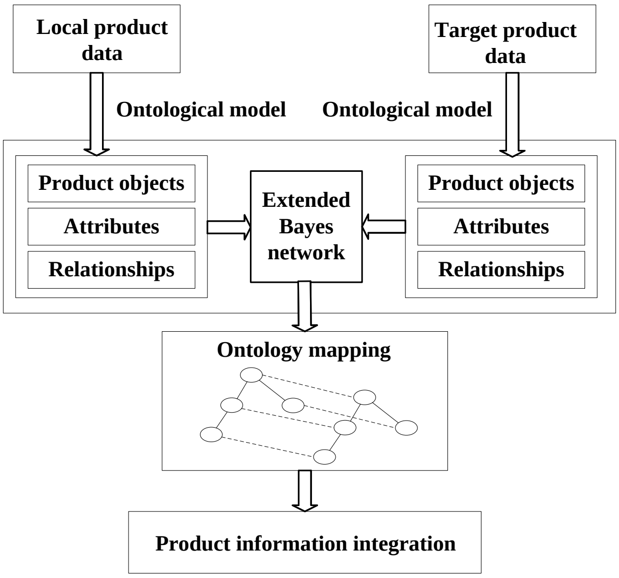

FIGURE 1 Retrieval and sharing of product information

Ontology mapping plays a key role in product information retrieval and integration [11]. Its main aim is to find semantic relationships between local product ontology and target product ontology. In this paper, product ontology mapping is implemented through an extended Bayesian network model, which can automatically discover the elements of target product ontology matching with the ones of the local ontology.

Product information integration is to integrate the product information of collaborative enterprises based on the establishment of ontology mapping after searching web-based product data.

3 Ontology Mapping and Extended Bayesian Network

3.1 BAYESIAN NETWORK

A Bayesian network is a graphical model that encodes probabilistic relationships among variables of interest [12]. The relevance of associated variables is characterized by joint probability that quantifies the interdependencies of the variables [13]. A typical Bayesian network is demonstrated in Fig. 2, where nodes represent random variables, directed edges between nodes represent causal relationships between variables [14].

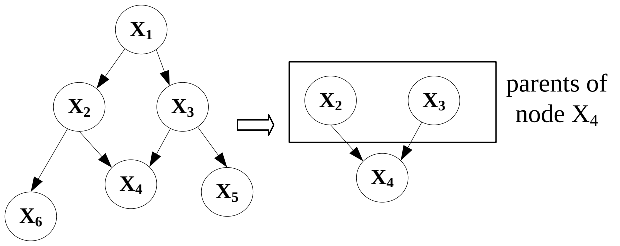

FIGURE 2 Example of a Bayesian network

We use P_{ai} to denote the parents of any node x_i in the network, and $P(x_i \mid P_{ai})$ to denote the conditional probability at the node x_i . For a set of variables $X = (x_1, x_2, ..., x_n)$, the joint probability for X is given by $P(X) = \prod_{i=1}^{n} P(x_i \mid P_{ai})$.

3.2 EXTENDED BAYESIAN NETWORK

To match the elements of target product ontology of collaborative enterprises with the ones of the local product ontology efficiently, we design an extended Bayesian network composed of the probabilistic inference and the semantic inference. The model shown in Fig. 3 is divided into three layers: the local product ontology layer, the root layer and the target product ontology layer.

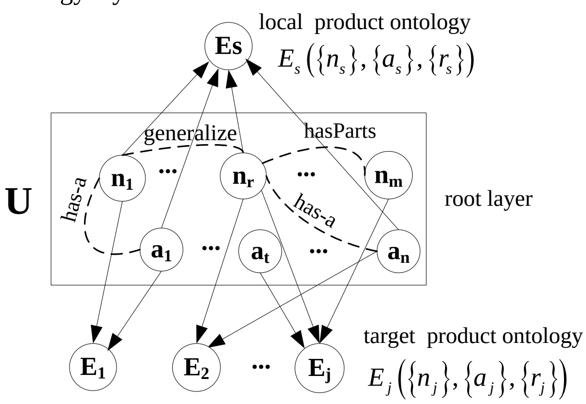

FIGURE 3 Extended Bayesian network

The root layer consists of terms of product objects, product object attributes and their semantic relationships. The nodes $n_r \subset U$ (r = 1,…,m) in the root layer are terms that represent names of product objects in collaborative enterprises. The nodes $a_t \subset U$ (t = 1,…,n) in the root layer represent attributes of product objects. Semantic relationships between these nodes signified by dotted lines includes synonym relationship, generalization relationship, composition relationship, and the relationship between product objects and their attributes (has-a). A generalization relationship is a relationship in which one term (the child) is based on another term (the parent). A composition relationship represents whole-part relationship among terms.

3.3 ONTOLOGY MAPPING PROCESS

Ontology matching can be seen as a process finding the matching elements of different product ontologies, which is the process of calculating the conditional probability $P(E_k \mid E_s)$, as follows:

$$P(E_k \mid E_s) = \eta \sum_u P(E_k \mid u) P(E_s \mid u) P(u),$$

where the set u is used to refer to any of the 2^{m+n} possible states of the root nodes. If $P(E_k \mid E_s)$ is greater than the specified threshold, the element E_k is considered to match with E_s.

First, we define the value of $P(u)$ as a constant ($1/2^{m+n}$) because there is no a priori preference for any set of terms and attributes (subsets of **U**), as follows:

$$P(u) = (\frac{1}{2})^{m+n}.$$

Second, according to vector space model, we can calculate $P(E_k \mid u)$, as follows:

$$P(E_k \mid u) = \frac{\sum_{i=1}^{m+n} w_{ik} \times w_{is}}{\sqrt{\sum_{i=1}^{m+n} w_{ik}^2} \times \sqrt{\sum_{i=1}^{m+n} w_{is}^2}},$$

where w_{ik} is the weight associated with the root node in the target product ontology and w_{is} is the weight associated with the root node in the local product ontology. It should be noted that different weight calculation methods would lead to different mapping strategies.

Third, in order to referring to the state of the variable k_i in u (subsets of root layer **U**), we define an indicator function as follows: $g_u(k_i) = \begin{cases} 1 & if\ k_i \in u \\ 0 & otherwise \end{cases}$.

Given the definition of $g_u(k_i)$, the probability of $P(E_s \mid u)$ is now defined as:

$$P(E_s \mid u) = \begin{cases} 1 & if\ \forall n_r, \forall a_t,\ g_{\{n_s\}}(n_r) = g_u(n_r) \\ & and\ g_{\{a_s\}}(a_t) = g_u(a_t) \\ 0 & otherwise \end{cases}.$$

4 Experiments

Fig. 4 shows the production structure tree of a drilling machine produced by a machine tool enterprise. The drilling machine is mainly composed of drill head, gearbox, electric motor, column and base unit. Most of parts of the product are produced by the enterprise itself, some parts such as electric box are produced by collaborative enterprises.

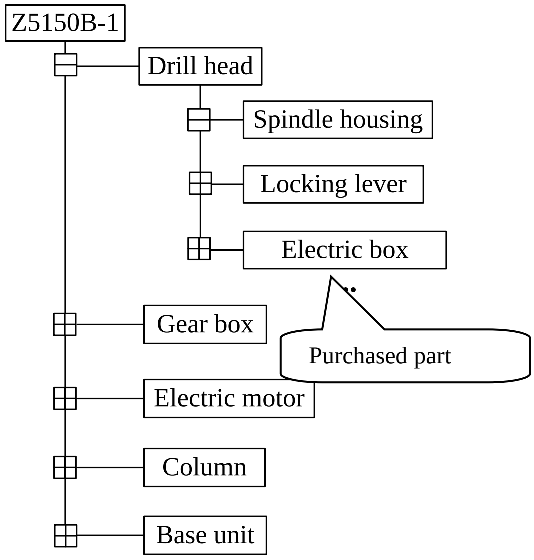

FIGURE 4 Production structure tree of a drilling machine

After importing target product ontology of a collaborative enterprise, we applied our extended Bayesian network aforementioned to ontology mapping. Table 1 lists matching elements of two product ontologies. The last column of the table indicates the credibility of the matching.

TABLE 1 Matching elements of two product ontologies

Element of local product ontology	Element of target product ontology	Credibility
Locking lever	Locking handle	1.00
Front cover	Front lid	0.91
Electric Box	Electric control cabinet	1.00
Engine	Electric motor	0.99
Ball bearing	Bearing	0.86
Table	Workbench	0.80
Gear box	Main spindle box	0.95
Base plate	Base	0.83
Spacer	Bushing	0.88

The adjusted product structure tree is shown in Fig. 5 according to the ontology mapping, where "Electric Box" and "Electric control cabinet" has established a matching relationship.

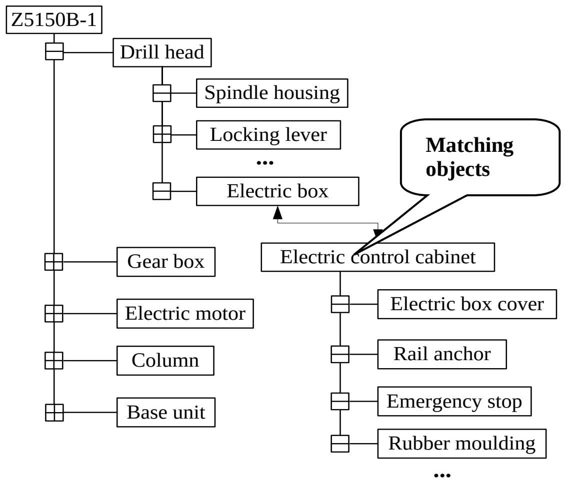

FIGURE 5 Adjusted product structure tree of the drilling machine

5 Conclusions

In this paper, we have proposed an extended Bayesian network model for collaborative enterprises product information retrieval and integration, which can efficiently deal with the problem of semantic differences of product descriptions in different enterprises. The model provides a novel approach to integrate the product information of collaborative enterprises based on ontology fuzzy matching without building complex grammar rules.

As future work, the research on the evaluation of the approach for product information retrieval remains necessary, which helps to further improve the proposed method.

References

[1] Calado P, Ribeiro-Neto B, Ziviani N, Moura E, Silva I 2003 Local versus global link information in the web *ACM Transactions on Information Systems* **21**(1) 42-63

[2] Sudarsan R, Fenves S J, Sriram R D, Wang F 2005 A product information modeling framework for product lifecycle management *Computer Aided Design* **37**(13) 1399-411

[3] Li W, Zhao T, Wang X 2010 Context-sensitive query expansion *Journal of Computer Research and Development* **47**(2) 300-4

[4] Wang S, Zhang K L 2005 Searching databases with keywords *Journal of Computer Science and Technology* **20**(1) 52-62

[5] Greiff W R, Croft W B, Turtle H 1999 PIC matrices: A computationally tractable class of probabilistic query operators *ACM Transactions on Information Systems* **17**(4) 367-405

[6] Revuelta-Martinez A, Rodriguez L, Garcia-Varea I, Montero F 2013 Multimodal interaction for information retrieval using natural language *Computer Standards & Interfaces* **35**(5) 428-41

[7] Kolomiyets O, Moens M F 2011 A survey on question answering technology from an information retrieval perspective *Information Sciences* **181**(24) 5412-34

[8] Boughanem M, Brini A, Dubois D 2009 Possibilistic networks for information retrieval *International Journal of Approximate Reasoning* **50**(7) 957-68

[9] Liao W, Ji Q 2009 Learning Bayesian network parameters under incomplete data with domain knowledge *Pattern Recognition* **42**(11) 3046-56

[10] Gruber T R 1995 Toward principles for the design of ontologies used for knowledge sharing *IEEE Transactions on Power Electronics* **43**(5/6) 907-28

[11] Panetto H, Dassisti M, Tursi A 2012 ONTO-PDM: Product-driven ontology for product data Management interoperability within manufacturing process environment *Advanced Engineering Informatics* **26**(2) 334-48

[12] Heckerman D, Mamdani A, Wellman P M 1995 Real-world applications of Bayesian networks *Communications of the ACM* **38**(3) 24-6

[13] Dogan I 2012 Analysis of facility location model using Bayesian networks *Expert Systems with Applications* **39**(1) 1092-104

[14] Hemmecke R, Lindner S, Studeny M 2012 Characteristic imsets for learning Bayesian network structure *Characteristic imsets for learning Bayesian network structure* **53**(9) 1336-49

Authors

Junxiang Tu, born in November 12, 1971, Jiangxi, China

Current position, grades: Doctor of Mechanical Engineering, lecturer in Fuzhou University
University studies: Mechanical Engineering in Huazhong University of Science & Technology
Scientific interest: Computer integrated manufacturing system
Publications: 1 Patent, 11 Papers
Experience: Lecturer of Fuzhou University, China, 2010-Present; Ph.D. Mechanical Engineering, Huazhong University of Science & Technology, China, 2010; Product engineer in Diamond Company, Guangzhou, China, 1997-2003; Assistant Engineer in Nanchang Zerowatt Electric Appliance Co., Nanchang, China, 1993-1997; B.S. Chemical and Machinery, Nanchang University, China, 1993