

## Comparison Study on Ontology Based Text Mining Method and Proposed the Solution for Selection of University Question Paper

Ms. Amruta Surana ,Mr.Shyam.Gupta

**Abstract**—University question Paper selection is an important task for Universities. When a large number of question papers are received, it is common to group them according to their similarities in each discipline. The grouped papers are then assigned to the appropriate experts for peer review. Current methods for grouping papers are based on manual matching of similar discipline areas and/or keywords. However, the exact discipline areas of the papers cannot often be accurately designated by the applicants due to their subjective views and possible misinterpretation. Text-mining methods have been proposed to solve the problem by automatically classifying text documents, mainly in English. However, these methods have limitations when dealing with non-English language texts, e.g., the language which is not having delimiters (Chinese) Papers. This paper presents a novel ontology-based text-mining approach to cluster question Papers based on their similarities in each areas. The method is efficient and effective for clustering question Papers with both English and the language which is not having delimiters (Chinese) texts.

**Index Terms**—Clustering analysis, decision support systems, ontology, question Paper selection, text mining.

### I. INTRODUCTION

Selection of Question Paper is an important and recurring activity in many Universities. It is a challenging multiprocess task that begins with a call for papers (CFP) by a University. The CFP is distributed to relevant communities such as universities or research institutions. The questions Paper are submitted to the University and then are assigned to experts for peer review. The review results are collected, and the proposals are then ranked based on the aggregation of the experts' review results. Fig. 1 shows the processes of question paper selection at the University, i.e., CFP, paper submission, proposal grouping, proposal assignment to experts, peer review, aggregation of review results, panel evaluation, and final awarding decision [1]. These processes are very similar in other Universities. In the University, the number of question paper received has more than doubled in the past four years, with over 110 paper submitted in one deadline in March 2010. Four to five reviewers are assigned to review each paper so as to assure accurate and reliable opinions on question papers. To deal with the large volume, it is necessary to group papers according to their similarities in each discipline and then to assign the Question Papers groups to relevant reviewers. Departments are classified according to areas, including mathematical and physical sciences, chemical sciences, life sciences, earth sciences, engineering and material sciences, information sciences, and management sciences.

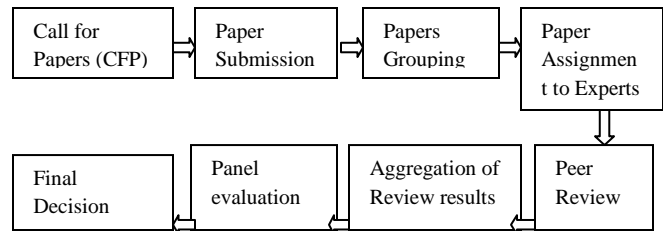


Fig. 1. Question Paper election processes in the University

The University is responsible for the selection tasks, and it dedicates the tasks to divisions or programs. Chairman then group the papers and assign them to external reviewers for evaluation and commentary. However, they may not have adequate knowledge in all disciplines, and contents of many Question Paper were not fully understood when the Papers were grouped. Therefore, there was an urgent need for an effective and feasible approach to group the submitted Papers with computer supports. An ontology-based text-mining approach is proposed to solve the problem.

The remainder of this paper is organized as follows. Section II reviews the literature on Question Paper selection and grouping of papers. The proposed method is described in Section III. Section IV validates and evaluates the method, and then discusses the potential application in the University. Finally, Section V provides the conclusion, and it points to future work.

### II. LITERATURE REVIEW

Selection of question Paper is an important topic in University management. Previous research deals with specific topics, and several formal methods and models are available for this purpose. For example, Chen and Gora [2] Proposed a fuzzy-logic-based model as a decision tool for project selection. Henriksen and Traynor [3] presented a scoring tool for project evaluation and selection. Ghasemzadeh and Archer [4] offered a decision support approach to project portfolio selection. Methods have been developed to group proposals for peer review tasks. For example, Hettich and Pazzani [5] proposed a text-mining approach to group proposals, identify reviewers, and assign reviewers to proposals. Current methods group proposals according to keywords. Unfortunately, proposals with similar research areas might be placed in wrong groups due to the following reasons: first, keywords are incomplete information about the full content of the proposals. Second, keywords are provided by applicants who may have subjective views and misconceptions, and keywords are only a partial representation of the research proposals. Third, manual grouping is usually conducted by division managers or program directors in funding agencies. They may have different understanding about the research disciplines and may not have adequate knowledge to assign proposals into the right groups. Text-mining methods (TMMs) [6], [7] have been designed to group proposals based on understating the English text, but they have limitations when dealing with other language texts, e.g., in the language which is not having delimiters (Chinese).

This paper presents a hybrid method for grouping Paper for Question Paper selection. It uses text-mining, multilingual ontology, optimization, and statistical analysis techniques to cluster question paper based on their similarities.

### III. ONTOLOGY BASED TEXT MINING FOR UNIVERSITY PAPER SELECTION

In the University, after question Paper are submitted, the next important task is to group papers and assign them to reviewers. The papers in each group should have similar characteristics. For instance, if the proposals in a group fall into the same primary discipline (e.g., computer Engineering) and the number of papers are small, manual grouping based on keywords listed in paper can be used. However, if the number of Papers is large, it is very difficult to group Papers manually.

Although there are several text-mining approaches that can be used to cluster and classify Papers [20]–[27], they are developed with a Focus on English text. TMMs which deal with English are not effective in processing the language which is not having delimiters (the language which is not having delimiters (Chinese) text) [28]. For example, (Chinese) text consists of strings of the language which is not having delimiters (Chinese) characters, while English text uses words. Also, Chinese text has no delimiters to mark word boundaries, while English text uses a space as word delimiter. Several methods were proposed to deal with the language which is not having delimiters (Chinese) text [29]–[32], but they are not efficient or sufficiently robust to process research proposals.

To solve the aforementioned problems, an ontology-based TMM (OTMM) is proposed. Ontology is a knowledge repository in which concepts and terms are defined as well as relationships between these concepts [38]–[41]. It consists of a set of concepts, axioms, and relationships that describe a domain of interests and represents an agreed-upon conceptualization of the domain's "real-world" setting. Implicit knowledge for humans is made explicit for computers by ontology [42]–[44]. Thus, ontology can automate information processing and can facilitate text mining in a specific domain (such as question paper selection). The proposed OTMM is used together with statistical method and optimization models and consists of four phases, as shown in Fig. 2. First, a research ontology containing the Question Papers funded in latest five years is constructed according to keywords, and it is updated annually (phase 1). Then, new research proposals are classified according to discipline areas using a sorting algorithm (phase2). Next, with reference to the ontology, the new papers in each discipline are clustered using a self-organized mapping (SOM) algorithm (phase 3). Finally, (phase 4) if the number of proposals in each cluster is still very large, they will be further decomposed into subgroups

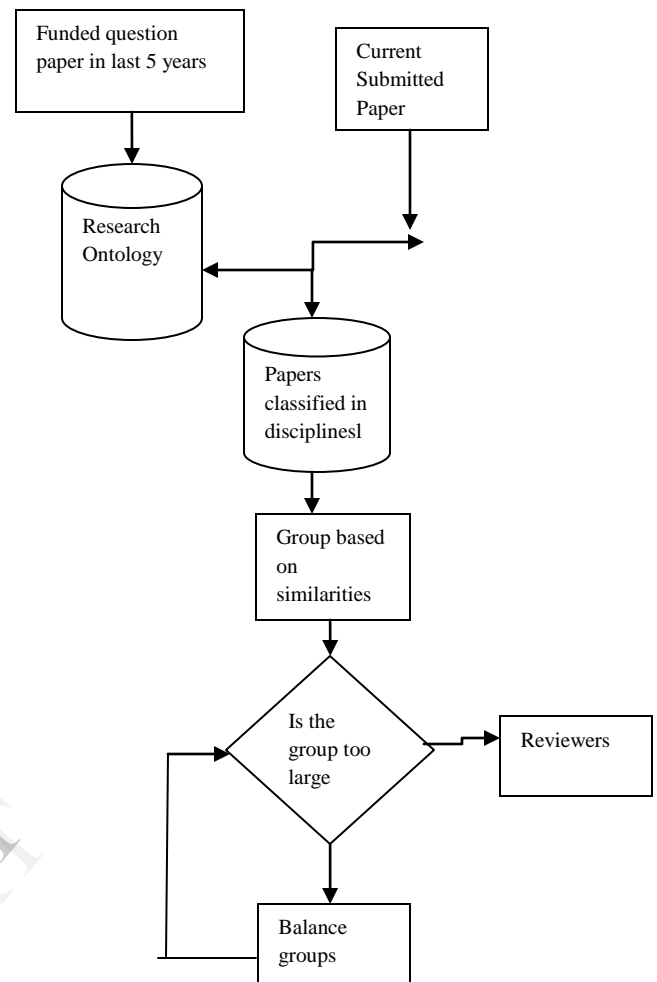


Fig. 2. Process of the proposed OTMM.

#### A. Phase 1: Constructing a question paper Ontology

University maintains a directory of discipline areas that form a tree structure. As domain ontology [41], question paper ontology is a public concept set of the question paper management domain. The question paper of different disciplines can be clearly expressed by ontology. Suppose that there are  $K$  discipline areas, and  $A_k$  denotes discipline area  $k(k = 1, 2, \dots, K)$ . Research ontology can be constructed in the following three steps to represent the topics of the disciplines.

Step 1) Creating the question paper of the discipline $_k$ , ( $k = 1, 2, \dots, K$ ). The keywords of the question paper in each year are collected, and their frequencies are counted (shown in Fig. 3). The keywords and their frequencies are denoted by the feature set

$$(No_k, ID_k, year, \{keyword_1, frequency_1, keyword_2, frequency_2, \dots, keyword_k, frequency_k\}),$$

where  $No_k$  is the sequence number of the  $k$ th record and  $ID_k$  is the corresponding discipline code. For instance, if discipline  $A_k$  has two keywords in 2007 (i.e., "Operating system" and "Data structure") and the total number of counts for them are 30 and 50, respectively, the discipline can be denoted by  $(No_k, ID_k, 2007, \{(Operating\ system, 30), (Data\ structure, 50)\})$ . In this way, a feature set of each discipline can be created. The keyword frequency in the feature set is the sum of the same keywords that appeared in this discipline during the most recent five years (shown in Fig. 4),

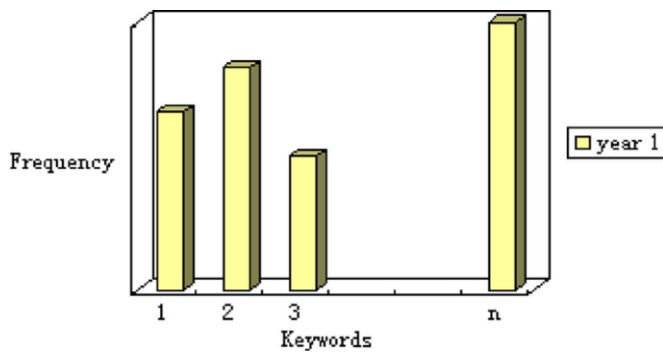


Fig. 3. Keywords of  $A_k$  in a year.

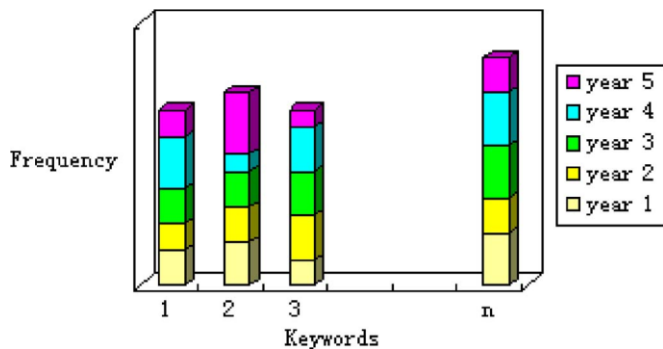


Fig. 4. Feature set of  $A_k$ .

Step 2) *Constructing the question paper ontology*. First, the research ontology is categorized according to scientific research areas introduced in the background. It is then developed on the Basis of several specific areas. Next, it is further divided into some narrower discipline areas. Finally, it leads to the topics in terms of the feature set of disciplines created in step 1. The question paper ontology is constructed, and its rough structure is shown in Fig. 5. Therefore, the research ontology allows more complex relationship between concepts besides the basic tree-like structure. Also, to deal with proposals with both English and Chinese text, it is designed as a multilingual ontology [45], which can process and share knowledge represented in multiple languages.

Step 3) *Updating the research ontology*. Once the Question paper funding is completed each year, the ontology is updated according to University policy and the change of the feature set. Using the research ontology, the submitted question papers can be classified into disciplines correctly, and question paper in one discipline can be clustered effectively and efficiently. The details will be given in the following two sections.

### B. Phase 2: Classifying question papers Into Disciplines

Papers are classified by the discipline areas to which they belong. A simple sorting algorithm is used next for papers' classification. This is done using the ontology as follows.

Suppose that there are  $K$  discipline areas, and  $A_k$  denotes area  $k$  ( $k = 1, 2, \dots, K$ ).  $P_i$  denotes papers  $i$  ( $i = 1, 2, \dots, I$ ), and  $S_k$  represents the set of papers which belongs to area  $k$ . Then, a sorting algorithm can be implemented to classify papers to their discipline areas, as shown in Table I.

### C. Phase 3: Clustering Question Paper Based on Similarities Using Text Mining

After the question papers are classified by the discipline areas, the papers in each discipline are clustered using the text-mining technique [18], [19]. The main clustering process consists of five steps, as shown in Fig. 6: text document collection, text document preprocessing, text document encoding, vector dimension reduction, and text vector clustering.

The details of each step are as follows.

Step 1) *Text document collection*. After the question paper are classified according to the discipline areas, the paper  $l$  documents in each discipline  $A_k$  ( $k = 1, 2, \dots, K$ ) are collected for text document preprocessing.

Step 2) *Text document preprocessing*. The contents are usually nonstructural. Because the texts of the question paper consist of the language which is not having delimiters (the language which is not having delimiters (Chinese)) characters which are difficult to segment, the question paper ontology is used to analyze, extract, and identify the keywords in the full text of the proposals.

Step 3) *Text document encoding*. After text documents are segmented, they are converted into a *feature vector* representation:  $V = (v_1, v_2, \dots, v_M)$ , where  $M$  is the number of features selected and  $v_i$  ( $i = 1, 2, \dots, M$ ) is the TF-IDF encoding [18] of the keyword  $w_i$ . TF-IDF encoding describes a weighted method based on inverse document frequency (IDF) combined with the term frequency (TF) to produce the feature  $v$ , such that to produce the feature  $v$ , such that  $v_i = tf_i * \log(N/df_i)$ , where  $N$  is the total number of proposals in the discipline,  $tf_i$  is the term frequency of the feature word  $w_i$ , and  $df_i$  is the number of proposals containing the word  $w_i$ . Thus, research proposals can be represented by corresponding feature vectors.

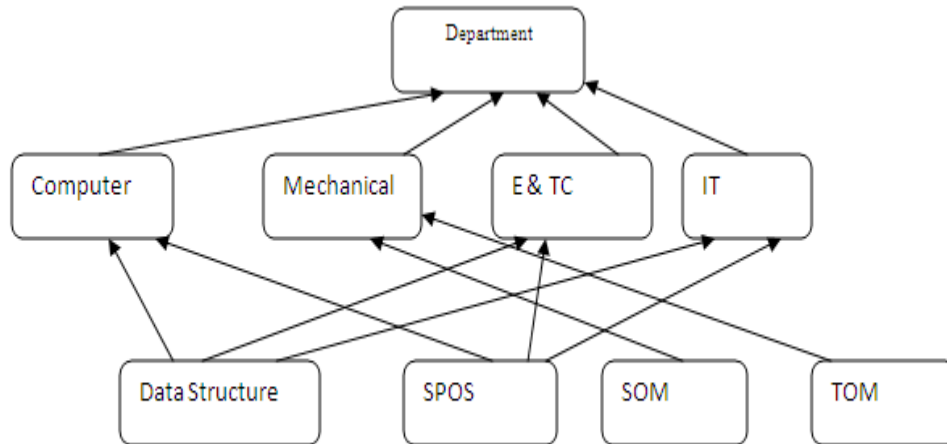


Fig. 5. Structure of the research ontology.

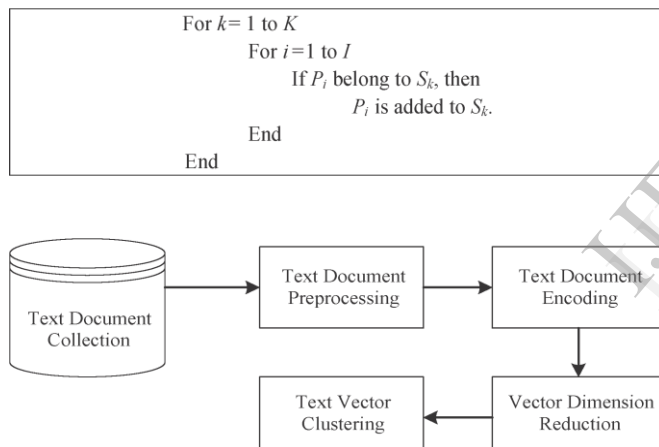
TABLE I  
SUMMARY OF THE SORTING ALGORITHM

Fig. 6. Main process of text mining

Step 4) *Vector dimension reduction*. The dimension of feature vectors is often too large; thus, it is necessary to reduce the vectors' size by automatically selecting a subset containing the most important keywords in terms of frequency. Latent semantic indexing (LSI) is used to solve the problem [18]. It not only reduces the dimensions of the feature vectors effectively but also creates the semantic relations among the keywords. LSI is a technique for substituting the original data vectors with shorter vectors in which the semantic information is preserved.

To reduce the dimensions of the document vectors without losing useful information in a proposal, a term-by-document matrix is formed, where there is one column that corresponds to the term frequency of a document. Furthermore, the term-by-document matrix is decomposed into a set of eigenvectors using singular-value decomposition. The eigenvectors that have the least impacts on the matrix are then discarded. Thus, the document vector formed from the term of the remaining eigenvectors has a very small dimension and retains almost all of the relevant original features.

Step 5) *Text vector clustering*. This step uses an SOM algorithm to cluster the feature vectors based on similarities of research areas. The SOM algorithm is a typical unsupervised learning neural network model that clusters input data with similarities. Details of the SOM algorithm [33], [34] can be summarized as shown in Table II.

#### D. Phase 4: Balancing Research Proposals and Regrouping Them by Considering Applicants' Characteristics

In this phase, when the number of proposals in one cluster is still very large (e.g., more than 20), the applicants' characteristics (e.g., affiliated universities) are considered. As mentioned in Sun *et al.* [15] and Fan *et al.* [35], the proposal group composition should be diverse. In the past, reviewers sometimes handled proposals improperly, having poor group composition (e.g., the same affiliation in a specific proposal group). Reviewers may feel confused and uncomfortable when evaluating proposals that may have poor group composition, so it is advisable that the applicants' characteristics in each proposal group should be as diverse as much as possible. Furthermore, the group size in each group should be similar. This is done as follows.

TABLE II  
SUMMARY OF THE SOM ALGORITHM

Step 1. Initialize network weight vectors  $w_i$ , initialize learning rate parameter  $\mu$ , define topological neighborhood function and initialize parameter  $N_q$ , set  $k = 0$ .

Step 2. Check stopping condition. If false, continue; If true, stop.

Step 3. For each training vector  $x$ , perform Steps 4 to 7.

Step 4. Compute the best match of a weight vector with the input

$$q(x) = \max \text{sim}(x, w_i) \quad \forall i$$

where  $\text{sim}$  can be calculated as cosine value of the angle between vector.

Step 5. For all the units in the specified neighborhood  $i \in N_q(k)$ , where  $q$  is the winning neuron, update the weight vectors according to:

$$w_i(k+1) = \begin{cases} w_i(k) + \mu(k)[x(k) - w_i(k)] & i \in N_q(k) \\ w_i(k) & i \notin N_q(k) \end{cases}$$

where  $0 < \mu(k) < 1$  (the learning rate parameter).

Step 6. Adjust the learning rate parameter.

Step 7. Appropriately reduce the topological neighborhood  $N_q(k)$ .

Step 8. Set  $k \rightarrow k+1$ ; then go to Step 2.

TABLE III  
SUMMARY OF THE GA ALGORITHM

Step 1. Initialize population and parameters, set  $k = 0$ .

Step 2. Check stopping condition. If false, continue; If true, stop.

Step 3. For one generation, perform Steps 4 to 6.

Step 4. Breed new offspring through crossover and mutation (genetic operations).

Step 5. Evaluate the fitness values of parents and offspring.

Step 6. Select best-ranking offspring to populate and replace worst-ranking parents to form a new generation.

Step 7. Set  $k \rightarrow k+1$ ; then go to step 2.

We first define an attribute set  $A$  that includes the applicant's characteristics. Then, the following optimization model is formed:

$$\max \sum_{g=1}^G \sum_{i=1}^{M-1} \sum_{j>i}^M d_{ij} x_{ig} x_{jg} - \sum_{g=1}^{G-1} \sum_{h>g}^G \left| \sum_{i=1}^M x_{ig} - \sum_{i=1}^M x_{ih} \right|$$

s.t.  $\sum_{g=1}^G x_{ig} = 1, \text{ for } i = 1, 2, \dots, M,$

$x_{ig} \in \{0, 1\}, \text{ for } i = 1, 2, \dots, M \text{ and for } g = 1, 2, \dots, G.$

where

- $x_{ig}$  decision variable. When a research proposal  $i$  is allocated to group  $g$ ,  $x_{ig} = 1$ ; otherwise,  $x_{ig} = 0$ ;
- $d_{ij}$  number of different values in the attribute set  $A$  between research proposals  $i$  and  $j$ ;
- $M$  number of research proposals in a cluster;
- $G$  number of proposal groups.

This may be a very complex optimization problem, and one solution method that could be used is genetic algorithm [35]. Details of the GA algorithm [36] (Fan *et al.* 2008) that are applied in our case are summarized as shown in Table III.

#### IV Comparison Study Of Text Mining Methods

- 1] LSA:-Latent Semantic analysis
- 2]PLSA:-Probabilistic Latent Semantic analysis
- 3]LDA:-Latent Dirichlet allocation
- 4]CTM:-Correlated Topic Model

TABLE IV

#### Comparison of four Text Mining Methods

Methods	Applications	Comments and Performance
LSA	1]Automatic Generation 2]Spam Filtering 3]Topic detection	essay LSA>PLSA, LSA>PLSA,LDA VSM>LSA
PLSA	1]Automatic Generation 2]Image Retrieval	essay LSA>PLSA High Level of visual features.
LDA	1]Automatic Generation 2]Experts identification	essay LSA,PLSA>LDA Experts for R&D
CTM	1]Query classification 2]Topic Detection	



TABLE IV

Char and Limitation of four Text Mining Methods

Model	characteristics	Limitations
LSA	1]Reduces Dimensionality of TF-IDF using singular value Decomposition 2]Captures Synonyms of word	1]Difficult to determine the no of topics 2]Difficult to label the topics in some cases
PLSA	1]PLSA partially handles Polysemy	1] No Probabilistic Model at the level of documentation
LDA	1] Handles Long length documents	1] Incapable to model relations among topics
CTM	1]Allows the occurrences of words in other topic	1] Requires complex computations.

## CONCLUSION

This paper has presented an OTMM for grouping of research proposals. A research ontology is constructed to categorize the concept terms in different discipline areas and to form relationships among them. It facilitates text-mining and optimization techniques to cluster research proposals based on their similarities and then to balance them according to the applicants' characteristics. The experimental results at the NSFC showed that the proposed method improved the similarity in proposal groups, as well as took into consideration the applicants' characteristics (e.g., distributing proposals equally according to the applicants' affiliations). Also, the proposed method promotes the efficiency in the proposal grouping process.

The proposed method can be used to expedite and improve the proposal grouping process in the NSFC and elsewhere. It uses the data collected from a research social network (ScholarMate; <http://scholarmate.com>) and extends the functions of the Internet-based Science Information System (<https://isis.nsf.gov.cn>). It also provides a formal procedure that enables similar proposals to be grouped together in a professional and ethical manner. The proposed

method can also be used in other government research funding agencies that face information overload problems.

Future work is needed to cluster external reviewers based on their research areas and to assign grouped research proposals to reviewers systematically. Also, there is a need to empirically compare the results of manual classification to text-mining classification. Finally, the method can be expanded to help in finding a better match between proposals and reviewers.

## ACKNOWLEDGMENT

The authors would like to thank the editors and the anonymous reviewers for their valuable comments and suggestions which have helped immensely in improving the quality of this paper.

## REFERENCES

- [1] Q. Tian, J. Ma, and O. Liu, "A hybrid knowledge and model system for R&D project selection," *Expert Syst. Appl.*, vol. 23, no. 3, pp. 265–271, Oct. 2002.
- [2] K. Chen and N. Gorla, "Information system project selection using fuzzy logic," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 28, no. 6, pp. 849–855, Nov. 1998.
- [3] A. D. Henriksen and A. J. Traynor, "A practical R&D project-selection scoring tool," *IEEE Trans. Eng. Manag.*, vol. 46, no. 2, pp. 158–170, May 1999.
- [4] F. Ghasemzadeh and N. P. Archer, "Project portfolio selection through decision support," *Decis. Support Syst.*, vol. 29, no. 1, pp. 73–88, Jul. 2000.
- [5] L. L. Machacha and P. Bhattacharya, "A fuzzy-logic-based approach to project selection," *IEEE Trans. Eng. Manag.*, vol. 47, no. 1, pp. 65–73, Feb. 2000.
- [6] J. Butler, D. J. Morrice, and P. W. Mullarkey, "A multiple attribute utility theory approach to ranking and selection," *Manage. Sci.*, vol. 47, no. 6, pp. 800–816, Jun. 2001.
- [7] C. H. Loch and S. Kavadias, "Dynamic portfolio selection of NPD programs using marginal returns," *Manage. Sci.*, vol. 48, no. 10, pp. 1227–1241, Oct. 2002.
- [8] L. M. Meade and A. Presley, "R&D project selection using the analytic network process," *IEEE Trans. Eng. Manag.*, vol. 49, no. 1, pp. 59–66, Feb. 2002.
- [9] M. A. Greiner, J. W. Fowler, D. L. Shunk, W. M. Carlyle, and R. T. Mcnett, "A hybrid approach using the analytic hierarchy process and integer programming to screen weapon systems projects," *IEEE Trans. Eng. Manag.*, vol. 50, no. 2, pp. 192–203, May 2003.
- [10] Q. Tian, J. Ma, J. Liang, R. Kowk, O. Liu, and Q. Zhang, "An organizational decision support system for effective R&D project selection," *Decis. Support Syst.*, vol. 39, no. 3, pp. 403–413, May 2005.
- [11] W. D. Cook, B. Golany, M. Kress, M. Penn, and T. Raviv, "Optimal allocation of proposals to reviewers to facilitate effective ranking," *Manage. Sci.*, vol. 51, no. 4, pp. 655–661, Apr. 2005.
- [12] A. Arya and B. Mittendorf, "Project assignment when budget padding taints resource allocation," *Manage. Sci.*, vol. 52, no. 9, pp. 1345–1358, Sep. 2006.
- [13] C. Choi and Y. Park, "R&D proposal screening system based on text-mining approach," *Int. J. Technol. Intell. Plan.*, vol. 2, no. 1, pp. 61–72, 2006.
- [14] K. Girotra, C. Terwiesch, and K. T. Ulrich, "Valuing R&D projects in a portfolio: Evidence from the pharmaceutical industry," *Manage. Sci.*, vol. 53, no. 9, pp. 1452–1466, Sep. 2007.
- [15] Y. H. Sun, J. Ma, Z. P. Fan, and J. Wang, "A group decision support approach to evaluate experts for R&D project selection," *IEEE Trans. Eng. Manag.*, vol. 55, no. 1, pp. 158–170, Feb. 2008.
- [16] Y. H. Sun, J. Ma, Z. P. Fan, and J. Wang, "A hybrid knowledge and model approach for reviewer assignment," *Expert Syst. Appl.*, vol. 34, no. 2, pp. 817–824, Feb. 2008.
- [17] S. Hettich and M. Pazzani, "Mining for proposal reviewers: Lessons learned at the National Science Foundation," in *Proc. 12th Int. Conf. Knowl. Discov. Data Mining*, 2006, pp. 862–871.
- [18] R. Feldman and J. Sanger, *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. New York: Cambridge Univ. Press, 2007.
- [19] M. Konchady, *Text Mining Application Programming*. Boston, MA: Charles River Media, 2006.
- [20] C. P. Wei and Y. H. Chang, "Discovering event evolution patterns from document sequences," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 37, no. 2, pp. 273–283, Mar. 2007.

- [21] T. H. Cheng and C. P. Wei, "A clustering-based approach for integrating document-category hierarchies," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 38, no. 2, pp. 410–424, Mar. 2008.
- [22] H. C. Yang and C. H. Lee, "A text mining approach for automatic construction of hypertexts," *Expert Syst. Appl.*, vol. 29, no. 4, pp. 723–734, Nov. 2005.
- [23] M. Nagy and M. Vargas-Vera, "Multiagent ontology mapping framework for the semantic web," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 41, no. 4, pp. 693–704, Jul. 2011.
- [24] W. Fan, D. M. Gordon, and P. Pathak, "An integrated two-stage model for intelligent information routing," *Decis. Support Syst.*, vol. 42, no. 1, pp. 362–374, Oct. 2006.
- [25] G. H. Lim, I. H. Suh, and H. Suh, "Ontology-based unified robot knowledge for service robots in indoor environments," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 41, no. 3, pp. 492–509, May 2011.
- [26] C. Lu, X. Hu, and J. R. Park, "Exploiting the social tagging network for web clustering," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 41, no. 5, pp. 840–852, Sep. 2011.
- [27] A. J. C. Trappey, C. V. Trappey, F. C. Hsu, and D. W. Hsiao, "A fuzzy ontological knowledge document clustering methodology," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 39, no. 3, pp. 806–814, Jun. 2009.
- [28] M. Zhang, Z. Lu, and C. Zou, "A Chinese word segmentation based on language situation in processing ambiguous words," *Inf. Sci.*, vol. 162, no. 3/4, pp. 275–285, Jun. 2004.
- [29] T. Ong, H. Chen, W. Sung, and B. Zhu, "Newsmap: A knowledge map for online news," *Decis. Support Syst.*, vol. 39, no. 4, pp. 583–597, Jun. 2005.
- [30] Y. Liu, C. Xu, Q. Zhang, and Y. Pan, "The smart architect: Scalable ontology-based modeling of ancient Chinese architectures," *IEEE Intell. Syst.*, vol. 23, no. 1, pp. 49–56, Jan./Feb. 2008.
- [31] D. A. Chiang, H. C. Keh, H. H. Huang, and D. Chyr, "The Chinese text categorization system with association rule and category priority," *Expert Syst. Appl.*, vol. 35, no. 1/2, pp. 102–110, Jul./Aug. 2008.
- [32] H. C. Yang, C. H. Lee, and D. W. Chen, "A method for multilingual text mining and retrieval using growing hierarchical self-organizing maps," *J. Inf. Sci.*, vol. 35, no. 1, pp. 3–23, Feb. 2009.
- [33] F. M. Ham and I. Kostanic, *Principles of Neurocomputing for Science & Engineering*. New York: McGraw-Hill, 2001. Vol. 2 Issue 1, January- 2013
- [34] J. Vesanto and E. Alhoniemi, "Clustering of the self-organizing map," *IEEE Trans. Neural Netw.*, vol. 11, no. 3, pp. 586–600, May 2000.
- [35] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Redwood City: Addison-Wesley, 1989.
- [36] Z. P. Fan, Y. Chen, J. Ma, and Y. Zhu, "Decision support for proposal grouping: A hybrid approach using knowledge rule and genetic algorithm," *Expert Syst. Appl.*, vol. 36, no. 2, pp. 1004–1013, Mar. 2009.
- [37] Y. Liu, X. Wang, and C. Wu, "ConSOM: A conceptual self-organizing map model for text clustering," *Neurocomputing*, vol. 71, no. 4–6, pp. 857–862, Jan. 2008.
- [38] D. Fensel, *Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce*. Berlin, Germany: Springer-Verlag, 2004.
- [39] J. Plisson, P. Ljubic, I. Mozetic, and N. Lavrac, "An ontology for virtual organization breeding environments," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 37, no. 6, pp. 1327–1341, Nov. 2007.
- [40] M. Cai, W. Y. Zhang, and K. Zhang, "ManuHub: A semantic web system for ontology-based service management in distributed manufacturing environments," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 41, no. 3, pp. 574–582, May 2011.
- [41] Y. Liu, Y. Jiang, and L. Huang, "Modeling complex architectures based on granular computing on ontology," *IEEE Trans. Fuzzy Syst.*, vol. 18, no. 3, pp. 585–598, Jun. 2010.
- [42] L. Razmerita, "An ontology-based framework for modeling user behavior—A case study in knowledge management," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 41, no. 4, pp. 772–783, Jul. 2011.
- [43] L. Zhou and D. Zhang, "An ontology-supported misinformation model: Toward a sigital misinformation library," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 37, no. 5, pp. 804–813, Sep. 2007.
- [44] Q. Liang, X. Wu, E. K. Park, T. M. Khoshgoftaar, and C. H. Chi, "Ontology-based business process customization for composite web services," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 41, no. 4, pp. 717–729, Jul. 2011.
- [45] O. Liu and J. Ma, "A multilingual ontology framework for R&D project management systems," *Expert Syst. Appl.*, vol. 37, no. 6, pp. 4626–4631, Jun. 2010.