RESEARCH ARTICLE OPEN ACCESS

An approach to the integration of knowledge maps

Ming Li[1], Mengyue Yuan[1], Yingcheng Xu[2,*]
[1](School of Business Administration, China University of Petroleum-Beijing, China)
[2](China nationalinstitute of Standardization, China)

ABSTRACT
Knowledge plays more and more important role in the heavy competition environment. In order to facilitate the inter-enterpriseknowledge sharing, the knowledge map in each enterprise needs to be integrated. In the paper, the knowledge map integration approach is proposed. Firstly, the explicit knowledge is integrated. In the integration, the documents in the original knowledge map are classified into the corresponding categories in the main knowledge map according to the relevance. The unclassified documents are clustered to derive new categories. Afterwards, experts, which are the owner of implicit knowledge, are classified based on the registered documents or the documents in the original categories. Finally, the ranking of experts in each category are determined. The illustrative example shows the proposed approach is feasible and performances well.
Keywords–knowledge map, knowledgemanagement, expert yellow pages

I. INTRODUCTION

Knowledge is the important strategic resource. Enterprises devote their efforts to manage these valuable resources. Knowledge management systems are implemented to help the collecting, storing and sharing of knowledge [1]. According to the forms, knowledge can be classified into explicit knowledge and implicit knowledge [2]. Explicit knowledge is the knowledge that can be codified. It often exists in documents, manual and handbooks. On the contrary, implicit knowledge is the knowledge that is difficult to be codified such as experience and intuition. This kind of knowledge often exists on the owners' brain. Because of the knowledge explosion, it is more and more difficult to find the appropriate knowledge. Knowledge map is commonly used for finding knowledge via the navigation. Many researches have been put on the knowledge map [4-14].

With the rapid changing environment, enterprises collaborate with each other to achieve the competition advantage [3]. In the collaboration, the knowledge map in each enterprise needs to be integrated to facilitate the finding of knowledge. However, most researches focus on the construction of knowledge map and few researches are put on the integration of knowledge maps. Therefore, in the paper, the knowledge map integration approach is proposed. In the approach, not only explicit knowledge but also implicit knowledge is integrated.The rest of the paper is organized as follows. In the following section, the related works are analyzed. Afterwards, the integration approach is proposed in the third section. In the fourth section, the illustrative example is given. Finally, the conclusion is provided.

II. RELATED WORKS

Many studies have been proposed to ease the burden of knowledge overload. Searching and browsing are two important ways to find the knowledge[4]. Search engine are often used for searching knowledge. With the search engine the unrelated knowledge can be removed from the candidate lists. It focuses on some specific knowledge and fitter for resolving a specific problem. Knowledge maps provide categories of knowledge. With the knowledge map, the knowledge can be gotten systematically. It has attracted the attention of researchers. Because the knowledge includes explicit knowledgeandimplicit knowledge, the research can also be classified into two categories.

In the research of explicit knowledge map, Ong et al. [5] proposed a method for constructing the hierarchical knowledge map. It focuses on the news. Lin and Hsueh [6] provided a method to construct the knowledge map in virtual communities of practice by using information retrieval and data mining techniques. Yang [7] developed a knowledge map for construction scheduling. Li et al. [8] proposed a method for the construction of personalized knowledge map. In the method, the documents are classifiedaccording the preferences of users. In order to find the hidden structure of explicit knowledge and facilitate the navigation, Li and Chang [9] proposed a layered thematic K-map system. Yoon et al. [10] provided an approach to construct knowledge maps. In the approach, the text mining method is used to extracting meaningful information, the network analysis is used to represent the relation between categories and measure the value of network. In the method proposed by Hao et al.[5], the text mining method is used to construct the knowledge map. Moreover, the social network analysis is used to find the important knowledge to alleviate knowledge overload. In order to facilitate the knowledge retrieval in knowledge map, Xiong et al. [11] provided a searchable knowledge map model.

In the research of implicit knowledge map, Yang and Huh [12] proposed the automatic profiling method based on the registered documents. In the method, activeness, relevance and usefulness are used to characterize the expertise of experts. Li et al. [13] proposed the personalized expert yellow pages construction method. In the method, experts are classified based on the relevance of the expertise with the category and ranked according to the preferences of users. Lichtnow et al. [14] proposed a method to identify the expertise of experts from their curricula vitae. In the method, the texts in the curricula are analyzed and the domain ontology is used compare the concepts.

III. APPROACH TO THE INTEGRATION OF KNOWLEDGE MAPS

In the integration of knowledge maps, both explicit knowledge and implicit knowledge needs to be integrated. The knowledge map is composed of documents and experts along with the categories. In each category, there are both documents and experts.

The knowledge map that will be integrated into is called the main knowledge map. The knowledge map to be integrated into the main knowledge map is called original knowledge map. There is probably more than one original knowledge map that covers the same set of documents [8].

The ith original knowledge map is denoted by KP_i. The sets of documents in the main knowledge map KP_0 and in the original knowledge map KP_i are represented by $SD_0$and SD_i, respectively. Meanwhile, the sets of experts in the main knowledge map KP_0 and in the original knowledge map KP_i are represented by $SE_0$and SE_i , respectively.

The steps to integrate knowledge maps are given as follows.

(1) Calculating the similarity between document d_p in the set SD_i and mth category c_m^0 in the main knowledge map KP_0.

$$SimDC_{p,m} = simiarity\left(d_p, c_m^0\right) \quad (1)$$

(2) Classifying the document d_p to the most relevant category c_*^0.

$$c_*^0 = \{c_t^0 | SimDC_{p,t} > \lambda, SimDC_{p,t} \geq SimDC_{p,\tau}, c_t^0, c_\tau^0 \in c^0\} \quad (2)$$

where, λ is the threshold.

(3) Clustering the unclassified documents.

The unclassified documents need to be clustered to derive new categories [8].

Firstly, the original knowledge maps that have the most similar classification preference with the main knowledge map are identified. The similarity $SimKK_{0,v}$ between knowledge mapsKP_0 and KP_v is determined whether the two documents are in the same category.

$$SimKK_{0,v} = \sum_{x=1}^{N_d} \sum_{y=1}^{N_d} \frac{|A_{xy}^0 \times A_{xy}^v|}{(N_d)^2} \quad (3)$$

where, A_{xy}^0 and A_{xy}^v means whether the document d_x and d_y are in the same category in the knowledge map KP_0 and KP_v. $N_d$is the number of documents.

Then the most similar original knowledge map can be identified with

$$KP_*^0 = \{KP_v | SimKK_{0,v} \geq SimKK_{0,i}, KP_v, KP_i \in SKP\} \quad (4)$$

where, SKP is the set of knowledge maps that covers the same set of documents.

Secondly, the unclassified documents need to be clustered. In the clustering, the similarity $SD_{i,j}$ between documents d_p and d_m is determined by both the textural similarity and the classification similarity.

$$SD_{i,j} = \zeta SimDD_{p,m} + (1-\zeta) SimKK_{0,*} \quad (5)$$

where, $SimDD_{p,m}$ is the textural similarity between documents d_p and d_m.

Then based on the KNN method [15], the new categories are defined and documents are clustered.

In the step, the documents are integrated into the main knowledge map. The integration of explicit knowledge is finished.

(4) Calculating the belonging degree of experts to the category.

The classification of experts is determined by the registered documents or the documents in the category that the expert belongs to in the original knowledge map.

The more registered documents that are classified into the category, the expert are more possible to be classified to the category.

$$BelR_{y,z} = \frac{N_{y,z}^r}{N_z}. \quad (6)$$

where, N_z is the number of documents in the category c_z^0, $N_{y,z}^r$ is the number of registered documents in the category c_z^0 by expert e_y.

For the experts that have no registered documents, we use the documents in the category that the expert belongs to in the original knowledge map to identify the similarity between the expert and the categories.

$$BelC_{y,z} = \frac{N_{y,z}^c}{N_z}. \quad (7)$$

where, N_z is the number of documents in the category c_z^0,$N_{y,z}^r$ is the number of documents that belongs to the same category with expert e_y in the original knowledge map in the category c_z^0.

(5) Assign the expert e_y to the most relevant category c_*^0.

$$c_*^0 = \{c_t^0 | Bel_{y,z} \geq Bel_{y,\tau}, c_t^0, c_\tau^0 \in c^0\}. \quad (8)$$

where, c^0 is the set of categories in the main knowledge map KP_0.

(6) Ranking the experts

There are many factors such as expertise, background, trust and familiarity that influence the preference of users in the selection of experts [13]. The expertise level of the expert can be obtained by the registered documents with the highest quality. If there are no registered documents, the expertise is set the default value. The experts in the original knowledge map are often in the other organizations and the owner of the main knowledge map probably doesn't know. So the most value of criteria trust, background and familiarity are difficult to be derived. In order to rank experts on the whole, these missing values are set the default value. With the continuous using of the knowledge map, the value of these factors can be derived precisely.

In the integration, the granularities of linguistic terms which are used to rate are probably not identical. We use the 2-tuple linguistic model to represent the ratings and these ratings with different granularity linguistic terms need to be unified by [16]

$$T_{t2}^{t1}\left(s_i^{n(t2)}, \alpha^{n(t2)}\right) = \Delta\left(\frac{\Delta^{-1}\left(s_i^{n(t2)}, \alpha^{n(t2)}\right)\cdot(n(t1)-1)}{n(t2)-1}\right) \quad (9)$$

where, $t1$ and $t2$ are the granularities of linguistic terms.

IV. ILLUSTRATIVE EXAMPLE

There are three knowledge maps which include one main knowledge map $K0 = \left\{\begin{matrix} C01(d01, d02), \\ C02(d03, d04, e02, e03), \\ C03(d05, d06) \end{matrix}\right\}$ and two original knowledge maps

$$K1 = \left\{\begin{matrix} C11(d07, d08, d11, e00), \\ C12(d09, d10, d12), \\ C13(d01, d02, d13, d14, e01) \end{matrix}\right\} \text{and}$$

$$K2 = \left\{\begin{matrix} C21(d07, d08, d11, e01), \\ C22(d01, d09, d12, d14, e01), \\ C23(d02, d10, d13) \end{matrix}\right\}.$$

The derived similarity between documents and the category in advance is shown in Table 1.

Table 1 Similarity between documents and categories

	$d07$	$d08$	$d09$	$d10$	$d11$	$d12$	$d13$	$d14$
$C01$	0.35	0.51	0.38	0.5	0.39	0.44	0.69	0.59
$C02$	0.64	0.19	0.26	0.38	0.22	0.28	0.25	0.28
$C03$	0.40	0.31	0.39	0.51	0.26	0.40	0.31	0.53

In the integration of documents, the similarity threshold is set 0.5. According to the similarity in Table 1, with Eq.(2), the document $d07$ is classified into category $C02$, document $d08$ is classified into category $C01$, document $d10$ is classified into category $C03$, document $d13$ is classified into category $C01$, document $d14$ is classified into category $C01$. Since the similarities are smaller than 0.5, the documents $d09$, $d11$ and $d12$ are not classified.

For the unclassified documents, with the Eq.(3), the similarities between knowledge maps are derived by

$$sim(K0, K1) = \sum_{x=1}^{N_d}\sum_{y=1}^{N_d}\frac{|A_{xy}^0 \times A_{xy}^1|}{(N_d)^2} = 0.097$$

$$sim(K0, K2) = \sum_{x=1}^{N_d}\sum_{y=1}^{N_d}\frac{|A_{xy}^0 \times A_{xy}^2|}{(N_d)^2} = 0.083$$

By using Eq.(4), the most similar original knowledge map with the main knowledge map $K0$is the original knowledge map $K1$.Then based on the KNN method, the new categories along with the corresponding set of documents can be identified.In the following, the experts are integrated into the main knowledge map. Experts in the original knowledge map $K1$ and $K2$ are integrated into the main knowledge map $K0$.

The author of document $d07$is $e00$, by using Eq.(6), the belonging degrees to thecategories in main knowledge map $K0$are derived by

$$BelR_{00,01} = \frac{N_{00,01}^r}{N_{01}} = 0$$

$$BelR_{00,02} = \frac{N_{00,02}^r}{N_{02}} = 0.2$$

$$BelR_{00,03} = \frac{N_{00,03}^r}{N_{03}} = 0$$

With the Eq.(8), expert $e00$ is assigned to $C02$. There are no related documents of expert $e01$. By using Eq.(7), the belonging degrees to thecategories in main knowledge map $K0$ are derived by

$$BelC_{01,01} = \frac{N_{01,01}^c}{N_{01}} = 1$$

$$BelC_{01,02} = \frac{N_{01,02}^c}{N_{02}} = 0$$

$$BelC_{01,03} = \frac{N_{01,03}^c}{N_{03}} = 0$$

With the Eq.(8), the expert $e01$ is classified into $C01$.The authors of documents along with the quality of documents in category $C02$ are shown as in Table 2.

Table 2 Quality of documents in the category $C02$

Experts	$e00$	$e02$	$e03$
Registered documents	$d07$	$d03$	$d04$
Ratings to documents	VH	VH	H
Granularities	7	5	5

In Table 2, we can see that the granularities of linguistic terms are different. By using Eq.(9), these ratings are unified as

$$T_7^5(VH, 0) = (L, 0.4)$$
$$T_7^5(H, 0) = (L, -0.2)$$

Based on the expertise, the ranking of the experts in the category $C02$is $e00 \succ e02 \succ e03$.

V. CONCLUSION

In the study, the approach to the integration of knowledge maps is proposed. Considering the different kind of knowledge, firstly, the explicit knowledge is integrated, and then the implicit knowledge is integrated. In the integration of explicit knowledge, firstly, the documents in the original knowledge maps are classified into corresponding categories. The unclassified documents are clustered based on the classification similarity and textual similarity. Then the experts are classified into each category based on the registered documents and the documents in the originalcategory. Finally the experts are ranked in each category. The illustrative example shows the proposed method is feasible and performances well.

ACKNOWLEDGEMENTS

The research is supported by the National Natural Science Foundation ofChina under Grant No. 71101153, 71571191,71301152, Science Foundation of ChinaUniversity of Petroleum, Beijing (No. 2462015YQ0722), Humanity and Social Science YouthFoundation of Ministry of Education in China (Project No. 15YJCZH081, 13YJC790112) and the Basic Scientific Research Funding(Project No. 242016Y-4700)

REFERENCES

[1] Alavi M, Leidner D E, Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues,*MIS quarterly, 25(1)*, 2001, 107-136.

[2] Mahr D, Lievens A, Virtual lead user communities: Drivers of knowledge creation for innovation, *Research policy, 41(1)*, 2012,167-177.

[3] Camarinha-Matos L M, Pantoja-Lima C,Cooperation coordination in virtual enterprises, *Journal of intelligent manufacturing, 12(2)*, 2001, 133-150.

[4] Hao J., Yan Y., GongL., Wang G., & Lin J, Knowledge map-based method for domain knowledge browsing,*Decision Support Systems,61*,2014, 106-114.

[5] Ong T H, Chen H, Sung W, et al, Newsmap: a knowledge map for online news, *Decision Support Systems, 39(4)*, 2005, 583-597.

[6] Lin F, Hsueh C, Knowledge map creation and maintenance for virtual communities of practice,*InformationProcessing&Management, 42(2)*,2006,551-568.

[7] Yang J B, Developing a knowledge map for construction scheduling using a novel approach, *Automation in construction, 16(6)*, 2007, 806-815.

[8] LiM,Yuan MY, Xiong HT, An approach to the construction of personalized knowledge map based on the collaborative tagging, *International Journal of Knowledge Engineering, 1(3)*, 2015, 209-213

[9] LiST., & ChangWC, Design and evaluation of a layered thematic knowledge map system, *Journal of Computer Information Systems, 49(2)*,2008, 92-103.

[10] Yoon B, Lee S, Lee G, Development and application of a keyword-based knowledge map for effective R&D planning, *Scientometrics, 85(3)*, 2010, 803-820.

[11] Xiong Q., Wang Y., Guo J., & Wu G. , A Searchable knowledge map based on ontology, *Proc. fourth International Conference on Semantics, Knowledge and Grid*, Beijing,IEEE, 2008, 457-460.

[12] Yang K W, Huh S Y, Automatic expert identification using a text categorization technique in knowledge management systems, *Expert Systems with Applications, 34(2)*,2008, 1445-1455.

[13] LiM., XuD., & ZhuYC, Construction of Personalized Expert Yellow Page based on Fuzzy Linguistic Method, *Advances in intelligent system research, 117*, 2015, 1489-1493

[14] Lichtnow D, Loh S, Junior L C R, et al, Using Text Mining on Curricula Vitae for Building Yellow Pages, 2008.

[15] Jiang S, Pang G, Wu M, et al, An improved K-nearest-neighbor algorithm for text categorization, *Expert Systems with Applications*, 39(1), 2012,1503-1509.

[16] Martı L, Herrera F, An overview on the 2-tuple linguistic model for computing with words in decision making: Extensions, applications and challenges,*Information Sciences, 207*, 2012, 1-18.