Construction of Personalized Expert Yellow Page based on Fuzzy Linguistic Method

Ming Li
School of Business Administration
China University of petroleum
Beijing, China
brightliming@outlook.com

Dan Xu
School of Business Administration
China University of petroleum
Beijing, China
momo_0025@163.com

Yanchun Zhu
School of Economics and Business Administration
Beijing Normal University
Beijing, China
kddzw@163.com
* Corresponding Author

Abstract—Expert yellow page is the knowledge map of experts. When selecting experts for consulting, besides expertise, learners take personalized factors such as trust, background and familiarity into consideration. In order to fulfill personalized requirements of learners, the construction approach of personalized expert yellow page based on fuzzy linguistic method is proposed. The expertise is identified based on ratings to submitted documents. In the aggregation of ratings, the expertise of the respondent and the degree of agreement are integrated to determine the weight of respondents. The max-min trust prediction method is used to propagate indirect trust degrees. The familiarity and the background are also defined to characterize the preference of learners. The incomplete weight information is required instead of precise values to facilitate the expression of opinions about weights of factors. Based on TOPISIS, the optimal model is constructed to derive the precise weight and then the ranking of experts whthin each category are obtained. The illustrative example shows the proposed method is feasible and efficient.

Keywords*- Knowledge Map; Expert Yellow Page; Fuzzy Linguistic Method; Knowledge Managment*

I. INTRODUCTION

Knowledge is classified into explicit knowledge and implicit knowledge [1]. The knowledge that can be codified is called explicit knowledge. It often exists in textbooks, documents, manuals and reports. On the contrary, the knowledge that cannot be codified is called implicit knowledge, which often exists in the owner's brain. The implicit knowledge is more important since it is difficult to be copied and imitated. Consulting the expert is an important way to learn the implicit knowledge [2]. Expert yellow page provides a way to find the appropriate expert by categories [3]. With the rapid changing environment, temporary groups or alliances are always formed. In these organizations, workers cannot know each expert in detail and they have their own personalized requirements on the experts. There is an urgent need of the expert yellow pages which reflects both the expertise of the expert and worker's own preferences.

In the construction of the expert yellow page, experts are classified according to the knowledge area of the expertise and are ranked according to the expertise level. The modeling of expertise relies on documents [4]. Documents with different quality play different roles. Ratings to documents reflect the quality and the accuracy of ratings depends on the expertise level of respondents. Therefore in the modeling of the expertise, the quality of documents along with the expertise level of respondents needs to be considered. The consulting process is the communication process, which involves the interaction between expert and learners. Besides the expertise, the personalized factors such as trust, background and familiarity, reflect the learner's preference and have great influence on effects of the communication process [5]. Providing the personalized expert yellow page can facilitate the finding of the appropriate expert.

In this paper, an approach to the construction of personalized expert yellow page is proposed. Firstly, the expertise is modelled based on ratings to the submitted documents along with the weight of respondents, which will be derived from the expertise level and the degree of agreement. The trust prediction method is used to propagate indirect trust degrees. Moreover, the familiarity and background are defined. Based on the incomplete weight information given by the learner, the optimal model is constructed to derive the precise weight information and then the ranking of the experts are derived. The remainder of this paper is organized as follows. In the next section, literature in fuzzy linguistic method is introduced. In Section 3 researchers present the approach to the construction of personalized expert yellow page. In Section 4, the illustrative example is given. Finally, conclusions are discussed.

II. FUZZY LINGUISTIC METHOD

The fuzzy linguistic method is popular to deal with the linguistic evaluation information. 2-tuple linguistic model is the often used model to represent the linguistic terms, which improves the accuracy of linguistic evaluation information processing and make the processing results interpretable. The 2-tuple linguistic model is briefly reviewed in the following [6-9].

Let $T_0 = \{t_i | i = 0,1, \cdots, g\}$ be a linguistic term set, $T = \{(t_1, \alpha_1), (t_2, \alpha_2), \cdots, (t_n, \alpha_n)\}$ be a set of 2-tuple linguistic variable, $U = \{u_1, u, \cdots, u_n\}$ be the numerical weight set, $U_l = \{(u_1, \alpha_1^u), (u_2, \alpha_2^u), \cdots, (u_n, \alpha_n^u)\}$ be the 2-tuple linguistic weights, s_i with the numerical value $\beta (0 \leq \beta \leq g)$ be the aggregation result.

Definition 1 [6-9]. By the function Δ, the 2-tuple linguistic tuple of t_i is derived and β is translated into the corresponding 2-tuple linguistic variable

$$\Delta (t_i) = (t_i, 0), t_i \in T \tag{1}$$

$$\Delta (\beta) = (t_i, \alpha) = \begin{cases} t_i & , i = Round(\beta) \\ \alpha = \beta - i, & \alpha \in [-0.5, 0.5) \end{cases} \tag{2}$$

The 2-tuple linguistic variable (t_i, α) is translated into an corresponding crisp value β by

$$\Delta^{-1} (t_i, \alpha) = i + \alpha = \beta \tag{3}$$

Definition 2 [6-9]. Arithmetic average operator is defined as

$$(\bar{t}, \bar{\alpha}) = \Delta\left(\frac{1}{n}\sum_{i=1}^{n} \Delta^{-1}(t_i, \alpha_i)\right) = \Delta\left(\frac{1}{n}\sum_{i=1}^{n} \beta_i\right), \bar{t} \in S, \bar{\alpha} \in [-0.5, 0.5) \tag{4}$$

Definition 3 [6-9]. Weighted average operator is defined as

$$(\bar{t}^u, \bar{\alpha}^u) = \Delta\left(\frac{\sum_{i=1}^{n} \Delta^{-1}(t_i, \alpha_i) \times u_i}{\sum_{i=1}^{n} u_i}\right) = \Delta\left(\frac{\sum_{i=1}^{n} \beta_i \times u_i}{\sum_{i=1}^{n} u_i}\right), , \bar{t}^u \in S, \bar{\alpha}^u \in [-0.5, 0.5) \tag{5}$$

Definition 4 [6-9]. Linguistic weighted average operator is defined as

$$(t_l^u, \alpha_l^u) = \Delta\left(\frac{\sum_{i=1}^{n} \Delta^{-1}(t_i, \alpha_i) \times \Delta^{-1}(u_i, \alpha_i^u)}{\sum_{i=1}^{n} \Delta^{-1}(u_i, \alpha_i^u)}\right) = \Delta\left(\frac{\sum_{i=1}^{n} \beta_i \times \beta_i^u}{\sum_{i=1}^{n} \beta_i^u}\right), t_l^u \in S, \alpha_l^u \in [-0.5, 0.5) \tag{6}$$

Definition 5 [6-9]. Let (t_i, α_i) and (t_j, α_j) be two 2-tuple linguistic variables,

1. if $i > j$, (t_i, α_i) is better than (t_j, α_j);
2. If $i = j$ and $\alpha_i > \alpha_j$ then (t_i, α_i) is better than (t_j, α_j);
3. If $i = j$ and $\alpha_i < \alpha_j$ then (t_i, α_i) is worse than (t_j, α_j); (7)
4. If $i = j$ and $\alpha_i = \alpha_j$ then (t_i, α_i) is equal to (t_j, α_j).

III. CONSTRUCTION OF PERSONALIZED EXPERT YELLOW PAGE

In the construction of personalized expert yellow page, the expertise is the key. It determines the classification of the expert. Besides the expertise level, in the ranking of the expert within a specific category, the personalized factors such as trust, familiarity and background are integrated.

A. Expertise

The expertise model is constructed based on submitted documents. The documents with high quality indicate the high expertise of the creator. Ratings given to documents about the quality reflect the quality of the document. The accuracy of ratings depends on the respondent's expertise. Higher expertise level represents the higher authority in the specific knowledge area. It makes the rating more accurate and more reliable. In the following, the method presented by Chen [10] is extended to calculate the weight of respondents.

Step1: Calculating the degree of agreement between the respondents R_p and R_q . The degree of agreement $S(R_p, R_q)$ of the opinions between respondents R_p and R_q is obtained by

$$S(R_p, R_q) = \sum_{d=1}^{N_d} r_d^p - r_d^q \tag{8}$$

where, r_d^p, r_d^q are ratings to the document D_d given by respondents R_p,R_q, N_d is the number of documents.

Step2: The average degree of agreement $A(R_p)$ of respondent R_p is derived by

$$A(R_p) = \frac{1}{n-1}\sum_{\substack{q=1 \\ p \neq q}}^{N_R} S(R_p, R_q) \tag{9}$$

where, N_R is the number of respondents.

Step3: The relative degree of agreement $RA(R_p)$ of respondent R_p is gotten by

$$RA(R_p) = \frac{A(R_p)}{\sum_{R_i \in R} A(R_i)} \tag{10}$$

where, R denotes the set of respondents.

Step4: The weight of the expertise level of respondents and the weight of the relative degree of agreement of the respondents are denoted by y_1 and y_2 , respectively, where $y_1 \in [0,1]$ and $y_2 \in [0,1]$.

The final weight w_j^p of the respondent R_p in knowledge area k_j can be got by

$$w_j^p = \frac{y_1}{y_1 + y_2} * \frac{v_{pj}}{G} + \frac{y_1}{y_1 + y_2} * RA(R_p) \tag{11}$$

where, v_{pj} denotes the expertise level of respondent R_p in the knowledge area k_j, G is the granularity of the linguistic term set.

Then the expertise level of the expert E_i in the knowledge area k_j can be derived by

$$VE_{ij} = \frac{\sum_{d=1}^{M_i}\sum_{p=1}^{N_r} w_j^p \times r_d^p \times VD_{ij}}{\sum_{d=1}^{M_i}\sum_{p=1}^{N_r} w_j^p \times r_d^p} \tag{12}$$

where, M_i is the number of documents that the expert E_i submitted, VD_{ij} denotes the belonging degree of the document d_i to the knowledge area k_j.

B. Trust

Learners are allowed to rate his/her trust in other learners and experts by using linguistic terms. The number of experts that the learner rates directly is limited. For the unrated experts, the trust can be predicted based on direct ratings. The max-min aggregation method among shorted paths [11] is extended in fuzzy linguistic environment to propagate the level of indirect trust in others.

For the strength of trust path to each learner $u_i \in Inlink\_Neighbors(U_t)$

$$ST_{si} = \max_{u_j \in outlink\_neighbors(U_s)} [\min\{(\tau_{sj}, 0), ST_{ji}\}] \quad (13)$$

Then, final estimated trust value is

$$T_{st} = \min\{ST^*_{s,i^*}, (\tau_{i^*,t}, 0)\} \quad (14)$$

where, ST^*_{s,i^*} is the highest strength of trust path, the most reliable $inlink\_neighbors$

$u_{i^*} = \arg\max_{u_i \in inlink\_neighbors(U_t)} ST_{si}$,

C. Background

Background means whether the learner has known the expertise in detail. The rating given to the documents which is submitted by the expert is used to characterize the familiarity with the expertise background. The degree that the learner U_i knows the background of the expert E_j can be obtained by

$$B_{ij} = (b_{ij}, \sigma_{ij}) = \frac{\sum_{k=0}^{N_j} r_{ik}}{N_j \times g} = \Delta\left(\frac{\sum_{k=0}^{N_j} \Delta^{-1}(r_{ik}, 0)}{N_j \times g}\right) \quad (15)$$

where, r_{ik} is the rating given by the learner U_i to the document D_{jk}, N_j is the number of documents that the expert E_j submitted.

D. Familiarity

Familiarity means whether they have contacted before. The more times that they have contacted, the more familiarity they are. Since the familiarity with the task may be deduced as time goes on, the time factor [12] is used in the calculations. The familiarity of the learner U_i with the expert E_j can be obtained by

$$F_{ij} = \sum_{k=1}^{N_{ij}} \frac{1}{e^{\tau(t_{now} - t_k)}} \quad (16)$$

where, N_{ij} is the contacted times between the learner U_i and the expert E_j, t_{jt} is the contact time, t_{now} is the present date, and τ is a tunable parameter.

E. Classification and ranking of experts

The expert is classified into the category according to the knowledge area with the highest expertise level. In the specific category, the ranking of experts are derived by extending the method [13], which is shown in the following.

Let $E = \{E_1, E_2, \ldots, E_t\}$ be the set of experts, $C = \{C_1 = \text{expertise}, C_2 = \text{trust}, C_3 = \text{background}, C_4 = \text{familarity}\}$ be the set of criteria. Suppose $R = (r_{ij})_{m\times 4}$ is the derived evaluation matrix of the leaner U, for the expert E_i. The information about criteria weights is incompletely known [14, 15]. Let $w = (w_1, w_2, \ldots, w_4) \in H$ be the weight vector of criteria, where $w_j \in [0,1]$, $\sum_{j=1}^{4} w_j = 1$. The weight information is given in the following forms. For $i \neq j$, $w_i \geq w_j$; $w_i - w_j \geq \alpha_i$, $\alpha_i > 0$; $w_i - w_j \geq w_k - w_l$, $j \neq k \neq l$; $w_i \geq \beta_i w_j, 0 \leq \beta_i \leq 1$; $\gamma_i \leq w_j \leq \gamma_i + \varepsilon_i, 0 \leq \gamma_i < \gamma_i + \varepsilon_i \leq 1$.

Step 1 With the function Δ, the linguistic evaluation value r_{ij} is transformed into the 2-tuple linguistic evaluation value (r_{ij},0).

Step 2 The Positive ideal solution (PIS) and negative ideal solution (NIS) are defined as

$$\begin{aligned} r^+ &= \left((r_1^+, a_1^+), (r_2^+, a_2^+), (r_3^+, a_3^+), r_4^+\right) \\ r^- &= \left((r_1^-, a_1^-), (r_2^-, a_2^-), (r_3^-, a_3^-), r_4^-\right) \end{aligned} \quad (17)$$

For the linguistic values,

$$\left(r_j^+, a_j^+\right) = \max_i\{(r_{ij}, a_{ij})\}, j = 1,2,3$$

$$\left(r_j^-, a_j^-\right) = \min_i\{(r_{ij}, a_{ij})\}, j = 1,2,3$$

For the numerical values,

$$r_j^+ = \min_i\{r_{i4}\}, j = 4; r_j^- = \max_i\{r_{i4}\}, j = 4$$

Step 3 Calculating the distances of each alternative from PIS and NIS. The idea of the TOPSIS method is that the best alternative is nearest to the PIS and is most far from the NIS. Considering the information about criteria weights is incompletely known, the objective optimization model is established to derive precise weights.

$$\begin{cases} \min r^+ = \Delta(\sum_{j=1}^{3} \left|\Delta^{-1}(r_{ij}, a_{ij}) - \Delta^{-1}(r_i^+, a_i^+)\right| w \\ \qquad + |r_{i4} - r_4^+| w_4) \\ \max r^- = \Delta(\sum_{j=1}^{3} \left|\Delta^{-1}(r_{ij}, a_{ij}) - \Delta^{-1}(r_i^-, a_i^-)\right| w \\ \qquad + |r_{i4} - r_4^-| w_4) \\ \qquad subject\ to: w \in H, i = 1,2,\cdots,m \end{cases} \quad (18)$$

Step 4 Relational degree of expert from PIS is obtained by

$$r_i = \frac{\sum_{j=1}^{3} \left|\Delta^{-1}(r_{ij}, a_{ij}) - \Delta^{-1}(r_i^-, a_i^-)\right| w_j + |r_{i4} - r_4^-| w_4}{\sum_{j=1}^{3} \left(2 \times \Delta^{-1}(r_{ij}, a_{ij}) + \Delta^{-1}(r_i^+, a_i^+) - \Delta^{-1}(r_i^-, a_i^-)\right) w_j + (2 \times r_{i4} + r_4^+ - r_4^-) w_4} \quad (19)$$

Step 5 Experts are ranked in an ascending order of relational degrees within each category.

IV. ILLUSTRATIVE EXAMPLE

Suppose there is a newly organized information system development group which includes four experts and three novices. Four knowledge areas are defined which are JAVA, Oracle, JS and CSS. The four experts are denoted by E1, E2, E3 and E4 and the three novices are denoted by U0, U1 and U2. In order to construct the expert yellow page for learner U0, experts need to be to be classified and ranked personalized.

The linguistic term set $S = \{S_0 = \text{Definitely Low (DL)}, S_1 = \text{Very Low (VL)}, S_2 = \text{Low (L)}, S_3 = \text{Average(A)}, S_4 = \text{High(H)}, S_5 = \text{Very High (VH)}, S_6 = \text{Definitely High(DH)}\}$ is given to express opinions. The submitted documents by experts and the classification information of documents are shown in Table 1.

TABLE 1. SUBMITTED DOCUMENTS AND ITS CLASSIFICATION INFORMATION

Expert	Documents	JAVA	Oracle	JS	CSS
E1	**D1**	H	A	VH	DH
	D2	H	VH	A	VL
E2	**D3**	L	A	VL	A
	D4	L	L	L	A
E3	**D5**	A	DH	VH	L
	D6	H	L	H	DH
E4	**D7**	A	VL	VH	VL
	D8	VL	VH	L	A

The expertise of respondents is shown in Table 2.

TABLE 2. EXPERTISE OF RESPONDENTS

Respondents	JAVA	Oracle	JS	CSS
U1	VH	L	VH	L
U2	VL	A	VL	A
U3	VH	H	VH	H
U4	DH	VH	DH	A

The τ is set 1/365 and the contact times along with the contact date is shown in Table 3.

TABLE 3. CONTACT DATE

E1	E2	E3	E4
353	282		333
321	215		
234			

The read and rated documents by the leaner are shown in table 4.

TABLE 4. READ AND RATED DOCUMENTS

D1	D2	D3	D4	D5	D6	D7	D8
A	H	L	A	VH	DH	L	L

The trust ratings given by learners are shown in Table 5. In the table researchers see that the there is no direct rating of trust between learner U0 and expert E3.

TABLE 5. TRUST RATINGS

Learners	E1	E2	E3	E4	U1	U2
U0	DH	H		L	L	A
U1			VH			
U2			H			

With equations (8) to (16), the evaluation results of four experts with regard to the criteria and the PIS and NIS are derived, which are shown in Table 6.

TABLE 6. EVALUATION RESULTS OF FOUR EXPERTS WITH REGARD TO CRITERIA ALONG WITH THE PIS AND NIS

	C1	C2	C3	C4
E1	4.14	2.55	0.58	6
E2	2.44	1.46	0.42	4
E3	4.53	0.92	0.92	3
E4	3.25	0.00	0.33	2
PIS	4.53	2.55	0.92	6
NIS	2.44	0.00	0.33	2

The incomplete weight information are given as

$$w_1 > w_2$$

$$w_3 - w_4 > 0.1$$

$$w_2 - w_3 > w_1 - w_4$$

By using equations (17) to (18), the precise weight is derived as $w = (1,1,0,0.1)$

With the equation (19), relational degrees of experts along with the rankings are obtained, which is shown in Table 7.

TABLE 7. RELATIONAL DEGREE AND RANKINGS OF EXPERTS

Experts	Relational degrees	Rankings
E1	0.92	4
E2	0.33	2
E3	0.62	3
E4	0.16	1

V. CONCLUSION

In the paper, the construction approach of personalized expert yellow page is proposed. In order to get accurate expertise level of experts, the expertise level in the knowledge area and the agreement level of opinions are incorporated into the calculation of weights of respondents. For unrated trusts, the trust prediction method is extended to derive trust degrees. Moreover, the background and familiarity are also defined to characterize personalized preferences of the learner. For the incomplete weight information given by the learner, the optimal model is constructed to get the precise weight information. Afterwards, based on the TOPSIS method, the ranking of the experts is derived. The illustrative example shows the proposed method is feasible and efficient.

ACKNOWLEDGMENT

The research is supported by the National Natural Science Foundation of China under Grant No. 71101153 and Science Foundation of China University of Petroleum, Beijing (No. 2462015YQ0722), Humanity and Social Science Youth Foundation of Ministry of Education in China (No. 13YJC790112)

REFERENCES

1. Alavi, M., & Leidner, D. E., “Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues,” MIS quarterly, pp.107-136. 2001.
2. Choi, S., Kim, J., & Ryu, K., “Effects of context on implicit and explicit lexical knowledge: An event-related potential study,” Neuropsychologia, 63, pp.226-234. 2014.

3. Liu, P., & Li, X., “The Application of Expertise Knowledge Map in Human Resource Management,” In Management and Service Science (MASS), 2011 International Conference on (pp. 1-4). IEEE. 2011.
4. Yang, K. W., & Huh, S. Y., “Automatic expert identification using a text categorization technique in knowledge management systems,” Expert Systems with Applications, 34(2), pp.1445-1455. 2008.
5. Duan, Y., Nie, W., & Coakes, E., “Identifying key factors affecting transnational knowledge transfer,” Information & management, 47(7), pp.356-363. 2010.
6. Herrera, F., & Mart ńez, L., “A 2-tuple fuzzy linguistic representation model for computing with words,” Fuzzy Systems, IEEE Transactions on, 8(6), pp.746-752. 2000.
7. Herrera, F., & Martinez, L., “The 2-tuple linguistic computational model: advantages of its linguistic description, accuracy and consistency,” International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 9(supp01), pp.33-48. 2001.
8. 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, pp.1-18. 2012.
9. Herrera, F., Herrera-Viedma, E., & Martínez, L., “A fusion approach for managing multi-granularity linguistic term sets in decision making,” Fuzzy sets and systems, 114(1), pp.43-58. 2000
10. Chen, S. M., “Aggregating fuzzy opinions in the group decision-making environment,” Cybernetics & Systems, 29(4), pp.363-376. 1998
11. Kim, Y. A., & Song, H. S., “Strategies for predicting local trust based on trust propagation in social networks,” Knowledge-Based Systems, 24(8), pp.1360-1371. 2011.
12. Zhang, J., Ackerman, M. S., & Adamic, L. , “Expertise networks in online communities: structure and algorithms,” In Proceedings of the 16th international conference on World Wide Web, ACM, New York, pp. 221-230. 2007
13. Wei, G. W., “Extension of TOPSIS method for 2-tuple linguistic multiple attribute group decision making with incomplete weight information,” Knowledge and Information Systems.25 (3), pp. 623-634. 2010
14. Kim, S. H., Choi, S. H., & Kim, J. K., “An interactive procedure for multiple attribute group decision making with incomplete information: Range-based approach,” European Journal of Operational Research.118.1, pp.139-152. 1999
15. Kim, S. H., & Ahn, B. S., “Interactive group decision making procedure under incomplete information,” European Journal of Operational Research.116.3, pp.498-507. 1999.