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 within 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 Management

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 pro-
cessing results interpretable. The 2-tuple linguistic model is briefly reviewed in the following [6-9].

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

**Definition 1** [6-9]. By the function \( \Delta \), the 2-tuple linguistic tuple of \( t_i \) is derived and \( \beta \) is translated into the corresponding 2-tuple linguistic variable

\[
\Delta(t_i) = (t_i, 0), t_i \in T
\]

**Definition 2** [6-9]. The 2-tuple linguistic variable \( (t_i, \alpha) \) is derived for the corresponding crisp value \( \beta \) by

\[
\Delta^{-1}(t_i, \alpha) = \begin{cases} 0, & \text{if } \alpha = \beta \\ i, & \text{if } \alpha \neq \beta \end{cases}
\]

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

\[
(f, \bar{a}) = \Delta \left( \frac{1}{n} \sum_{i=1}^{n} \Delta^{-1}(t_i, \alpha_i) \right) = \Delta \left( \frac{1}{n} \sum_{i=1}^{n} \beta_i \right), f \\
\in S, \bar{a} \in [-0.5, 0.5)
\]

**Definition 4** [6-9]. Weighted average operator is defined as

\[
(f^w, \bar{a}^w) = \Delta \left( \frac{\sum_{i=1}^{n} \beta_i u_i \alpha_i}{\sum_{i=1}^{n} u_i} \right)
\]

**Definition 5** [6-9]. Let \( (t_i, \alpha_i) \) and \( (t_j, \alpha_j) \) be two 2-tuple linguistic variables,

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

### 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

\[
S(R_p, R_q) = \sum_{i=1}^{N_d} r^p_d - r^q_d
\]

where, \( r^p_d \), \( r^q_d \) 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

\[
A(R_p) = \frac{1}{n-1} \sum_{q=1}^{N_p} S(R_p, R_q)
\]

where, \( N_p \) is the number of respondents.

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

\[
RA(R_p) = \frac{A(R_p)}{\sum_{R \in R} A(R)}
\]

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 \( w_p \) of the respondent \( R_p \) in knowledge area \( k_j \) can be got by

\[
w_p = \frac{y_1}{y_1 + y_2} + \frac{y_2}{y_1 + y_2} = RA(R_p)
\]

where, \( w_p \) 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

\[
VE_{ij} = \sum_{d=1}^{M_i} \sum_{p=1}^{N_p} w^p_d \times r^p_d \times \mu_D_d
\]

where, \( M_i \) is the number of documents that the expert \( E_i \) submitted, \( \mu_D_d \) 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 un-rated 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 \( U_i \in Inlink\_Neighbors(U_i) \),

\[
ST_{sl} = \max_{u_i \in \text{outlink\_neighbors}(U_i)} \left[ \min(\{r_{ij}, 0\}, ST_{il}) \right]
\]

Then, final estimated trust value is

\[
T_{sl} = \min[ST_{sl}, (t_{il}, 0)]
\]

where, \( ST_{sl} \) is the highest strength of trust path, the most reliable \( U_i \) and the \( ST_{il} \) is the rating given by the learner \( U_i \) to the document \( D_j \), \( N_j \) is the number of documents that the expert \( E_j \) submitted.

C. Background

Background means whether the learner has known the expertise in detail. The more times that they have contacted, the more familiarity with the expertise background. The degree of familiarity with the task may be obtained by

\[
F_{ij} = \frac{1}{\sum_{k=1}^{N_{ij}} \Delta^{-1}(r_{tk})} \tag{16}
\]

where, \( N_{ij} \) is the contacted times between the learner \( U_i \) and the expert \( E_j \), \( t_{tk} \) is the contact time, \( t_{now} \) is the present date, and \( \tau \) is a tunable parameter.

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

\[
B_{ij} = (b_{ij}, \sigma_{ij}) = \frac{\sum_{k=0}^{N_j} r_{tk}}{N_j \times \Delta} \tag{15}
\]

where, \( r_{tk} \) is the rating given by the learner \( U_i \) to the document \( D_j \), \( N_j \) is the number of documents that the expert \( E_j \) submitted.

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 \( E = \{E_1, E_2, \ldots, E_i\} \) be the set of experts, \( C = \{C_1 = \text{expertise}, C_2 = \text{trust}, C_3 = \text{background}, C_4 = \text{familiarity}\} \) be the set of criteria. Suppose \( R = \{r_{ij}\}_{m \times 4} \) is the derived evaluation matrix of the learner \( U_i \) for the expert \( E_k \). 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 \( i \neq j, w_i \geq w_j; w_i - w_j \geq \alpha_i, \alpha_i > 0; w_i - w_j \geq w_k - w_l, k \neq l \); \( w_j \geq \beta_i, \beta_i, 0 \leq \beta_i \leq 1 \); \( y_i \leq w_j \leq y_i + e_i, 0 \leq y_i < y_i + e_i \leq 1 \).

Step 1 With the function \( \Delta \), 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

\[
\begin{align*}
\text{r}^+ &= \{(r_{1,1}^+, a_{1,1}^+), (r_{2,1}^+, a_{2,1}^+), (r_{3,1}^+, a_{3,1}^+), (r_{4,1}^+, a_{4,1}^+) \} \\
\text{r}^- &= \{(r_{1,1}^-, a_{1,1}^-), (r_{2,1}^-, a_{2,1}^-), (r_{3,1}^-, a_{3,1}^-), (r_{4,1}^-, a_{4,1}^-) \}
\end{align*}
\]

For the linguistic values,

\[
\begin{align*}
(r_{ij}^+, a_{ij}^+) &= \max\{(r_{ij}, a_{ij}) | j = 1, 2, 3 \} \\
(r_{ij}^-, a_{ij}^-) &= \min\{(r_{ij}, a_{ij}) | j = 1, 2, 3 \}
\end{align*}
\]

For the numerical values,

\[
\begin{align*}
(r_{ij}^+, a_{ij}^+) &= \min\{r_{ij} | j = 4 \} \\
(r_{ij}^-, a_{ij}^-) &= \max\{r_{ij} | j = 4 \}
\end{align*}
\]

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.

\[
\begin{align*}
\text{min} r^+ &= \Delta \sum_{j=1}^{4} \Delta^{-1}(r_{ij}, a_{ij}) - \Delta^{-1}(r_{ij}^+, a_{ij}^+) | w_j \\
&+ \|r_{t4} - r_{t4}^+| w_4 \\
\text{max} r^- &= \Delta \sum_{j=1}^{4} \Delta^{-1}(r_{ij}, a_{ij}) - \Delta^{-1}(r_{ij}^-, a_{ij}^-) | w_j \\
&+ \|r_{t4} - r_{t4}^-| w_4 \\
\text{subject to:} w \in H, i = 1, 2, \ldots, m
\end{align*}
\]

Step 4 Relational degree of expert from PIS is obtained by

\[
\begin{align*}
\sum_{j=1}^{4} (2 \times \Delta^{-1}(r_{ij}, a_{ij}) + \Delta^{-1}(r_{ij}^+, a_{ij}^+)) | w_j \\
&+ (2 \times \|r_{t4} - r_{t4}^+| w_4 \\
\sum_{j=1}^{4} (2 \times \Delta^{-1}(r_{ij}, a_{ij}) + \Delta^{-1}(r_{ij}^-, a_{ij}^-)) | w_j \\
&+ (2 \times \|r_{t4} - r_{t4}^-| w_4 \\
\end{align*}
\]

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 \( S_T = \{S_0 = \text{Definitely Low (DL)}, S_1 = \text{Very Low (VL)}, S_2 = \text{Low (L)}, S_3 = \text{Average (A)}, S_4 = \text{High (H)}, S_5 = \text{Very High (VH)}, S_6 = \text{Definitely High (DH)} \} \) is given to express opinions. The submitted documents by experts and the classification information of documents are shown in Table 1.
The expertise of respondents is shown in Table 2.

<table>
<thead>
<tr>
<th>Respondents</th>
<th>JAVA</th>
<th>Oracle</th>
<th>JS</th>
<th>CSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>U1</td>
<td>VH</td>
<td>L</td>
<td>VH</td>
<td>L</td>
</tr>
<tr>
<td>U2</td>
<td>VL</td>
<td>A</td>
<td>VL</td>
<td>A</td>
</tr>
<tr>
<td>U3</td>
<td>VH</td>
<td>H</td>
<td>VH</td>
<td>H</td>
</tr>
<tr>
<td>U4</td>
<td>DH</td>
<td>VH</td>
<td>DH</td>
<td>A</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>E1</th>
<th>E2</th>
<th>E3</th>
<th>E4</th>
</tr>
</thead>
<tbody>
<tr>
<td>353</td>
<td>282</td>
<td>333</td>
<td>234</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Documents</th>
<th>JAVA</th>
<th>Oracle</th>
<th>JS</th>
<th>CSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>H</td>
<td>A</td>
<td>VH</td>
<td>DH</td>
</tr>
<tr>
<td>D2</td>
<td>H</td>
<td>VH</td>
<td>A</td>
<td>VL</td>
</tr>
<tr>
<td>D3</td>
<td>L</td>
<td>A</td>
<td>VL</td>
<td>A</td>
</tr>
<tr>
<td>D4</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>A</td>
</tr>
<tr>
<td>D5</td>
<td>A</td>
<td>DH</td>
<td>VH</td>
<td>L</td>
</tr>
<tr>
<td>D6</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>DH</td>
</tr>
<tr>
<td>D7</td>
<td>A</td>
<td>VL</td>
<td>VH</td>
<td>VL</td>
</tr>
<tr>
<td>D8</td>
<td>VL</td>
<td>VH</td>
<td>L</td>
<td>A</td>
</tr>
</tbody>
</table>

TABLE 1. SUBMITTED DOCUMENTS AND ITS CLASSIFICATION INFORMATION

TABLE 2. EXPERTISE OF RESPONDENTS

TABLE 3. CONTACT DATE

TABLE 4. READ AND RATED DOCUMENTS

TABLE 5. TRUST RATINGS

<table>
<thead>
<tr>
<th>Learners</th>
<th>E1</th>
<th>E2</th>
<th>E3</th>
<th>E4</th>
<th>U1</th>
<th>U2</th>
</tr>
</thead>
<tbody>
<tr>
<td>U0</td>
<td>DH</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>U1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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>
<thead>
<tr>
<th>Experts</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1</td>
<td>4.14</td>
<td>2.55</td>
<td>0.58</td>
<td>6</td>
</tr>
<tr>
<td>E2</td>
<td>2.44</td>
<td>1.46</td>
<td>0.42</td>
<td>4</td>
</tr>
<tr>
<td>E3</td>
<td>4.53</td>
<td>0.92</td>
<td>0.92</td>
<td>3</td>
</tr>
<tr>
<td>E4</td>
<td>3.25</td>
<td>0.00</td>
<td>0.33</td>
<td>2</td>
</tr>
</tbody>
</table>

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>
<thead>
<tr>
<th>Experts</th>
<th>Relational degrees</th>
<th>Rankings</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1</td>
<td>0.92</td>
<td>4</td>
</tr>
<tr>
<td>E2</td>
<td>0.33</td>
<td>2</td>
</tr>
<tr>
<td>E3</td>
<td>0.62</td>
<td>3</td>
</tr>
<tr>
<td>E4</td>
<td>0.16</td>
<td>1</td>
</tr>
</tbody>
</table>

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