Research on Ontology Perception Technology in Recommendation of Professional Books

Liu Li
Library of North China Electric Power University
1850508590@qq.com

Abstract: A book recommendation method in the paper is proposed by using the professional book ontology perception method. By establishing a model framework based on ontology technology, the ontology triggering mechanism of professional books is combined to obtain relevant information parameters and improve the higher recommendation accuracy, so as to more accurately calculate the regional weights and preferences of professional books, and complete the professional books and the recommendation process. Through the evaluation and testing of the theoretical methods, the feasibility and accuracy of the research content in the paper is proved, which contributes to the improvement of book recommendation service.

1. Introduction
With the advent of the era of informational digital libraries, the management of library books has focused more on the application and development of new technologies. At present, the research content of book management in the world mainly focuses on the effect of book recommendation. By using book comparison relationship to define and locate, more books related information content is integrated into the book ontology, which can result in the complexity of book recommendation research. Therefore, in the library management, how to obtain the relevant information of the book for analysis and make reasonable evaluation and recommendation becomes a difficult point in the field of book management research[1-2].

2. The establishment of professional book model based on ontology
The professional book ontology perception model in the paper is proposed, which can mainly consists of three parts: dynamic positioning module, knowledge base and perceptual trigger module, as shown in Figure 1.
In the professional book ontology perception model, the main function of the dynamic positioning module is to obtain various types of dynamic information and establish associated information attributes from it, add the basic information to the system log as the original information for registration, and provide the user location context to the sensing trigger module. The various functions in the perceptual trigger module are mainly embodied in the process of intelligently processing the information provided by the dynamic positioning module, which is triggered by the direct information triggering into the ontology related information triggering. The function of the knowledge base is to process the sensing trigger information module. The associated information is optimized and recommended for processing. In the recommendation module, the user is configured to build an ontology-based book preference model that includes preferences for user-specific attributes. The former builds a time-based forgetting pattern, which uses the user's recent interest to represent it. Secondly, the professional book recommendation effect will be fed back into the system log to form a regional group model to express the common preferences of readers who are similar to reading preferences for regional books[3-4].

3. Professional book recommendation method based on ontology sensing technology

3.1 Mapping of ontology-aware structures

The professional knowledge books in the library exist in a specific directory in the database, and the professional directory is used as the mapping area as the basic unit of the ontology aware structure, that is, any information triggered to the professional class directory and the root directory as a The original information trigger point, which will form a body-aware trigger within the profession to obtain an information variable. In other professional catalogs, if the final result is related to the ontology knowledge of this catalog, it will trigger the generation of relevant information and form an associated neural network organization structure, which will continuously map out the regional environment with the concept of the region, as shown in Figure 2[5].

Figure 1. Professional book ontology perception model
When the information relationship is in the mapping area, the perceptual-based book recommendation problem is converted into the book recommendation problem based on the ontology location area, that is, when the related information of the computer professional books and the educational professional books coincides, the temporary ontology sensing module "language" is the relative mapping area is generated as a relation. The overall structure will form a mapping area, the region set is: Region={region1, region2,...}, the concept set of the region region1 is Cregion1={c1, c2,..., cn}; the user set is U={u1, u2,..., un}; the concepts set of ontology O is C={c1,c2,...,cn}; the relationship set between concepts in ontology O is set as R; the book set is set as B={b1,b2,...,bn }[6-7].

3.2 Recommendations for professional books

According to the user's model of interest in professional books, the feature vector model is first established according to the initial information of the system, which includes user information, search information, and borrowing information. A more detailed information processing process can be performed after a certain characteristic phase is formed, so the feature H is taken as the basic unit, the phase unit is in the time T, and the feature set H= \{h_{t1},h_{t2},...,h_{tn}\}.

The user interest model relies on the preference information of professional books, through the user borrowing records and the user and system interaction log. According to the "class-instance" relationship between the leaf concept and the book in the construction ontology O, the concept set of the user preference can be directly obtained through the user borrowing the record and the book name in the interaction log. Formally, the user's short-term preference is marked as: 

\[ P_u = \{<c_1,w_1>, ..., <c_n,w_n>\} \]

The calculation of the concept weight affects the user's borrowing behavior and click viewing behavior, the proposed formula is:

\[ W_{ck} = \frac{1}{m1} \sum_{j=1}^{m1} \alpha_{ij}^k + \frac{1}{m2} \sum_{j=1}^{m2} \alpha_{ij}^k \cdot A_j \]

Where \( W_{ck} \) is the weight of the concept \( c_k \). \( m1 \) and \( m2 \) are the number of books viewed by borrowing and clicking of the user current time period, respectively. \( a_i^k, a_j^k \in [0,1] \) represents a professional book borrowed during an observation[8].

The concept of setting the borrowing/clicking book directly belongs to \( c_k \), then \( a_i^k, a_j^k = 1 \), otherwise \( a_i^k \) and \( a_j^k \) depend on the semantic similarity between the concept \( c_k \) and the concept of the book i/j. \( A_j \) is a constant to assess whether the user is really interested in clicking on a book, which is defined as:
\[ A_j = \{ \begin{cases} \text{readTime} \geq \text{Th} \\ 0 \text{ otherwise} \end{cases} \} \] (2)

Where readTime is calculated by the user when the book borrowing time (the value is sorted by the library borrowing data), and the threshold value Th is set as:

\[ \text{SimConcept}(c_i, c_k) = \frac{2 \times \text{depth}[NCT(c_i, c_k)]}{\text{depth}(c_i) + \text{depth}(c_k)} \] (3)

Where SimConcept \((c_i, c_k)\) is the similarity between the concepts \(c_i\) and \(c_k\); NCT \((c_i, c_k)\) is the common nearest ancestor of the concepts \(c_i\) and \(c_k\), and the depth \((c_i)\) is the depth of the concept \(c_i\) in the ontology[9].

After the user continuously updates the borrowing information, the user's short-term preference \(P_{st}\) is extended. The approach taken is based on the display semantic relationship with other directly related concepts in the ontology. The concept set of concept \(c_i\) directly related (including \(c_i\) itself) is defined as \(r(c_i) = \{c_{i1}, c_{i2}, \ldots, c_{ik}\}\), then for any concept \(c^k_i\), the concept weight calculation formula is:

\[ w(c^k_i) = \begin{cases} w(c_i) & \text{if } c^k_i = c_i \\ 0 & \text{otherwise} \end{cases} \] (4)

Where \(w(c^k_i)\) is the weight of the concept \(c^k_i\); \(w(r)\) denotes the known concept \(x\) in the user's initial preference concept set, and the \(r(x, y)\) relationship exists, then the user's preference probability for the concept \(y\) (here denoted as the semantic relationship weight). It should be noted that when \(r\) is an "instance relationship", \(w(r)\) takes a value of 0, which is the extension of the concept to the book instance. The extended concept set \(Ex\_concepts\) is obtained by extending the concept of the short-term preference \(P_{st}\) concept set={\(c_1, c_2, \ldots, c_n\)} in the user[10].

\[ Ex\_concepts = \bigcup_{c^k_i \in Econcepts} r(c_i) \] (5)

After obtaining the extended user short-term preference, a long-term preference concept set is formed. As the knowledge is continuously updated, the accuracy of the calculation will continue to increase.

3.3 Building a regional group model to expand recommendations

Depending on the close relevance of the book's knowledge domain, users with borrowing records in the same book profession can be considered as users with similar reading interests for books in the region. The regional group model is constructed to improve the diversity of recommendations by merging different knowledge domains to approximate the user's long-term preference for knowledge domain books.

The mapping structure between the ontology concept \(C\) and the Region, the user's long-term preference \(P\) is divided into multiple semantic divisions. Each semantic partition corresponds to a regional cluster, which represent a subset of user interests. Formally, the user's long-term preference in the region regionI can be labeled as: \(P_{region} = \{C_{region}, W_{region}\\}. In order to reduce the computational complexity and ensure the interest of most group members, the concept set selection in the regional group model adopts a conceptual feature selection strategy based on the smallest total distance. We can define a concept dictionary for each region \(Lexion_{region} = \{c_1, c_2, \ldots, c_k\} \), which contains all the concepts of the region. Then, the \(W_{region}\) of the cluster user can be uniformly expressed as a vector \(V = \{f_1, f_2, \ldots, f_i\}\), where \(f_i\) is a corresponding preference weight, and the value is defined as: if \(c_i\) is included in \(C_{region}\), then \(f_i = 1\); Otherwise \(f_i = 0\). The result of combining the same dimension element \(a_i (i = 1, 2, \ldots, N, a_i \in \{0,1\})\) in N user preference vectors \(V\) is defined as:

\[ \text{MergeResult} = \begin{cases} \sum_{a_i} \text{count}_i & \text{count}_i \geq \text{count}_0 \\ \text{count}_0 & \text{count}_i > \text{count}_0 \end{cases} \] (6)

Where \(\text{count}_1\) is the total number of occurrences of '1' for the dimension element in the N user preference vectors, and \(\text{count}_0\) is the total number of occurrences of '0'.

4
For each conceptual feature, the formula (6) is used for merging. After eliminating the concept that the combined result is 0, the weighting formula for each conceptual feature selected is:

\[ Weight_i = \sum_{j=1}^{n} w_{ij} \]  

(7)

Where Weight\(_i\) is the weight assigned to c\(_i\) in the regional group model; w\(_{ij}\) is the weight of c\(_i\) in the j-th user; n is the number of concepts selected.

The regional group model is built based on the user's long-term preferences. After updating the long-term user preferences, the zone group model for the affected area also needs to be updated. Similarly, this paper also sets the update period of the regional group model to T. That is, every time period T. Based on the borrowing record and the interaction log, the affected areas are searched and the regional group model of these areas is reconstructed.

4. Experiment and analysis

4.1 Experimental design

(1) The experiment is based on the library's collection of professional books, which are set to "A, B, C, D, E, F, G, H, I" books, A, B, C professional books as the main In part, other types of books are presented as relevant factors that affect the ontology triggering.

(2) Using the annual cycle in the library as a time chain, the user's data in one cycle is analyzed. When the user logs into the library account and queries the book, the user's system interaction log is obtained by recording the book clicked and viewed by the user, and other data samples (user borrowing records, personal information, and book information) are directly extracted from the library system data.

(3) Parameter setting

According to the set weight value of professional books, when the user's short-term preference is semantically extended, the values of different semantic relations is \( r \in R \) and \( w(r) \) are shown in Table 1.

| R          | W(r) | W(r\(^{-1}\)) |
|------------|------|---------------|
| contains   | 0.61 | 0.59          |
| instanceOf | 0    | 0             |
| composeOf  | 0.54 | 0.55          |
| similarTo  | 0.65 | 0.61          |
| subclassOf | 1    | 0.28          |

The diversity of the recommended model recommendation books is to verify the diversity of the recommended books in the model, and to compare the preferences of the combined regional groups using the diversity of the collections. The collection diversity of books is defined as the average dissimilarity between all books in the recommendation list.

At the time of evaluation, the reading records of the last N time ranges are selected for each user as the evaluation criteria (according to the set T value, each time range is 30 days), and the history before this time range is used for the construction of the user preference model and the regional group model. Through the continuous extension of each user's time range as a professional book recommendation effect, as shown in Figure 3:
Perceptual recommendation
Ontology trigger recommendation
Comparison of recommended books in the experimental period

As described in Figure 3, professional book recommendation methods based on ontology-aware triggering techniques perform better than traditional book recommendation methods that the ontology-aware triggering method improves the accuracy of 27.65% on average, which indicates that the ontology-aware triggering method proposed by the paper can improve the accuracy of recommendation. Although compared with the traditional method, the recommendation effect on the original information material is low in the initial stage of user registration, but the shape of the model is more perfect in a short period of time, and the recommended effect is continuously improved. In the future, in order to further improve the recommendation effect, the paper will discuss the value of the weight coefficient.

4.2 Evaluation of the validity of the recommendation model

In order to verify the recommended accuracy of the proposed method, the proposed method is compared with the existing comprehensive collaborative filtering technology and the recommended method of ontology. The experimental setup compares the recommended accuracy and relevance of the ontology-aware trigger book. The evaluation indicators use the precision of the widely used.

\[ \text{Precision} = \frac{|\text{RB} \cap \text{PB}|}{|\text{PB}|} \quad (8) \]

Where The RB marks the book recommended to the user, the PB marks the book preferred by the user, and the |RB| indicates the cardinality of the set RB. The experimental results are shown in Figures 4:
It can be found that the similarity between the actual effect of the ontology-aware triggering method proposed in this paper and the model comparison reference is higher, which verifies the unification of the theoretical and practical effects of the method. Although the effect trend in the model comparison test is higher than the actual application effect, it is mainly because the accuracy of the parameter weight calculation method needs to be improved, and the user's search for other types of professional books in the process of borrowing books will affect the actual effect., so the verification effect of this method can already meet the needs of users.

5. Conclusions

A professional book recommendation model in the paper is proposed based on ontology-aware triggering, which integrates the user's basic information, knowledge type preference, long-term knowledge accumulation and related types of knowledge acquisition, so that books can be recommended according to the knowledge of the user's needs. The experimental results show that the model can improve the precision of the recommendation. There are some shortcomings in the technique of the value of the weight coefficient and the interference of the exclusive information, but the experimental results of the theory and method prove the feasibility. In the future research, we will continue to increase the depth of research to further improve the quality of recommendations and meet the needs of users’ books.

References:
[1] Optimizing ontology alignments through a Memetic Algorithm using both MatchFmeasure and Unanimous Improvement Ratio[J]. Xingsi Xue,Yuping Wang. Artificial Intelligence . 2015
[2] Visual support for work assignment in process-aware information systems: Framework formalisation and implementation[J]. Massimiliano de Leoni,Michael Adams,Wil M.P. van der Aalst,Arthur H.M. ter Hofstede. Decision Support Systems. 2012 (1)
[3] Discrete particle swarm optimisation for ontology alignment[J]. Jürgen Bock,Jan Hettenhausen. Information Sciences . 2010
[4] Research on domain ontology in different granulations based on concept lattice[J]. Xiangping Kang,Deyu Li,Suge Wang. Knowledge-Based Systems . 2011
[5] Model-driven design and development of semantic Web service applications[J]. Marco Brambilla,Stefano Ceri,Federico Michele Facca,Irene Celino,Dario Cerizza,Emanuele Della Valle. ACM Transactions on Internet Technology (TOIT) . 2007 (1)
[6] An Overview of Semantic Web Services Composition Approaches[J]. Yasmine Charif,Nicolas Sabouret. Electronic Notes in Theoretical Computer Science . 2006 (1)
[7] Comparison among five evolutionary-based optimization algorithms[J]. Emad Elbeltagi,Tarek
Hegazy, Donald Grierson. Advanced Engineering Informatics. 2005 (1)

[8] HTN planning for Web Service composition using SHOP2[J]. Evren Sirin, Bijan Parsia, Dan Wu, James Hendler, Dana Nau. Web Semantics: Science, Services and Agents on the World Wide Web. 2004 (4)

[9] Resource constraints analysis of workflow specifications[J]. Hongchen Li, Yun Yang, T.Y. Chen. The Journal of Systems & Software. 2003 (2)

[10] Learning to Match the Schemas of Data Sources: A Multistrategy Approach[J]. AnHai Doan, Pedro Domingos, Alon Halevy. Machine Learning. 2003 (3)