ISBN 978-981-11-3671-9 Proceedings of 2017 the 7th International Workshop on Computer Science and Engineering (WCSE 2017) Beijing, 25-27 June, 2017, pp. 675-681 doi: 10.18178/wcse.2017.06.118

Petroleum, Oil, and Lubricants Information Acquisition and Processing in Hall for Workshop of Meta-Synthetic Engineering

Bixin Li¹, Jin Huang² and Yan Lei¹⁺

¹ Logistical Engineering University ² No.285 hospital

Abstract. Being the first and important step for organizing POL (Petroleum, Oil, and Lubricants, POL) service, POL information acquisition and processing is an indispensable part of POL support service. It offers POL service information, which is critical to understand the intention of POL support service, and make timely and accurate judgment on the organization of POL support service. Without timely and accurate POL information acquisition and processing, it leads to the organization of POL support service in chaos and disorder. However, in reality, combat system has the complex characteristics of the sudden and variability in time domain, the crisscross in spatial domain, the congestion and overlap in frequency domain, and the uneven density in energy domain, posing a challenge to POL information acquisition approach based on Hall for Workshop of Meta-Synthetic Engineering. With system theory as the core of our approach, the proposed approach decomposes the complexity of information acquisition and processing, and incorporates human computer interaction to present a framework featuring the combination of qualitativeness and quantitativeness. This work presents a new perspective on POL information acquisition and processing, and suggests fresh directions of research on an extensively studied topic.

Keywords: information acquisition, meta-synthetic engineering, POL service, system theory.

1. Information Acquisition and Processing System in Hall for Workshop of Meta-Synthetic Engineering

Information acquisition and processing system in Hall for Workshop of Meta-Synthetic Engineering (HWMSE)^[1] consists of workshop subjects, information support platform, network environment, application system, etc. The following will give an introduction of each part.

1.1. Workshop subjects

Acquiring timely, accurate and reliable POL information for POL support is on the premise that we apply HWMSE to study and discuss POL support issues in military actions. In the workshop, it involves many subjects showing different behavioural features. Workshop subjects mainly contain POL divisions of combat troops, POL support units, institutes of scientific POL research and training, institutes of local force mobilization^[2].

1.2. Information support platform

In HWMSE, after specifying workshop subjects, it requires computer technology to process big POL support data and information for the workshop. HWMSE information support platform^[3] offers database (relevant data and information, and database management system), text library (POL support plan, proposal, and paperwork), method library (the methods to solve various POL support issues), model library (parameters, models, algorithms and model library management system), human-computer interaction

⁺⁺ Corresponding author. Bixin Li. Tel.: + 8613627691722.

E-mail address: libixin000000@163.com.

technology (text, audio, natural language input, etc.), access servers and system terminals for realizing distributed structure in different places.

1.3. Network environment

According to the requirement of HWMSE, network structure can be classified into two types: on-site meeting structure and distributed structure^[4]. In case of on-site meeting structure, all workshop staff members should gather at the pre-assigned site. Teaching workshop system uses easy implemented local area network with simple connection structure. In addition, in local area network, we can also use more advanced high-speed network to ensure the demand of HWMSE as shown in Fig.1.

When workshop staff members cannot gather at a pre-assigned site and their locations are distrusted in different geographical areas, we should adopt distributed network structure. In this case, we use wide area network to connect workshop staff members and workshop resource. The whole system is implemented by distributed computer network structure, conducting HWMSE in different locations and areas as shown in Fig. 2.



Fig. 1: Network structure of on-site meeting.

Fig. 2: Distributed network structure.

1.4. Application system

HWMSE application system^[5] makes the best use of different workshop means to unearth and inspire workshop staff member's experience and intelligence, realizing timely, reliable, real-time, fast POL support information acquisition. First of all, HWMSE application system should provide easy-use workshop tools, such as text, pictures, audio, video, file sharing, online communication, electronic whiteboard, screen sharing, remote collaboration, etc. Secondly, in the distributed HWMSE, in order to achieve the goal of smooth communication and group decision among workshop staff members, application system should use various group workshop modes such as anonymous mechanism. Thirdly, it should offer POL support workshop classification tool, modeling methods and model analysis methods for various classifications. In addition, model analysis methods can assemble existing models from model library in demand. Fourthly, HWMSE application system should have visualization function, transforming text into intuitive, comprehensive, easy-to-remember visual graphics and thus assisting workshop staff members in solving the issues of group information capturing and memorizing caused by massive information. Finally, application system should have different solutions on measuring the importance of information to make the final information choice decision.

2. Dynamic Self-learning POL Support Information Acquisition Method

Self-learning POL support information acquisition is based on HWMSE by repeating self-learning process to dynamically acquire POL support information as shown in Fig. 3. Specifically, we use natural

language recognition technology to automatically transform the workshop content discussed by workshop staff members, in the theme of unit action POL support, into text. Then, we import POL support feature dictionary. Based on the dictionary, we adopt topic extraction and judgment technology to extract current topic of the workshop from the generated text and judge topic change situation. After that, we use the key words of the current topic as retrieval words, and input them into system for retrieval. Finally, we recommend the retrieval information to workshop staff members. The above process will be repeated until workshop staff members acquire satisfactory POL support information.



Fig. 3: Framework of self-learning information acquisition.

3. POL Support Core Vocabulary Generation Module

In HWMSE, POL support core vocabulary generation module is based on POL support feature vocabulary, providing the vocabulary of POL support and its related majors^[6]. The main process is shown as follows:

Before workshop, according to the target problems in this workshop, workshop host chooses a few key words, that best match the target problems, as the feature key words. Based on these feature key words, the module offers related workshop material and generates POL support feature vocabulary by self-learning using FWB (FieldWords-Bootstrapping, FWB) model.

3.1. POL Support Feature Vocabulary

POL support feature vocabulary is a semantic knowledge set that is accumulated from the work, training, scientific research, practice in the field of POL support. The set has a distinct technicality and special property.

3.2. Generation process of POL support feature vocabulary

It is costly to manually construct POL support feature vocabulary. Furthermore, at the initial stage, manual construction usually cannot systematically and comprehensively cover the feature vocabulary from POL support workshop. It is necessary for automatically generating POL support feature vocabulary. Thus, we present a feature vocabulary acquisition method using FWB model. This method uses computer learning technology to process, recognize, judge, filter, search massive POL service data, summarizing and producing POL support feature vocabulary. The process using FWB model to generate feature vocabulary is described as below (see Fig. 4):

(1) Initialize the POL support seed words;

(2) Based on the initial POL support seed words, this step utilizes FWB model to generate new POL support words and add them into the current feature vocabulary. Meanwhile, update POL support seed words;

(3) If there is no new POL support feature word generated, the learning will terminate;

(4) Initialize new seed words in the current POL support feature vocabulary. Repeat the above steps to construct POL support feature vocabulary until satisfy the threshold.



Fig. 4: Acquisition process of Domain vocabulary in HWMSE.

4. Topic Extraction, Judge and Update module

In the HWMSE discussion of POL support organization and plan, we use the input means (such as keyboard, handwriting, audio, natural language, etc.) to recognize the conversation of workshop staff as the text displaying on each user terminal^[7]. First, Topic extraction, judge and update module accept the text showing workshop content from the workshop system. The next, based on POL support feature vocabulary, this module uses weighted analysis method to process the workshop text and extract key words from the text as the current topic. Moreover, it computes the similarity between the extracted topic and all existing topics. Finally, it judges whether this topic is a new one in terms of similarity. If the extracted topic is a new one, this module will input the topic key words into the retrieval system and recommend the retrieval information to workshop staff; otherwise, keep current topic and recommendation information, and proceed to the next workshop stage.

4.1. Topic extraction in POL support workshop

In POL support workshop, we set a time interval T, we divide the workshop content at the time interval T into paragraphs (using *ICTCLAS* tool), delete and save these paragraphs to generate bag of words format file. Based on bag of words, we conduct the workshop topic extraction.

We introduce vector space model to represent POL support workshop text, and apply tf-idf method to compute the feature weight of each word. The process of algorithm is shown as follows:

(1) Self-learning feature set

Assume the set of POL support feature words generated at the time interval T is FeatureSet, namely POL support feature set.

$$FeatureSet = (feature_1, feature_2, \cdots, feature_r)$$
(1)

In Equation 4.1, z represents the number of feature words in FeatureSet_t. Feature_i denotes the i-th feature word, which is atomic and sliced by *ICTCLAS* tool. Feature_i is the basic and only element in vector space model.

As the proceeding of POL support workshop, new workshop content will produce new feature words, that is, the feature set will be continuously extended, updated and enriched. Thus, this set has the self-learning ability.

$$FeatureSet = (feature_1, feature_2, \cdots, feature_7)$$
⁽²⁾

(2) Feature word computation

Let R_t is the workshop support workshop text at the time interval t. R_t is a vector of FeatureSet_t:

$$R_{i} \rightarrow (weight(R, feature_{i}), weight(R, feature_{i}), \cdots, weight(R, feature_{i}))$$
 (3)

Where, weighe(R,feature_i) is the weight of Feature_i in workshop text R_t . The weight computation equation in POL support workshop is shown as follows:

$$weight(R, feature_i) = \frac{(\lg(tf(R, feature_i)) + 1)\lg(Z / df(feature_i)))}{\sqrt{\sum_{i=1}^{z} [(\lg(tf(R, feature_i)) + 1)\lg(Z / df(feature_i))]^2}}$$
(4)

Where, the e is the display frequency of Feature_i in R. Z is the number of POL support workshop text at the time interval t. $df(feature_i)$ is number of workshop text containing Feature_i.

If Feature_i displayed in the previous workshop text, we should accumulate its displaying number, that is, $df_t(feature_i) = df_{t-1}(feature_i) + 1$.

(3) Feature word output

After the computation of weights, we output the feature words in descending order of weight as the current topic.

$$topic = ((topicword_i, weight_i)), k = 1, 2, \cdots, n$$
(5)

According to the requirement, we can set the n value. The lower value of k is, the higher weight of topicword_k is. A feature word with a higher weight means that it is more reliable to represent the current POL support workshop topic.

4.2. POL Support Workshop Topic Judge and Update

We adopt Single-Pass clustering model to judge and update current workshop topic. Single-Pass model has two types: s-c and s-s type. s-c type compares the whole previous workshop text with the current text, and judge their core matching degree to realize the update of workshop topic. In contrast, s-s type compares each previous workshop text with the current text, and judge whether to update the workshop topic according to each matching result. The difference of two types is similar to that of topic judging detailed and estimated methods.

Let use s-s type of Single-Pass model as an example:

A: compute the similarity of current text R_t and previous text R_j :simi $(R_t, R_j) = cos(R_t, R_j)$

B: obtain text R_{max} , then we have the simi(R_t , R_{max})=max simi(R_t , R_j).

C: when $simi(R_t, R_{max})$ >TopicThreshold, there is no Topic_{new} generated, where TopicThreshold is the initial threshold. R_t is added to the cluster R_{max} , waiting for the next judge.

D: when $simi(R_t, R_{max})$ <TopicThreshold, there is $Topic_{new}$ generated, where TopicThreshold is the initial threshold. We construct a cluster R_t , select several feature words from $Topic_{new}$ in terms of weights, and search them in the system.

4.3. Text Analysis Model of the Combination of Feature Words and Regular Words

Based on POL support feature vocabulary, we utilize self-learning methods to acquire POL support feature words. With these words, we analyze POL support workshop text to improve the extraction ability of core content (POL support information) and realize accurate track of workshop topic.

The principle is that we analyze the matching degree of feature terms in the workshop text and POL support feature words. If a feature term matches a POL support feature word, we will increase the weight of this feature term and update the weight(d,term_i) value in tf-idf model using the following equation.

$$weight(d, term_i) = tf - idf(d, term_i)(1+\alpha)$$
(6)

Where, α is the additional coefficient. A term with a higher value of α means that it is more important.

5. POL Support Information Retrieval and Recommendation Module

5.1. POL support information retrieval

We input the new feature words of $Topic_{new}$ into the workshop system for retrieval, and retrieve some POL support information related to feature words.

The retrieval results have a strong correlation with current workshop topic, covering and supporting the useful information of current workshop topic. If retrieval results have many items, it requires workshop staff to select several representative ones. To ensure no loss of important POL support information, we organize the workshop staff as a collaborative team, and establish feedback mechanism. In this way, we solve the selection problem in retrieval information by using collaborative mode.

5.2. POL support information recommendation

Based on information retrieval results using the key terms, a workshop member can choose part of results to read and evaluate the importance of this result according his or her professional experience. Meanwhile, the workshop system will record the reading situation and importance evaluation of the workshop member, and summarize the evaluation of each member as an importance value according to their authority degrees. After that, the system will use reorder the retrieval results in descending order of the importance values, and recommend the high-ranking results to those workshop members who do not read the retrieval results. With the complementary information of member-to-member and member-to-machine, it forms a well self-learning mechanism to implement the acquisition of POL support critical information.

Recommendation model of POL support information:

(1) Input the number and importance values of retrieval information items;

(2) x_i is the importance value of the i-th retrieval information item (see Equation 5.1);

$$x_i = \sum_{k=1}^n \lambda_k \theta_k, \quad i = 1, 2, \cdots, M$$
(7)

Where, M is the number of retrieval information items, n is the number of workshop staff members, λ_k is the important value of the i-th retrieval information item evaluated by the k-th workshop member, and θ_k is the authority degree of the k-th workshop member.

(3) Reorder retrieval information by computing the importance values;

According to the weight of each link, the system reorders each link in terms of importance values. The computation of the a-th link is shown as below:

$$l_a = \arg\max\{x_i\}, i = 1, 2, \cdots, M$$

(4) Filter information;

According to the importance value of each link, we choose the top m links (l_1, l_2, \dots, l_m) . Where, m=1,2,3,....

(5) Recommend the important POL support information l_1, l_2, \dots, l_m to workshop staff.

6. Conclusion

Taking HWMSE as the base, this paper proposes a dynamic self-learning POL support information acquisition approach. This approach successfully transform the process of POL support workshop into a continuous self-learning and knowledge-accumulating process.Our approach is effective to address many problems in POL support information acquisition, such as many information sources, massive volume, various formats, coexistence of authenticity ,processing uncertainty, etc. Our work presents a new perspective on acquiring POL information accurately and reliably.

However, our approach still has some limitations on core vocabulary acquisition, potentially affecting the accuracy of core vocabulary acquisition and thus causing topic update and information recommendation unsatisfactory. Our future work plans to solve this problem via the three aspects; 1)gather and improve POL support data; 2)refine FWB feature words acquisition model; 3)improve quality control of feature works extracted by acquisition model to ensure the accuracy of information acquisition.

7. References

- [1] C. L. Han, L. Xu, and W. Chen, "Research on workshop hall for meta-synthetic engineering of course construction "office software advanced applications", *Computer Era*. 2014, 7: 62-64,.
- [2] Y. Bi, X. Lai, Y. Zeng, and F. Cheng, "Research on workshop hall for meta-synthetic engineering of emergency handling," *China CIO News*. 2016,**8**:136-138.
- [3] T. Yu, "Research on Hall for Workshop of Meta-Synthetic Engineering of Equipment Support System," *Ordnance Industry Automation*.2005, **24** (6): 65-65.
- [4] B. Wang, X. Chang, and Y. Li, "Concept Exploration of Hall for Workshop of Meta Synthesis Engineering for Space Military System," *Journal of the Academy of Equipment Command & Technology*.2003, **14** (1): 46-49.
- [5] X. Li, and G. Ma, "BIM Technology for Hall for Workshop of Meta Synthesis Engineering," *Real Estate Information of China*.2015, **12**: 23-25.
- [6] B. Yi, "Military Theory Innovation in Hall for Workshop of Meta Synthesis Engineering," *Military Operations Research and Systems Engineering*.2011, **41** (4): 15-19.
- [7] D. Li, "Application Prospect of Hall for Workshop of Meta-synthetic Engineering," *Co-operative Economy & Science*.2010, 1: 50-51.