Scaffold Learners’ Qualitative Browsing of Major Topics in Learning Forum

Chih-Kai Chang
Department of Information Management
Da-Yeh University, Taiwan
chihkai@mail.dyu.edu.tw

Abstract

After the system database contains an accumulation of discussion messages with date tagged from different classes within a course, an instructor may want to have an overview of learners’ discussions, such as the top ten hot topics, without reading every message in the learning forum. Hence, it is necessary to model the relationships between discussion messages during a specific time interval for showing the major discussion topics. Consequently, an instructor can use the information to refine his/her teaching strategies in the classroom or virtual classroom. This paper proposes a statistical model of word usage, which originates from information retrieval technology, to group learners’ discussion messages by concept similarity into time dependent topics. Consequently, the methodology can generate comprehensible topics for an instructor to sketch out what learners were discussing by concept-based message groups and when the important messages were posted by a temporal interface.

1: Introduction

To motivate learners’ reflection, an instructor may retrieve earlier discussion messages as the topic of a discussion activity. On the basis of this observation, we developed a concept-similar mechanism to reuse the discussion messages of a web-based learning forum as recommended readings to scaffold learners by filtering, encoding, and indexing previous learners’ discussion portfolios [1]. On the other hand, the concept-similar clustering mechanism can support an instructor to summarize learners’ discussions without reading every discussion message in the learning forum.

Unfortunately, most of the existing asynchronized discussion styles are thread-based on the Web. This means that an initial message was first posted and subsequently followed by messages on the same subject. Discussion messages can be presented in different ways, for instance, by time, author, or subject order. However, discussion messages of a learning topic may distribute in many subjects with diverse date tag. An instructor can not use those orders to determine what the most significant topics are because he/she wants to find out the topics for a learning reflection activity. Furthermore, an instructor may be only interested in the “hot” topics related to the course. Unfortunately, incalculable numbers of discussion messages are generated and stored in asynchronous learning forums daily. Hence, it is necessary to find a method of managing discussion messages supporting a learner or an instructor to get summarized information to do something without excessive effort.

This paper presents a method for retrieving major topics in discussion messages and constructing a temporal interface of major topics in a web-based forum by the support of prior learners’ discussion portfolios in a distance learning system. First, a concept-based mapping method is used to reorganize prior learners' discussion portfolios so that learners can benefit from them without reading every accumulated message. Then, the temporal distribution of a concept is demonstrated to scaffold a learner or an instructor to outline the discussion process. Finally, an instructor can manipulate time interval to know what and when the most important topics were contained in prior learners' discussion messages.

2: Discussion Messages Grouping

The goal of concept-based grouping is to reorganize learners’ discussion messages by the degree of the concept similarity between each message. The basic problem is how to determine the relevance of a discussion message to a specific concept. Once the discussion messages are retrieved with respect to a learning concept, they can be ranked according to a relevance score. To determine the relevance score of discussion messages to a concept category, there should be a ranking method that is based on conceptual similarity, rather than on words shared between a message and concept descriptions. This method cannot assume that there is any shared vocabulary between the concept category and the discussion messages. Yang and Chute reported how to solve the problem by singular value decomposition (SVD) method [2].

The intuitive meaning of the solution of the problem is to find a matrix that can be used to measure the “distance” between a concept category and every discussion message in the learning forum. Because matrix is derived from discussion messages of some classes within the same course,
it is trained to assign properly a message to a concept. Hence, the system can reorganize the discussion messages according to the concept similarity to show the major topics during a specified cover period. Then, an instructor can outline the discussion process by reorganizing the discussion messages into a concept-based style.

3: Browsing of Major Topics

The process of browsing major topics is similar to assign messages to proper concept category by the degree of concept relevance. First, a group of discussion messages is selected during a specific time interval. For instance, the discussion messages with timestamp less then or equal to twenty days are selected as input data. However, the number of discussion messages may be too large to read if the duration of the time interval is too long. Consequently, it is necessary to classify concept-related messages into a same concept category. In other words, the system can use concepts instead of discussion messages by ranking the relevance for messages to a concept.

However, the number of concept category may be too large to serve as a “major” topic during a specific time interval. Hence, we tried to order concepts appeared during the time interval and selected top ranked concepts as “major” topic in the learning forum. To order concepts, the following attributes of a concept are compared in sequence:

1. Number of learners involved in the discussion messages of a concept category.
2. Maximum time span, which is the time difference between the latest and earliest messages of a concept.
3. Number of discussion messages in a concept category.

Therefore, the top concepts of the time interval were of course the “major” topic in the learning forum.

Thus the system is ready to construct the temporal interface for demonstrating the distribution of major topics in the learning forum. Assume that an instructor desires four major topics displayed on the temporal interface. The top four concepts, which are denoted as A, B, C, and D, are selected as major topics. Before displaying the top four concepts on temporal interface, the system should compute four attributes of a concept, that is minimum, first quartile, median, third quartile, and maximum timestamp. The minimum timestamp indicates the earliest discussion message of a concept in the learning forum during the specific time interval. The first quartile, median, and third quartile timestamp mean that there was respectively 25%, 50%, and 75% of discussion messages posted before this timestamp. The maximum timestamp indicates the latest discussion message of a concept in the learning forum during the specific time interval. An example is shown in Figure 1.

<table>
<thead>
<tr>
<th>Concept</th>
<th>Start</th>
<th>First Quartile</th>
<th>Third Quartile</th>
<th>Last</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2003/2/6</td>
<td>2003/2/13</td>
<td>2003/2/18</td>
<td>2003/2/19</td>
<td>09d</td>
</tr>
<tr>
<td>B</td>
<td>2003/2/7</td>
<td>2003/2/15</td>
<td>2003/2/19</td>
<td>2003/2/20</td>
<td>12d</td>
</tr>
<tr>
<td>C</td>
<td>2003/2/10</td>
<td>2003/2/16</td>
<td>2003/2/20</td>
<td>2003/2/21</td>
<td>15d</td>
</tr>
<tr>
<td>D</td>
<td>2003/2/12</td>
<td>2003/2/19</td>
<td>2003/2/23</td>
<td>2003/2/24</td>
<td>16d</td>
</tr>
</tbody>
</table>

Figure 1. The temporal interface of major topics in a learning forum.

4: Conclusion

This report has proposed a methodology for reorganizing discussion messages by a concept-based method. If an instructor’s provision of a web-based, asynchronized forum for his or her course becomes continual, a “concept-based” text retrieval methodology to quickly search topics and scaffold learners by prior learners’ learning experiences could be constructed. In other words, the proposed methodology can overcome the learners’ vocabulary gap problem, which is the weak point in a general query system. Consequently, this research first groups the discussion messages into concept categories by a dominant, classic method in information retrieval. Hence, we can retrieve major discussion topics as the summarized information to an instructor.

Then, the temporal interface of a learning forum can not only illustrate how many messages were posted during a time interval, but can also demonstrate what learners were discussing during that period. Hence, an instructor or a learner can qualitatively browse what learners were discussing and when the important messages were posted.

5: Acknowledgements

This research was supported by the National Science Council of Taiwan under contract number NSC 91-2520-S-212-001.

6: References
