The effectiveness of an Intelligent Annotation Sharing System on e-learning

Chun-Hsiung Lee*, Gwo-Guang Lee, Yungho Leu
Department of Information Management, National Taiwan University of Science and Technology, Taipei 106, Taiwan

A R T I C L E   I N F O
Keywords:
e-Learning
Annotation system
Knowledge sharing
Scaffolding theory

A B S T R A C T
Reading is a very important part in learning process. When reading the teaching materials of textbooks in a traditional way, students usually underline the main points and take notes to help memorizing, thinking and understanding the contents of the teaching materials. With the progress of network technology, e-learning has gradually become a new learning trend. However, the digital e-teaching materials of e-learning are always the texts that cannot be changed by students as an easier reading format.

In this paper, we propose an algorithm named Expert Keywords Annotation Alignment Algorithm (EKAAA) and based on which we have developed an Intelligent Annotation Sharing System (IASS) as an auxiliary tool for students to read the e-teaching materials. Based on the cluster to which a student belongs, the annotation sharing system adaptively provides the student a suitable sharing model. The models serve as a "scaffolding" to guide the students' learning, intending to achieve the purposes of auxiliary learning and knowledge sharing. Finally, we use statistics to analyze the effectiveness of the Intelligent Annotation Sharing System on e-learning.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

Traditionally, when we are reading the teaching materials of textbooks, we usually have the habit of underlining and notetaking to help ourselves with memorizing, thinking and understanding the contents of the teaching materials. However, with the progress of network technology, e-learning has gradually become a new learning trend. Although the acquisition, transmission and storage of electronic documents enjoy obvious advantages when compared with the traditional paper documents, students find it unavailable to conduct any annotation act on the electronic documents while reading them. However, annotation is a traditional learning strategy commonly used by general students. Therefore, the reading pattern with annotation unavailability is the greatest obstacle to e-learning.

Although some websites provide students with single-sided annotation act, such an interactive pattern is of limited help to students. This is because the students of low-score cluster with poorer learning ability are always incapable to rearrange the main points, the contents they have learned and the notes, and the students of high-score cluster are unable to share with other students the annotation process of their reading. Therefore, the students of low-score cluster cannot acquire the knowledge of the students of high-score cluster through the knowledge sharing mechanism.

In view of this, the study attempts to develop an Intelligent Annotation Sharing System (IASS) as an auxiliary tool for students to read the e-teaching materials. According to different clusters of students, adaptive annotation sharing patterns are provided to them for learning and sharing knowledge by themselves, hoping to achieve the purpose of adaptive learning.

Based on the above phenomena, the study is undertaken with the following aims:

(1) Scaffolding theory is employed as the foundation to construct an Intelligent Annotation Sharing System (IASS) as an auxiliary tool and adaptive sharing platform for students to read the teaching materials.

(2) Expert Keywords Annotation Alignment Algorithm (EKAAA) is proposed to examine whether the annotation contents of students contain the keywords selected by experts, and evaluate the annotation act of students in a more objective and reasonable way.

(3) Taking EKAAA as the foundation and by means of Data Mining, the study establishes adaptive Annotation Patterns, and recommends intelligent annotation sharing model (IASM) to different clusters of students for reference.

(4) Statistical methods are employed to analyze different clusters of students. After the students have been guided by IASM, the paper studies whether it is helpful to enhance the learning effects of students.

2. Literature review

The study develops the IASS of digital learning platform and employs constructivism as the theoretical foundation, intending
to let students use the annotation tool to attach personal reading process to systematic rearrangement, turn the teaching materials to be useful knowledge, and share the knowledge among different clusters of students to achieve the purpose of knowledge sharing.

Constructivism is a theory of knowledge, as well as a theory of cognitive learning. The implication it suggests is to transfer knowledge from the instructor to the student. It emphasizes that knowledge is an active construction of students, instead of the passive acceptance or absorption of information by students (Chang, xxxx).

The so-called “annotation” refers to the special signs made on the contents of teaching materials. When a reader is undertaking a reading activity, if he/she puts some meaningful annotated signs on the book, it will be helpful to the subsequent readers (Marshall, 1997). Thus, it is very important to find out main points and retrieve useful information from a large amount of reading materials to help reading (MacLellan, 1997).

Why do students have such “annotation” act in the learning process? Slotte et al. (2003) think that there are two main reasons: one is that students think that the process of notetaking can help them learn, and the other is that students think that the “notes” produced from annotation act will be very useful to their reviews in the later days. Hidi and Anderson (1996) indicate that putting summarized signs on the representative information and contents of an essay is of great help to the comprehension of the essay. Focusing on annotation act, Nokelainen, Kurhila, Miettinen, Floreen, and Tirri (2003) make a preliminary study of learning by adopting the methods of pre-test, log analysis, questionnaire, and term-end test, and study the influence of annotation system on learning. The study also finds that the testees generally think that the annotation system provided by the study is helpful to learning, and the self-made annotations are more helpful to students themselves than those made by their classmates (Nokelainen et al., 2003). A study of Quade (1996) also reveals that in the computerized instruction environment, notetaking in computer has better learning effect than notetaking by pencil and paper to Quade (1997).

Marshall (1997) argues that according to the appearance, form and position of annotated signs, annotated signs can be divided into two kinds: telegraphic signs and explicit signs. Telegraphic signs refer to the underlined and colored annotations added to the contents of teaching materials, whereas explicit signs refer to the personal notes added to the teaching material.

In the study of website annotation system made by Hwang and Wang (2005), they use the sharing function to stimulate the learning motive of students. The research results show that the learning way of annotation sharing can better improve the learning effect of students than the annotations of individuals (Hwang & Wang, 2005).

Although many scholars use annotation system to do researches, they hardly mention how to share the valuable annotation process of a student with other students who need sharing. Hence, the study attempts to employ scaffolding theory proposed by the psychologist, Vygotsky. According to Vygotsky, the development of “cognition” is divided into two levels: one is the real level of development, and the other is the potential level of development. The former refers to the level that students can independently solve problems, whereas the latter refers to the level that students need the guidance of or cooperation with other people (teachers, the more outstanding classmates) to solve the problems encountered in the active knowledge construction process (Vygotsky, 1978). The distance between these two levels is called by Vygotsky as the zone of proximal development (ZPD) (Vygotsky, 1978). Therefore, if instruction can be close to the ZPD of students, it can effectively help students to be promoted from the original development level to a higher development level. It also refers that the teacher has adopted a temporary support structure to assist students in developing their abilities. This kind of guidance is called “scaffolding” (Vygotsky, 1978). A scaffolding chart is shown in Fig. 1.

3. Research approach

The study mainly proposes using Expert Keywords Annotation Alignment Algorithm (EKAAA) to inspect whether the annotation contents of students contain the keywords selected by experts, find out through the annotation association rules of Data Mining the Annotating Pattern of the students of high-score cluster to serve as the learning scaffolding for the students of different clusters, and further achieve the purposes of knowledge sharing and learning. The complete flow chart is shown in Fig. 2.

3.1. Pre-test

In order to understand the difference in annotation act of the students of different levels while reading the teaching materials, the study took 110 students of a high school as the research targets to carry out the research. In the beginning of a school term, LMS was used to carry out a pre-test of “VB program language” curriculum. The pre-test result was used as the reference of clustering. In order to enhance the prediction functions of the test, the paper stresses the cross polarization discrimination (XPD) analysis of questions, as shown in Table 1.

From the CPD index shown in Table 1, the CPD of Q9 is lower than 0.2, implying that it is not a good question and has to be eliminated. Therefore, the study has nine valid questions.

3.2. Cluster

According to the arguments of Kelley (1939), under normal distribution the most suitable rate is 27% for high-score and low-score clusters respectively. Hence, the pre-test results of students achieved in the study are divided into 3 clusters, as shown in Fig. 3.

3.3. Import of annotation process

After the students have made annotation, the system automatically imports the annotation process to the database to serve as the source of information for “Expert Keywords Annotation Alignment” in future.

3.4. Expert Keywords Annotation Alignment Algorithm (EKAAA)

Since the study provides very complete contents of teaching materials, most of the students add such “telegraphic” annotations...
as underlines and colors to the teaching materials, but rarely make the annotation of notes themselves. Therefore, the study takes “telegraphic” annotations as the main research area.

In order to find out the telegraphic Annotation Patterns of different clusters for automatically recommending them to and sharing them with the students of different clusters for reference, first of all, the annotation process of all the students had to carry out pre-handling work, with the following steps adopted:

Step 1: Rule out the full-text annotationsIt refers that the student makes annotation on all the teaching materials. This is an unreasonable phenomenon.

Step 2: Judge the validity of annotationAre the annotation contents of students valid (the main points) to teachers? We have to examine whether the contents of a section annotated by the student contain the keywords in “expert keywords database.” If positive and the alignment is successful, then the annotation of the section is valid; otherwise, it is an invalid annotation. Suppose \( C = \{C_1, C_2, \ldots, C_N\} \) is a set of the contents of a sections annotated by a student, and \( K = \{K_1, K_2, \ldots, K_N\} \) is a set of keywords of the section selected by experts. If \( K \) is a subset of \( C \), then \( K \subseteq C \). Suppose a student makes annotation of the contents of section \( C_1 \), if \( K_1 \subseteq C_1 \), then we can say that \( C_1 \) has contained \( K_1 \), implying that the annotation of the section for alignment is considered valid. On the contrary, if \( K_1 \nsubseteq C_1 \), then we can say that \( C_1 \) has not contained \( K_1 \), implies that the annotation of the section for alignment is considered invalid.

The chart of “Expert Keywords Annotation Alignment” Algorithm is shown in Fig. 4 as follows:

In Fig. 4, when a student thinks a certain section important when reading the digital e-teaching materials [Item (1) in Fig. 4], the student can use the annotation tool to do the e-annotation act. In this study, the system will automatically record the annotation process of the student and inspect the rationality of annotation. When the student carries out full-text annotation, it is considered an unreasonable annotation act [Item (2) in Fig. 4] because full-text annotation is of no help to the learning of students at all; otherwise, it is regarded as a reasonable annotation act [Item (3) in Fig. 4]. Therefore, the system further compares the annotation contents \( C_0 \) with the expert keywords database \( (K) \) [Item (5) in Fig. 4]. If \( K \subseteq C \), we can say that \( C \) has contained \( K \), implying that the annotation of the section for alignment is valid [Item (4) in Fig. 4]. The flow chart of “Expert Keywords Annotation Alignment” is shown in Fig. 5.

Therefore, the study uses the algorithm to draw a flow chart of “Expert Keywords Annotation Alignment,” with its virtual codes indicated as follows:

```void Main( ) // Main Program
   Call Reader_Annotation _Section (Section_Count)
End
```

// Reader Annotation Section
```
Sub Reader_ Annotation _Section (Section_Count)
   for (i=1;i<Section_Count;i++) do
      begin
         if Content_Array(i) Is Remarked Then // the section is annotated.
            if Content_Array(i) NotALL Is Remarked Then // all annotations are removed.
               Annotation_Content_Array(i) = Content_Array(i) // valid annotation
         end
      end
```

Table 1
Cross polarization discrimination (XPD) index of questions

<table>
<thead>
<tr>
<th>CPD</th>
<th>Question</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>Q5</th>
<th>Q6</th>
<th>Q7</th>
<th>Q8</th>
<th>Q9</th>
<th>Q10</th>
</tr>
</thead>
<tbody>
<tr>
<td>PH</td>
<td></td>
<td>0.93</td>
<td>0.93</td>
<td>0.93</td>
<td>0.93</td>
<td>0.71</td>
<td>0.86</td>
<td>0.64</td>
<td>0.14</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td>PL</td>
<td></td>
<td>0.71</td>
<td>0.57</td>
<td>0.36</td>
<td>0.21</td>
<td>0.21</td>
<td>0.14</td>
<td>0.21</td>
<td>0.21</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>CPD</td>
<td></td>
<td>0.22</td>
<td>0.36</td>
<td>0.57</td>
<td>0.72</td>
<td>0.72</td>
<td>0.57</td>
<td>0.65</td>
<td>0.43</td>
<td>0.14</td>
<td>0.79</td>
</tr>
</tbody>
</table>

Fig. 2. Flow chart of the study.

Fig. 3. Rates of clusters under normal distribution.

Fig. 4. Flow chart of the study.

Fig. 4. Expert Keywords Annotation Alignment chart.

Fig. 5. Flow chart of "Expert Keywords Annotation Alignment".
Annotation_Count += 1
end
End Sub
// Expert Keywords Database and Reader Annotation Alignment
Sub Expert_Keyword_Alignment_Content( )
for (i=1;i<Annotation_Count;i++) do
begin
if Annotation_Content_Array(i) Like '% Expert_Keywords (i) %'
then // compare database with annotation
Valid_Annotation(i)=Annotation_Content_Array(i) // successful alignment
Valid_Annotation_Count +=1
end
End Sub
// Get the first 27% Annotation Patterns of students from each cluster
Sub Get_HighScore_Annotation_Pattern( )
for (i=1;i<3;i++) do // 3 clusters in total
begin
for (j=1;j<Valid_Annotation_Count;j++)
Select top 27 % Valid_Annotation(i,j) From Ann_Content
Order by PreScore
HighScore_Annotation_Pattern(i,j) = DataMining(Valid_Annotation(i,j))
end
End Sub
3.5. Data Mining
The association rules of Data Mining were used to find out the Annotation Patterns of the first 27% students from each cluster. Take the students of high-score cluster for example, for the annotation process of the first 27% students of “high-score cluster,” “annotation association rules” are employed by the students to find out the important annotation contents they think important, as well as the “Annotation Pattern” having successfully passed the “Expert Keywords Annotation Alignment,” so as to form the common annotation model of the students of high-score cluster.
Suppose that the annotations made by five students of the first 27% students of “high-score cluster” while reading the teaching materials are found successfully containing the keywords selected by experts. The successful annotation alignment process of students is shown in Table 2. In the table, the annotation of student S1 contains the expert keywords of K1, K2 and K3, and so on and so forth.
1: The student has made annotation and is successful in alignment.
2: The student has made no annotation or is failed in alignment.
3.6. Sub-cluster
In order to verify whether IASM can enhance the learning effects of the students of different clusters, the study had to sub-cluster the students of each cluster into “experimental group” and “control group,” as shown in Table 5. The way of sub-clustering is as follows:
By the way of random selection, each cluster was sub-clustered into two groups: “experimental group” and “control group.”

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Successful annotation alignment process of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keyword</td>
<td>Student</td>
</tr>
<tr>
<td></td>
<td>S1</td>
</tr>
<tr>
<td>K1</td>
<td>1</td>
</tr>
<tr>
<td>K2</td>
<td>1</td>
</tr>
<tr>
<td>K3</td>
<td>1</td>
</tr>
<tr>
<td>K4</td>
<td>0</td>
</tr>
<tr>
<td>K5</td>
<td>0</td>
</tr>
</tbody>
</table>

Note: 0, the student has made no annotation or is failed in alignment.

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Successful annotation alignment sequence of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keyword</td>
<td>Student</td>
</tr>
<tr>
<td>S1</td>
<td>K1, K2, K3</td>
</tr>
<tr>
<td>S2</td>
<td>K1, K2, K3</td>
</tr>
<tr>
<td>S3</td>
<td>K1, K2, K3, K4</td>
</tr>
<tr>
<td>S4</td>
<td>K2, K3, K4, K5</td>
</tr>
<tr>
<td>S5</td>
<td>K3, K4, K5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4</th>
<th>C2-itemsset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Itemset</td>
<td>No. of times of success</td>
</tr>
<tr>
<td>K1,K2,K3</td>
<td>3</td>
</tr>
<tr>
<td>K1,K2,K4</td>
<td>1</td>
</tr>
<tr>
<td>K2,K3,K4</td>
<td>2</td>
</tr>
<tr>
<td>K3,K4,K5</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Table of number of students of experimental group and control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>Cluster</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Experimental group</td>
</tr>
<tr>
<td></td>
<td>Control group</td>
</tr>
<tr>
<td>No. of students</td>
<td>30</td>
</tr>
</tbody>
</table>
3.7. Post-test

Post-test was taken as a reference for evaluation of learning effects.

3.8. Test

In order to test whether there is significant difference between experimental group and control group in each cluster after IASM has been taken for sharing among the students of different clusters, the study used $t$-test to carry out data analysis. Suppose the significant difference is $\alpha$. If $P$-value $< \alpha$, then $H_0$ is rejected, implying that there exists a significant difference between the means of experimental group and control group. If $P$-value $\geq \alpha$, then $H_0$ is

<table>
<thead>
<tr>
<th>Cluster and group</th>
<th>Mean</th>
<th>Standard</th>
<th>$t$ value</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-score cluster</td>
<td>Experimental group</td>
<td>77.13</td>
<td>9.83</td>
<td>0.453</td>
</tr>
<tr>
<td>Control group</td>
<td>75.53</td>
<td>9.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medium-score cluster</td>
<td>Experimental group</td>
<td>50.28</td>
<td>9.95</td>
<td>2.141</td>
</tr>
<tr>
<td>Control group</td>
<td>44.28</td>
<td>9.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low-score cluster</td>
<td>Experimental group</td>
<td>32.93</td>
<td>8.90</td>
<td>3.685</td>
</tr>
<tr>
<td>Control group</td>
<td>23.07</td>
<td>5.32</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* $p < 0.05$.  
** $p < 0.01$.  

![Fig. 6. Structure chart of the system.](image)

![Fig. 7. Intelligent Annotation Sharing Model (IASM).](image)
not rejected, implying that there exists no significant difference between the means of experimental group and control group. The test results are shown in Table 6.

3.9. Analysis

As found in the t-test of independent samples, after IASM is taken, there is a significant difference in both medium-score cluster and low-score cluster (both P-values are < .05), but no significant difference in high-score cluster.

4. Intelligent Annotation Sharing System

4.1. Structural chart of the system

In Fig. 6, through the operation interface of browser, students can read the digital teaching materials (HTML documents) in the LMS system. In the reading process of students, they can use "annotation tool" to annotate the teaching materials or insert words of explanation. Then the system will store them in the annotation database. In this way, when the students read the same chapter next time, the teaching materials being annotated last time will be presented. Besides, the annotation contents of other students of high-score cluster can be seen through the "Intelligent Annotation Sharing Model." The annotation contents of the students of other high-score clusters can be watched, thereby achieving the purpose of knowledge sharing.

Scaffolding learning theory suggests that different support frames of learning should be constructed according to the abilities and learning progress of students (Yang, 2000). For the annotation process of the students of different clusters, the study uses Data Mining technique to establish adaptive annotation sharing models of the students of different clusters. The model is called by the study the "Intelligent Annotation Sharing Model" (IASM), and its

Fig. 8. Interface of annotation tool used by students for reading e-teaching materials.

Fig. 9. Annotation sharing interface.
purpose is to verify whether the model can enhance the learning effects of the students of different clusters. After the students of different clusters have been guided by “scaffolding”, a learning purpose structural chart is formed, as shown in Fig. 7.

Notes:

ZPD1: Annotation sharing model of low-score cluster
ZPD2: Annotation sharing model of medium-score cluster
ZPD3: Annotation sharing model of high-score cluster

4.2. Practice environment interface of system

4.2.1. Interface of annotation tool used by students for reading e-teaching materials

When students are reading e-teaching materials, they can use the annotation tool of underlining or notetaking at any time to assist their reading. The system is able to store automatically the annotation process of students in the database, as shown in Fig. 8.

4.2.2. Annotation sharing interface

The annotated documents of students not only are helpful to the original annotator. Through the “Annotation Sharing Model,” the student himself/herself or other students can also refer to the annotation contents of the students of high-score cluster to achieve the purpose of knowledge sharing, as shown in Fig. 9.

5. Conclusions

A complete e-teaching system not only provides suitable teaching materials for students, but is also requested to be a mechanism allowing students to undergo annotation act and online sharing of annotation. Hence, the study develops an Intelligent Annotation Sharing System (IASS) that is suitable for students to use, and can serve as an auxiliary tool for students in reading the teaching materials.

The three main contributions deduced from the study are:

1. An Intelligent Annotation Sharing System (IASS) is developed for students as an auxiliary tool for the reading of teaching materials and for knowledge sharing, with an aim to compensate the insufficiency of traditional annotation system.

2. Expert Keywords Annotation Alignment Algorithm (EKA AA) is proposed to inspect whether the annotation contents of students contain the keywords selected by experts in order to rule out the unreasonable annotation contents of students, such as full-text annotation or random annotation, etc.

3. The Annotation Pattern of high-score cluster acquired through the annotation association rules of Data Mining serves as a scaffolding of learning to guide the students of different clusters. Thus, the study also discovers that except the students of “high-score cluster,” having used “Annotation Pattern” to guide their learning, the students of “experimental group” in “medium-score cluster” and “low-score cluster” have significant improvement. Therefore, the researcher further interviewed the students of high-score cluster, and found that most of the students of high-score cluster belonged to field dependence since they had their own ways of learning and less depended on external assistance.

Although the results acquired from the statistical analysis of “medium-score cluster” and “low-score cluster” show that “experimental group” has more significant difference than “control group,” the P-value of “low-score cluster” is less than 0.001, whereas the P-value of “medium-score cluster” is 0.037, implying that the students of “low-score cluster” have a greater need of learning guidance by “Annotation Pattern” than the students of “medium-score cluster.” Therefore, when a majority of students encounter difficulties in the process of active construction of knowledge, the learning effects of students resulted from the guidance of adaptive “Annotation Pattern” are significantly different from the learning effects of students without the guidance of adaptive “Annotation Pattern.”

Finally, there is one noteworthy thing that after the students of low-score cluster have been guided by “Intelligent Annotation Sharing Model” (IASM), if their learning effects have been significantly enhanced, the instructor can also suggest the students of “low-score cluster” to try to take the IASM of “medium-score cluster” for reference. However, they are not suggested to use the IASM of high-score cluster because the learning of students has to follow in order and advance step by step.

References


