A novel approach to monitoring and creating significant learning experiences using social tag cloud navigation

Jun Ming Chen[a*], Meng Chang Chen[b], Yeali S. Sun[a]
[a]*Department of Information Management, National Taiwan University, Taiwan*
[b]*Institute of Information Science, Academia Sinica, Taiwan*
*d96725001@ntu.edu.tw

Abstract: Prior knowledge learning for increasing the efficiency of the learning processes is an important issue. Traditional studies on prior knowledge generation during language learning have focused on providing additional supplemental materials from reading materials that are manually generated by educators. This is time-consuming and hence personalized prior knowledge recommendation is difficult to perform and monitor. To cope with these problems, we purpose a novel tag-cloud visualization learning approach to automatically monitor running activities and student status. In addition, we incorporate cloud tags into a prior knowledge learning system (TAK), which provides students an engaging way to reinforce meaningful topics, identifies suitable supplementary materials through tag cloud navigation, and helps students reevaluate their reading process. Our experimental results demonstrate that our approach not only significantly improves the efficiency of prior knowledge learning but also helps teachers assist students in improving their reading ability.

Keywords: Social tagging, Reading comprehension, Information retrieval, Teaching and Tutoring

Introduction

In the past, effective reading comprehension requires the integrated interaction of derived text information and preexisting reader knowledge [6, 10], especially with learners of foreign languages such as English. Studies have found that students who are unable to link new knowledge with prior knowledge have problems understanding, recalling, and accessing the new knowledge later [3, 4, 9]. This further suggests it is important to assist students in obtaining relevant prior knowledge, as this can enable them to engage meaningfully with their learning material.

Additionally, keeping students focused is a concern for many teachers. Several studies have demonstrated that teachers, due to either a lack of administrative support or time constraints, have difficulty evaluating student literacy effectively [1, 2, 3, 4]. In some cases, it is very time-consuming to complete the evaluation process, and this leads to significant learning obstacles for students. Thus, identifying and understanding the level of a student's knowledge is important for teachers to help students learn efficiently [5].

To cope with this problem, this paper proposes a novel tag-cloud visualization learning approach. This novel approach uses a social tagging analysis mechanism and tag cloud visualization, which gives students an overview of the knowledge represented and helps them quickly grasp the structure and concepts of English articles. Moreover, it also provides a monitoring interface that enables teachers to navigate to potentially relevant information, and assists teachers in analyzing interactions from student tag behaviors and evaluating student learning performance.

1. Background and motivations

1.1 A typical Web 2.0 phenomenon: Tag cloud

One of the ways in which the Internet helps facilitate knowledge conversion and sharing is through the use of Web 2.0 social tagging. Tags usually represent words or phrases that enable users to easily add metadata to online content [1]. Tag clouds have emerged as an important new interface paradigm, quickly gaining popularity in social information sharing sites such as Flickr and Delicious, which need to find visually appealing ways to summarize vast amounts of information. Fig. 1 shows an example of a tag cloud on Delicious, where tags on the site are organized by popularity.

Figure 1: Example of a tag cloud (Extracted from Delicious)

Recently, researchers are increasingly turning to tag applications such as social networking sites to enhance classroom learning and develop a new generation of learning architecture [1, 2, 3, 4]. Thus, by aggregating tag annotations and corresponding resources from documents, the tag cloud provides navigational clues and helps people easily browse and discover on-line resources [11]. Meanwhile, a tag cloud can also provide learners a visually appealing picture of main themes by using tags to visually analyze the frequency of words in a text. An extension application to tag clouds involves using semantic web technologies to generate tag clouds of semantic concepts, which provides a new way to extract meaningful learning results and analyze learning behavior. This information is helpful to teachers for evaluating student learning achievements and determining learning status.

1.2 A tag-based prior knowledge recommendation system (TAK)

Several studies have been conducted to analyze the learning behaviors of students by using Web 2.0 social tagging techniques to collect implicit information for reading comprehension [1, 2, 3, 4, 7]. A tag-based learning environment called TAK [3, 4] has been developed to improve prior knowledge construction and learning. This system provides effective recommendations, especially through social tagging to spread prior knowledge generation. However, it is difficult for teachers to observe ongoing, running activities and student behavior without assistance.

To cope with the above problems, in the following sections, a novel tag-cloud visualization learning approach is proposed and implemented to extend the TAK system into an effective and efficient mechanism for assisting teachers in tracing and analyzing the information search behaviors of students. Moreover, this study also attempts to harnesses Web 2.0 principles to find new ways for using tag clouds to get a quicker grasp of the main focus of reading materials.

2. I-Cloud: A tag-cloud visualization learning approach

To improve English reading comprehension and assist teachers in monitoring student progress, this study proposes an integrated tag-cloud based learning and monitoring approach, based on fundamental skills of analysis, contextualization, conceptualization, and visualization, called I-cloud. Fig. 2 illustrates the framework of our approach, which consists of three parts: data preprocessing, tag cloud for social navigation, and tag cloud for learning alarm.

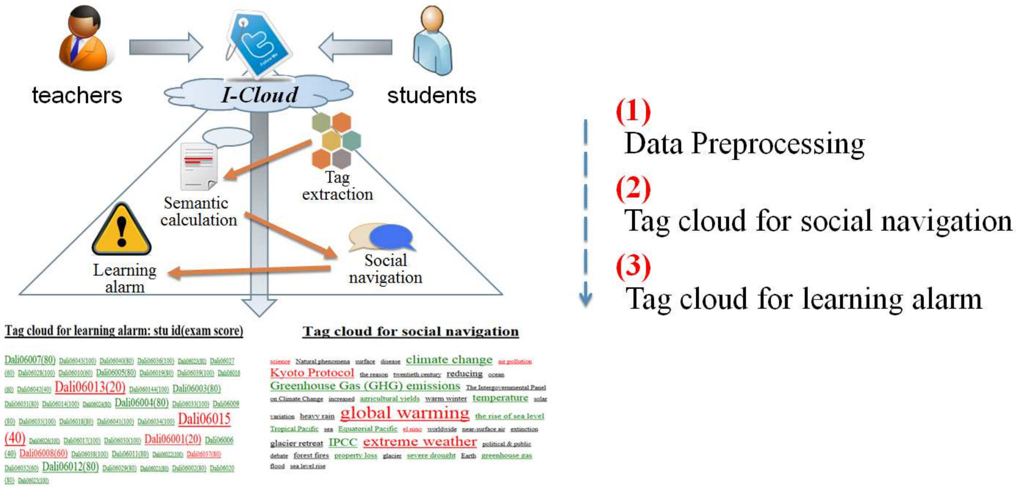

Figure 2: framework for the support approach

2.1 Data Preprocessing

The preprocessing of student tags and articles involves translating tags and reading text into a vector of terms, based on the Vector Space Model (VSM). Each article is preprocessed by 1) tokenizing, 2) removing stop words, 3) and Porter Stemming. In order to identify meaningful paragraphs and locate article topics, we use Latent Semantic Analysis (LSA) to measure the importance of each paragraph and the relationship between paragraphs. The tag is also considered a measure of the importance of a paragraph in the article. Meanwhile, we calculate a tag score to identify the importance of each subsequent paragraph by using the diagonal matrix elements with the LSA process and TF-IDF of the vector-based information retrieval scheme [4]. After this scoring process, we derive the tag score to identify the importance of each subsequent paragraph.

2.2 Tag cloud for social navigation

A tag cloud can be regarded as a collection of main topical terms mined from the articles. It enables the reader to identify the context of the retrieved text data and quickly determine if it is of interest. In this subsection, we describe the tag-cloud generation technique for student tags of well-tagged items. Here, we applied an information retrieval scheme and collaborative filtering (CF) to mine topical terms intelligently, which are then presented in the form of a cloud of tags with varying sizes emphasizing degrees of relevance. This process is summarized below:

(1) To identify the importance of each subsequent paragraph, the tag score ($W_t^{i,j}$) is calculated from the pre-processing process. $W_t^{i,j}$ represents the tag score of the t_{th} tag of i_{th} paragraph in the j_{th} article. It is regarded as the evaluation of the importance of a

paragraph and can help students to construct the article structure. Top K (user-given parameter) recommended tag sets $Top_K(\text{t}) = \sum_{t\in Top_k}\sum_{i\in I} Top_K(\text{W}_t^{i,j})$ are considered for recommending topic tags to students, and help students understand and identify key ideas within the reading (K set in our experiments to 5).

(2) Before identifying appropriate supplementary materials for a student, we recommend a tag cloud to each student. The basic idea of our recommendation approach is collaborative filtering (CF), information retrieval and tag usefulness [12]. The underlying assumptions of CF are that people with similar interests will prefer similar tags, and that those individuals who agreed in the pass tend to agree again in the future [13]. As in user-based CF [8], our approach predicts the tag score likelihood of neighborhoods of students to reflect meaningful tags to be recommended to the user. Tag usefulness is determined by powerful coverage of the most frequently used tags. Thus, the recommendation score of tag t for social navigation (SN) for student u is determined by:

$$\text{SN(u,t)} = e^{-\gamma}\times[\alpha\times\sum_{v\in N(u)} \text{w(u,v)} \times \text{w(v,t)} + (1-\alpha)\times W_{Top_K}(\text{t})\times\sum_{t\in T(u)} \text{w(u,t)}] + \sum_{u=1}\left(\frac{\log(d_{t,j})}{m^2}\right)$$

where γ is a propagation decay factor; α represents a parameter that controls the relative weight between people and tags (initial α is 0.5); w(u,v) and w(u,t) are the relationship strengths of u to student v and tag t, as given by the student profile; w(v,t) represents the relationship strengths between student v and tag t; W_{Topk}(t) represents the tag weight for recommending topic tags; $d_{t,j}$ represents the frequency with which the t_{th} tag has been used to describe the j_{th} article; and lastly m is the number of different tags assigned to the j_{th} article.

(3) For each tag t, the normalized weight SN_{norm}(u,t) is used to calculate its font size. The cloud accomplishes this by increasing and decreasing the font size of each tag according to its weight. Thus, when the student clicks on a given tag in the tag cloud the system selects suitable supplementary materials based on our previous study, the tag-based prior knowledge recommendation system (TAK) [3, 4], and presents them to the student. This function is an engaging way of visualizing meaningful information and helping rethink the reading process for students.

In brief, this stage adopted a social tagging approach, based on a tag cloud navigation schema, to help us examine and understand each aspect of student learning behavior. The social navigation interface is shown in Fig. 3(a). A tag is regarded as a single keyword or a phrase that describes the topic, theme, or idea of the article. Students can add as many tags as their wish, and use the tag cloud visualization tool to discover interesting clues and refine their thoughts or ideas of the reading. Thus, when students click on a given tag in the tag cloud, the system serves as a useful reference guide, as well as selects suitable supplementary materials for students by analyzing the characteristic of the tag cloud.

2.3 Tag cloud for learning alarm

To help teachers evaluate a student’s learning situation and monitor the learning achievements of individual students more easily, the learning alarm is combined with the tag-cloud monitoring mechanism into a tag-based prior knowledge recommendation (TAK) system. A tag-cloud monitoring mechanism can be extremely useful to help teachers keep up with all relevant student activity and drill down into their detailed learning portfolio by tag cloud navigation. This process is summarized as below:

(1) Students read an article and take an exam, and the score is used to evaluate their comprehension. We assume the students' scores approximate their level of comprehension, and thus our tag score mechanism is a combination of a student's exam score with their tagging preferences for the same article.

(2) To estimate the error in interpreting an individual's exam score, the standard error of measurement (SEM) is calculated by using a reliability coefficient and the standard deviation of the exam, as follows:

$$\mathrm{SEM} = \mathrm{SD} \times \sqrt{1 - \frac{\mathrm{Q}}{\mathrm{Q}-1}\left(1 - \frac{\mathrm{M}(\mathrm{Q}-\mathrm{M})}{\mathrm{Q}\cdot\sigma^2}\right)}$$

where SD is the standard deviation of the exam, Q is the number of exam items, M is the average exam score of every student who takes the exam, and σ^2 is the variance of every student who takes the exam. The larger the SEM, the less reliable the test. Thus, we use the standard error of measurement (SEM) to recalculate the exam score.

(3) Not all tags carry the same amount of information, and so we use the entropy of a tag to measure its potential characteristics. The value is normalized into a range of 0 and 1.

$$Entropy(t) = -\sum_{u\in U} p_j(u,t)\times\log(p_j(u,t)) \quad ; \quad p_j(u,t) = \frac{TF_j(u,t)}{\sum_{t'\in T} TF_j(u,t')}$$

Here, $P_j(u,t)$ represents the probability of tag t appearing in the j_{th} article for student u, and $TF_j(u,t)$ denotes the number of times tag t appears in the j_{th} article for student u. Tags with higher entropy have lower tag scores. Tags with higher tag scores are important, because these tags help describe the topics or ideas they annotate in a more accurate way. The learning score in the j_{th} article that the u_{th} participant annotates can be represented as follows:

$$\mathrm{LA}_j(\mathrm{u}) = \lambda \times \frac{1}{\mathrm{SEM_j}} \times \left(\frac{\mathrm{S}_\mathrm{u}^j}{\mathrm{Avg}(\mathrm{S}^j)}\right) + (1-\lambda)\times(1 - Entropy(t))$$

where λ represents a parameter that controls the relative weight between the exam score and a tags' importance, S_u^j is the exam score of the j_{th} article that the u_{th} student receives, and $\mathrm{Avg}(\mathrm{S}^j)$ is the average score of the j_{th} article for all students. Here, large values signify a strong alert status where a teacher should be aware of the early conditions of a learning problem.

(4) A threshold θ is used to determine the selection warning sign in order to discover an abnormal learning status from tagging behaviors and reading comprehension degrees. When a teacher logs into the proposed system, they can determine the threshold to warn of any potential abnormal learning status of individual students. When $\mathrm{LA(u)} < \theta$, a user's name in the tag cloud is shown with a green font color, and the larger font size represents better learning performance. In contrast, a user's name in red font attracts the teacher's attention and the larger font size represents lower learning performance.

To summarize, Fig. 3(b) shows the learning alarm interface presenting the current status of a student's learning performance and tagging behavior, recently forwarded to the alert monitor. When students who have a learning disability are diagnosed by analyzing the tag cloud for social navigation, this tool can assist teachers in providing focused feedback and questions to students as soon as possible. This helps activate prior knowledge, probe students' conceptual understanding, and leads to deeper understanding.

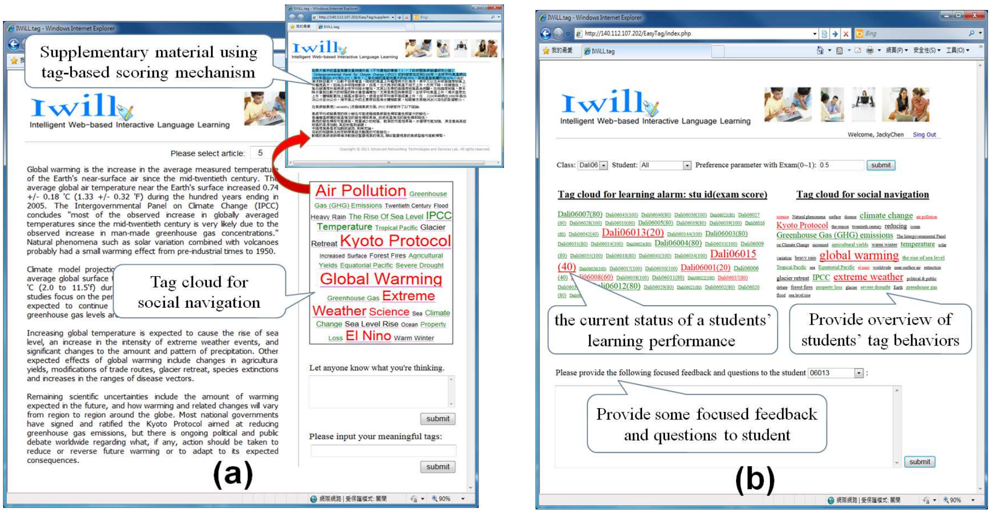

Figure 3: Interface of I-cloud (a) Tag cloud for social navigation (b) Tag cloud for learning alarm

3. Experiments

The goal of our experiment is to quantify the influence of I-cloud practices on an on-line learning environment. In this section, a series of investigations is conducted to evaluate the usefulness of I-cloud.

3.1 Student data and results

In order to evaluate the efficacy of tag cloud for social navigation, an experiment was conducted from April 2012 on reading activity at a senior high school in Taiwan. 86 participants were divided into two groups. A pre-test and post-test experimental design employed before-and-after surveys to demonstrate the usefulness of our approach among participants. Each group was given a pre-test evaluation and then a post-test after two weeks. The experimental students demonstrated significant improvement in reading comprehension ability by taking advantage of the tag-cloud for social navigation (Table 1; $p < .05$). Fig. 4 shows the learning curve of students that is representational of an average rate of knowledge gained over time. Students first start learning how to use an on-line reading system in the first 10 minutes of the reading activity. The control group on a flat learning curve demonstrates that the rate of knowledge gained is slowly spaced out over time. That implied that subjects with a flat curve are often very difficult to learn, so that students can't quickly grasp the concepts and use the knowledge. In the contrast, the experimental group gains knowledge quickly after 10 minutes and I-cloud helps them quickly grasp the concepts and knowledge of reading materials. In addition, students' reflections on the activity were collected using a questionnaire and personal interviews. The results of 43 questionnaires were conducted after post-test. Each question underwent a discriminate validity test by using factor analysis. The coefficients from the experimental results show that these factors were sufficiently reliable for representing student-tagging behaviors. The major findings are presented as follows:

(1) 74% of students agreed that the I-cloud interface was easy to use.
(2) 88% of students agreed that tag cloud offers a way to navigate through a structure of article based on tags, and it can help students quickly grasp the structure and concepts of English articles. Some students indicated that these tags help them easily realize new information from the article.
(3) 79% of students thought that tag cloud navigation reduced the amount of search time and enhanced their search experience.
(4) 93% of students agreed that tag cloud can enable students to select different views on the tag cloud, such as system tags or other student tags, and then help judge difference viewpoints and explicate their thinking processes.
(5) 86% of students agreed that the tag cloud provides navigational clues through corresponding supplementary materials, which help students understand prior knowledge in the article.

Table 1: Paired t-test of the pre-test and post-test results

Test	Group	N	Mean	Std. Deviation	Std. Error Mean	t-test
Pretest	Control Group	43	57.79	28.0794	4.2821	t = .588
	Experimental Group	43	54.65	25.5530	3.7443	p = .56
Posttest	Control Group	43	60.70	14.5807	2.2235	t = -2.487
	Experimental Group	43	66.86	11.8534	1.8076	**p = .017***

*p < .05

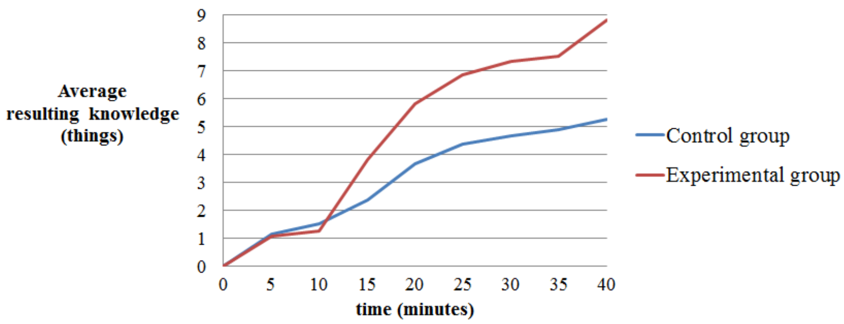

Figure 4: Learning curve simplified

3.2 Teacher data and Results

To evaluate the effectiveness of the tag cloud for learning alarm, five participating teachers from a senior high school in Taiwan were invited to experience the use of I-cloud. We designed a questionnaire that was used to collect feedback from the teachers. The major findings are presented as follows:

(1) All teachers agreed that the interface of I-cloud is helpful for teachers to realize the learning status of students. This interface can assist teachers in evaluating the reading ability of the students, such that constructive suggestions can be given to the students, and tutoring strategies can be improved accordingly. Moreover, I-cloud provides a new way to assist teachers that exploit tag information on the students' tag cloud to get their attention and bring them back into focus.
(2) 80% of the teachers agreed that the interface is intuitive and easy to use. Only one teacher indicated that the interface needed more guidance.
(3) 80% of the teachers believed that tag cloud for social navigation can enhance the ability of the students for prior knowledge learning.

4. Conclusions and Future Work

In this work, we propose a novel method for monitoring and creating significant learning experiences. The results of this study showed that experiences with applying social tagging and tag cloud helped students increase their reading comprehension and quickly grasp the structure and concepts of English articles. In addition, the experimental results demonstrated that this novel approach is also helpful in assisting teachers in evaluating student learning achievements by tag-cloud visualization. These results provide suggestions and references for the design of efficient web 2.0 supported collaborative learning activities in the future. Furthermore, we hope to examine the differences between the effectiveness of a more declarative approach, and make it easier to use and combine different tools and data sources for generating learning suggestions. These suggestions should further ensure that social tagging applications in collaborative learning environments improve reading and recommendation incrementally in future experiments.

Acknowledgements

This work was supported by the National Science Council of Taiwan, R.O.C., under contract numbers NSC100-2221-E-001-015-MY3.

References

[1] Chen, J. M., Chen, M. C., & Sun, Y. S. (2010). A novel approach for assisting teachers in assessment of student reading ability in web-based learning environment. International Conference on Computers in Education (ICCE), 260-264.

[2] Chen, J. M., Chen, M. C., & Sun, Y. S. (2010). A novel approach for enhancing student reading comprehension and assisting teacher assessment of literacy. Computers & Education, 55(3), 1637-1382.

[3] Chen, J. M., Chen, M. C., & Sun, Y. S. (2012). A hybrid tag-based recommendation mechanism to support prior knowledge construction. the 12th IEEE International Conference on Advanced Learning Technologies (ICALT).

[4] Chen, J. M., Chen, M. C., Sun, Y. S. & Chen, Y. Y. (2011). A novel approach for enhancing student reading comprehension by activating prior knowledge. International Conference on Computers in Education (ICCE), 279-286.

[5] Chen, L. H. (2011). Enhancement of student learning performance using personalized diagnosis and remedial learning system. Computers & Education, 56(1), 289-299.

[6] Corredor, J. (2006). General and domain-specific influence of prior knowledge on setting of goals and content use in museum websites. Computers & Education, 47(2), 207-221.

[7] Guan, Z., Wang, C., Bu, J., Chen, C., Yang, K., Cai, D., & He, X. (2010). Document Recommendation in Social Tagging Services. In Proceedings of the 19th international conference on World Wide Web, 391-400.

[8] Guy, I., Zwerdling, N., Ronen, I. Carmel, D., Uziel, E. (2010). Social media recommendation based on people and tags. SIGIR 2010: 194-201.

[9] Lin, Y. C., Lin, Y. T., & Huang, Y. M. (2011). Development of a diagnostic system using a testing-based approach for strengthening student prior knowledge. Computers & Education, 57(2), 1557-1570.

[10] Wang, T. H., Wang, K. H., Huang, S. C., & Chen, S. Y. (2004). Web-based Assessment and Test Analyses (WATA) system: Development and evaluation. Journal of Computer Assisted Learning, 20(1), 59–71.

[11] Sinclair, J. & Cardew-Hall, M. (2008). The folksonomy tag cloud: when is it useful? J. Inf. Sci., 34(1), 15-29.

[12] Yusef, H. M., & Víctor, H. S. (2006). Improving Tag-Clouds as Visual Information Retrieval Interfaces, International Conference on Multidisciplinary Information Sciences and Technologies, InSciT2006, 25-28.

[13] Zhen, Y., Li, W., & Yeung, D. (2009). TagiCoFi: Tag Informed Collaborative Filtering. In Proceedings of the 3rd ACM Conference on Recommender Systems (RecSys'09), 69-76.