Hirashima,T. et al. (Eds.) (2010). Workshop Proceedings of the 18th International Conference on Computers in Education. Putrajaya, Malaysia: Asia-Pacific Society for Computers in Education.

Mining Learners' Disorientation in Web-based Learning

Yu-Cheng SHIH[a], Yung-Chi HSU[a], Sherry Y. CHEN[ab*]
[a]*Graduate Institute of Network Learning Technology, National Central University*
[b]*Department of Information Systems and Computing, Brunel University, UK*
* sherry@cl.ncu.edu.tw

Abstract: Web-based learning systems provide flexible ways for learners to develop their own learning strategies. However, some learners may experience more disorientation problems when they need to decide the learning strategies by themselves. In other words, not all of learners are suitable to use the Web-based learning systems. Therefore, it is necessary to investigate how learners' individual differences influence their disorientation problems. To this end, this study addresses this issue. Unlike other previous studies, we will not only examine how learners' individual differences influence their disorientation but also investigate how the learners' disorientation influences their learning performance. The other novelty of this study is the fact that we use a data mining approach to conduct data analyses. The results indicate that learners' computer experience is a major factor to influence their disorientation. Moreover, their disorientation may also influence their learning performance. The results are useful for developing Web-based learning systems to reduce learners' disorientation.

Keywords: Disorientation, Individual differences, Learning performance, Data mining

Introduction

The major advantage of the Web-based learning systems corresponds to their flexibility [2]. More specifically, learners can obtain information from the Web-based learning systems by different ways [17]. More specifically, the Web-based learning systems provide a nonlinear way to help learners develop their own learning strategies.

On the other hand, learners have diverse background, including subject knowledge, cognitive styles, Internet skills, and computer experiences, and so on [3]. Thus, not all learners feel comfortable to develop learning strategies by themselves. In other words, some learners may feel confused when they use the Web-based learning systems. Therefore, reducing learners' disorientation problems needs to consider their individual differences. In other words, there is a need to investigate how individual differences affect learners' disorientation problems when using the Web-based learning systems. Such investigation is important because understanding learners' disorientation is helpful to develop the Web-based learning systems that meet learners' real needs [8] and it can provide guidelines for designing Web-based learning systems that can accommodate learners' individual differences and their corresponding disorientation problems.

In addition to identifying the relationships between learners' individual differences and their disorientation, another important issue is how learners' disorientation problems influence their learning performance [13]. In summary, learners' disorientation problems may not only be influenced by their own individual differences but also affected their learning performance. In other words, it is helpful to understand learners' disorientation because such understanding can be useful to develop the Web-based learning systems that can enhance learners' performance.

Therefore, we not only need to investigate the relationships between learners' individual differences and their disorientation, but also should pay attention to analyzing how the learners' disorientation problems influence their learning performance. To this end, the study presented in this paper investigates these two issues. Such investigation can contribute not only to provide the knowledge of how to reduce learners' disorientation problem from an individual difference perceptive, but also to provide the understanding of how to enhance learners' performance by reducing their disorientation problems. To conduct a deep investigation, a data mining approach is adopted because our previous research shows that it can discover some unexpected relationships [5]. This paper is organized as follows. In section 2, the literature review is described. In section 3, the methodology used to achieve our aims is described. In section 4, the experimental results are presented. The conclusions are summarized in the last section.

Literature Review

Previous studies indicate that non-linear learning strategies may cause learners' disorientation problems. In other words, learners may get lost or become disorientated in the Web-based systems [17]. In an early study, Perlman [9] indicated that the learners' disorientation problems can include uncertainty over what learners had and had not read, the lack of organizational cues to find the information they wanted. Based on these disorientation problems, recently, a number of studies investigate the relationship between learners' individual differences and the degree of disorientation when using the Web-based systems. Ford and Miller [7] indicate that females more easily get lost than males. In addition to gender differences, Mohageg [15] suggested that novices might more easily get lost than experts. Therefore, there is a need to provide novices with suitable content structure to reduce their disorientation problems. Moreover, Last et al. [13] indicate that the level of students' prior knowledge may influence their disorientation problems. As showed in their work, students with high prior knowledge reported positive feelings about using the system and vice versa. Moreover, they also indicated that the students' disorientation problems may also influence their performance. In summary, such findings provided the guidelines for the improvement of the design of the Web-based systems so that learners can avoid getting lost.

Although the aforementioned studies indicate that learners' disorientation problems, which may be influenced by their individual differences, have great impacts on their performance, they mainly conduct data analyses with statistical techniques to which are used to prove known relationships by verifying existing knowledge [16]. However, the problems of the statistical techniques lie within the fact that such techniques cannot detect unexpected relationships [20]. To address this issue, a data mining approach is a better candidate because it can search for valuable information in large volumes of data [11] and uncover some unexpected relationships. The main difference between statistical analyses and data mining lies in the aim that is sought. The former is used to verify existing knowledge to prove a known relationship [16] while the latter is aimed at finding unexpected relationships [20]. Due to the fact that the data mining approach can provide such benefits, recently, several studies used the data mining approaches to evaluate the effects of learners' behaviors. These studies demonstrate that clustering is a useful data mining method (e.g.,[14]). The clustering method is mainly used to partition rough data into several groups base on their similar features [19]. It is suitable for investigating differences and similarities among learners' behavior. For example, an early study by Hay et al. [12] adopted a clustering algorithm to extract sequences with similar behavioral patterns. Their results indicated that learners' navigation patterns are distinctive among different clusters. It implies that clustering can provide the structural properties to identify dissimilarities between each group. Moreover, our recent research [4] adopts a clustering method to investigate the relationship between human factors and students' learning behaviors in the Web-based learning systems. The finding of their work indicates that prior knowledge and subject content are two potential factors, which can distinguish learning behaviors among different clusters of learners.

As mentioned above, the clustering method is a useful tool for investigating the learners' behaviors in Web-based learning systems. Thus, this study adopts the clustering method not only to evaluate the relationship between learners' individual differences and their disorientation problems but also to analyze the relationship between learners' disorientation problems and their performance.

2. Methodology

This section describes the methodology used to reach the two-folds aims of this study and introduce techniques applied to analyze corresponding data.

Web-based learning systems**:** In this study, the subject content of a Web-based learning system is *Interaction design* (Figure 1). The Web-based learning system allow learners to develop nonlinear learning strategies to seek information.

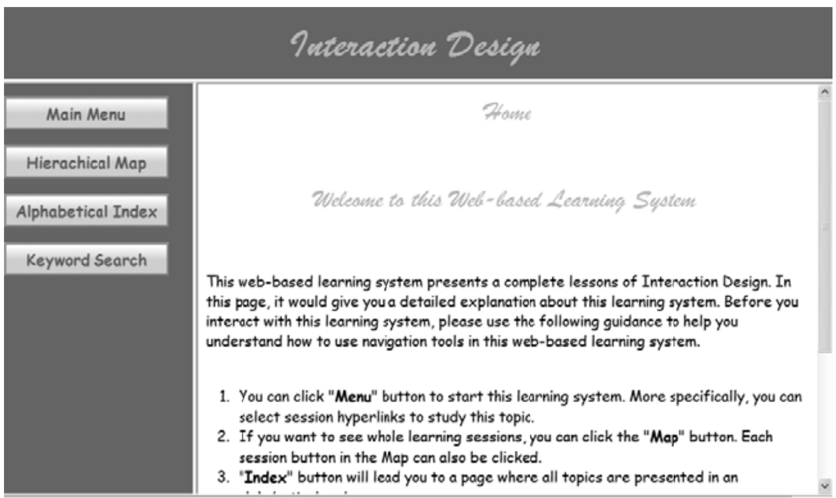

Figure 1. The Web-based learning system.

Participants**:** 50 students from a university in Taiwan participated in our study. These participants were undergraduate or graduate students and they took part in this study voluntarily. All participants had different Internet skills and computer experience of computer aid learning tools to help us identify our aims.

Procedure**:** The procedure was composed of four steps (Figure 2). Each participant had a log file, which recorded his/her personal information, i.e. Internet skills and computer experience of computer aid learning tools. In addition, the participants needed to take the pretest and posttest, which were designed to evaluate participants' learning performance before and after using the Web-based learning system. More specifically, the participants' learning performance was measured based on the differences between the score of posttest and the score of pretest. Both tests included 20 multiple-choice questions, each with three different answers and an "I don't know" option. After learners interacting with the Web-based learning system, they needed to fill out a questionnaire, which was used to measure students' disorientation problems. The questionnaire was developed by an analysis of previous studies on disorientation problems ([9], [18] and [1]). There were 16 questions and each one used a five-point Likert scale ranging *very much, quite a lot, average, not much,* and *not at all*. The smaller of the values represented the low levels of learners' disorientation problems and vice versa. The details of the questions are showed in Table 1.

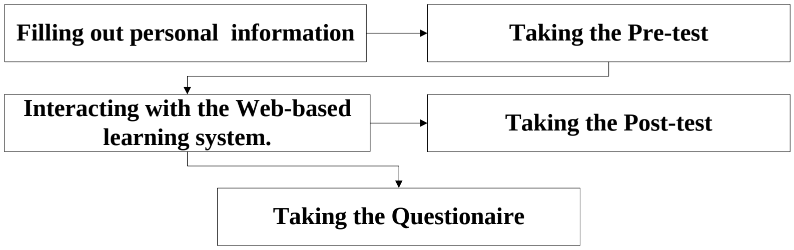

Figure 2. Procedures.

Table 1. The details of questionnaire.

ID	Question
1	I felt the structure of this system not clear.
2	It is very difficult to know which buttons/icons corresponded to the information I wanted.
3	I sometimes got lost because the buttons and icons made me confused.
4	I would have found it more helpful to be given a suggested route through this system.
5	It is hard to find a route for a specific task in this system.
6	I got confused because same page was available through different links or buttons.
7	When proceeding through this system, I was confused which options I wanted, because it provided too many choices.
8	The links provided in this system cannot help me discover the relationships between different concepts.
9	It is not easy to choose any learning topics according to my needs and progress.
10	It is not easy to choose learning sessions according to my needs.
11	It is not easy to review something that I have learnt in this system.
12	I felt confused with the links that show different colors.
13	I was lost with using back/forward buttons.
14	I have difficulties in browsing pages containing texts and links in the same pages.
15	When I navigated in this system, I often forgot where I was.
16	Sometimes I found it hard to keep track which bits I had learnt.

Data Analysis**:** In this study, a clustering method was used to analyze the participants' individual differences, the participants' learning performance and the participants' responses to the questionnaire. Among a variety of clustering techniques used for clustering method, a K-means [10] algorithm is widely used to partition the data into several clusters according to their similar features. The major principle of the K-means algorithm uses k initial centers, each of which is assigned to each cluster, to partition data into k clusters. Each pattern in the cluster is decided based on the nearest distance between the pattern and each cluster center. However, a major limitation of using the K-means algorithm is that the number of clusters needs to be predefined. In other words, there is a need to identify the most suitable number of clusters to perform the K-means algorithm. Such an issue can be treated as parameter exploration [6], which decides the suitable value of parameters. The parameter exploration is useful when dataset is not large. Thus, the issue of the K-means algorithm is suitable for this study because the dataset was not large. Therefore, the parameter exploration was applied to decide the parameters of the K-means algorithm in this study. Subsequently, the number of clusters is set for the large range of value to investigate the robustness in the clustering results. The suitable number of clusters is determined based on not only the smallest distance between the features in same cluster, but also the largest distance between the features in different clusters. After doing so, the suitable number of the K-means algorithm is three.

3 Results and Discussions

The purpose of clustering is to group learners based on their responses to the questionnaire. The higher score from the questionnaire means the situation that they will more easily get lost and vice versa. After performing the K-Means algorithm, three clusters are produced. The percentage of learners within each cluster is satisfactory for the total number of 50 instances. Clusters can be characterized as well balanced: Cluster 0 (N = 21): 42%, Cluster 1 (N= 17): 34%, Cluster 2 (N = 12): 24% (show as Figure 3). As showed in Table 1, the learners are grouped according to the following trends.

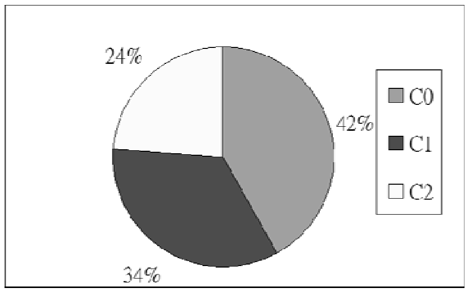

Figure 3 the result from K-Means algorithm.

—Cluster 0 (C0). Learners feel comfortable when using the system. In other words, they are clear with the structure, buttons and icons. On the other hand, they think it would be helpful if a suggested route can be given through this system.

—Cluster 1 (C1). Learners may feel a little confused when using the system. Because the back/forward buttons are unclear to them and it is difficult to keep track which bits they had learnt.

—Cluster 2 (C2). Learners may more easily feel confused when using the system. They are unclear with the structure, buttons and icons and it is difficult to keep track what topics they had learnt. Moreover, they also think that too many options are provided so they do not know how to choose them.

The relationships between students' disorientation problems and their learning performance are examined based on the mean value of the gain scores of each cluster and the mean value of the scores obtained from the questionnaire (Table 2). As shown in this table, the cluster with low degree of disorientation problems may demonstrate high learning performance and vice versa. More specifically, learners in Cluster 0 (C0) get the lowest score from the questionnaire but they have the highest gain score. In contrast, other clusters may get higher scores from the questionnaire but their gain scores are decreased.

Table 2 Score from questionnaire for Each Cluster

Cluster (C)	The score of questionnaire		Performance		Instances
	Mean	STD	Mean	STD	
C0	2.33	0.46	5.29	3.24	21(42%)
C1	3.35	0.36	4.82	2.83	17(34%)
C2	5.23	0.47	3.25	3.57	12(24%)

The aforementioned results indicate that learners' disorientation problems greatly influence their learning performance. Thus, there is a need to further investigate the relationship between learners' disorientation problems and their learning performance.

Beside, the aims of this study also tend to investigate how learners' individual differences influence their disorientation problems. Therefore, the learners' performance and their individual differences are needed to taken into account. The details of the discussion will be showed below.

4.1 Learning performance

In addition to Table 2, the relationships between learners' disorientation problems and their learning performance are further illustrated in Figure 4. As shown in Figure 4, the relationship between learners' disorientation problems and their learning performance are in inverse. More specifically, the learners that never feel confused may obtain high performance while the learners that easily feel confused may have low learning performance. A possible reason is that learners that may not feel confused can easily obtain information they want from the Web-based learning system. As showed in C0, learners which never feel confused are clear with the structure, buttons and icons. In other words, they cannot only easily recognize where to find the topic but also easily understand other relevant topics. It implies that such learners can identify logical relationships between each topic from the Web-based learning system. Thus, this is the reason why they can demonstrate high performance. In contrast, the learners who easily feel confused may feel difficult to locate the relevant information they need. As showed in C1 and C2, both of them easily feel confused because they cannot recognize what topic they have learned or how to locate relevant topics. In other words, they feel confused with the logical relationships between each topic. Thus, their performance is not as good as learners in C0.

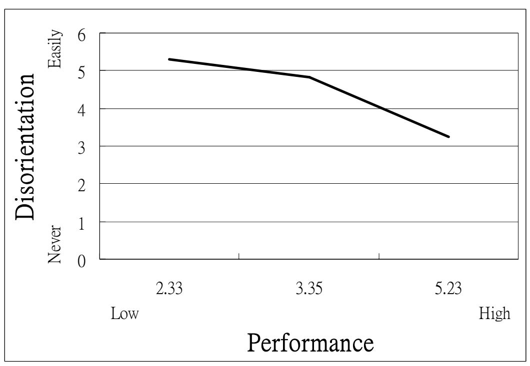

Figure 4 Relationship between learners' disorientation problems and their learning performance.

4.2 Individual Differences

In addition to learners' learning performance, we also investigate how learners' individual differences influence their disorientation. Among various individual difference elements, learners' computer experience is taken into account. The differences of learners' computer experience in each cluster are illustrated in Figure 5. Like Figure 4, the relationship between learners' computer experience and their disorientation problems are also in inverse. More specifically, learners that never feel confused may have more computer experience and vice versa. In other words, learners with much computer experience may feel comfortable with the Web-based leaning system. A possible reason is that learners who have sufficient computer experience can easily understand each function provided by the Web-based learning system so they have less confusion. As showed in C0, learners, who have a high level of computer experience and are clear with the functions, never feel confused. More specifically, they not only understand the meanings of each function but also understand how to use the functions to locate information they need. In other words, learners can easily identify the differences between each function. Thus, they can easily locate information they want by using each function. It also echoes the findings in Section 4.1, which shows that such learners can demonstrate high performance. On the other hand, learners with less computer experience may easily feel confused with each function in the Web-based Learning system. As showed in C1 and C2, both of them easily feel confused because they do not understand the meanings of the functions or how to locate relevant topics by using proper functions. In other words, they may feel confused to identify the differences between each function in the Web-based learning system. Thus, they cannot easily locate the topic they want by using each function. This may be reason why their performance is not as good as learners in C0. (See Section 4.1).

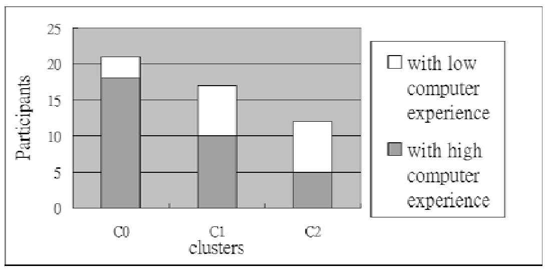

Figure 5 Relationship between learners' computer experience and their disorientation problem.

In summary, the learners' individual differences critically influence their disorientation problems, which may also influence their learning performance. More specifically, learners who have more computer experience may never feel confused to use the Web-based learning system. Thus, they may easily demonstrate high performance. Conversely, learners with less computer experience may easily feel confused when using the system. Thus, it is hard to improve their performance. Figure 6 presents a framework, which summarizes the findings of this study.

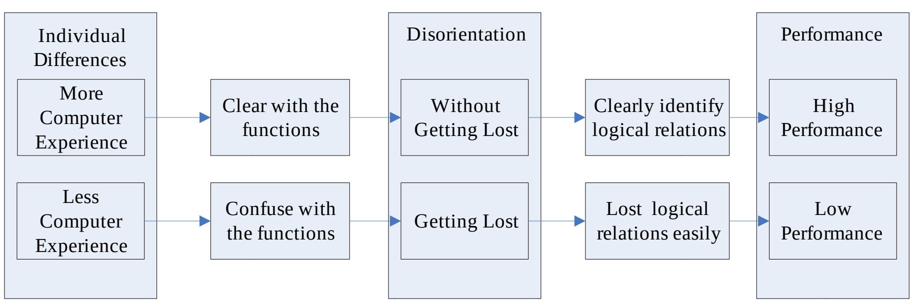

Figure 6 A proposed framework.

5. Concluding remarks

This study investigates learner's disorientation problems with a clustering method. In summary, we found that the learners' computer experiences greatly affect their disorientation problems. Moreover, their disorientation problems also have considerable impacts on their learning performance. In brief, it is necessary to consider learners' computer experience to reduce their disorientation problems, The findings from this study can not only be used to reduce disorientation problems but also to improve their learning performance. In other words, the findings from this study have showed the importance of understanding learners' disorientation problems when using Web-based learning systems. However, this is a small-scaled study so future works should consider a larger sample to provide additional evidence as to learners' disorientation problem but also adopt other data mining methods, such as association rules, classification. Such results can be incorporated into the findings of this study so that reliable user models can be developed to support the development of web-based learning systems.

Acknowledgements

This work is funded by National Science Council, ROC (NSC 98-2511-S-008 -012 -MY3 and NSC 99-2511-S-008 -003 -MY2).

References

[1] Ahuja, J. S. & Webster J. (2001). Perceived disorientation: an examination of a new measure to assess web design effectiveness. *Interacting with Computers*, 14, 15-29.
[2] Chen, S. Y. & Macredie, R. D. (2004). Cognitive modeling of student learning in Web-based instructional programmes. *International Journal of Human-Computer Interaction*, *17*(3), 375–402.
[3] Chen, S. Y., Fan, J. P. & Macredie, R. D. (2006). Navigation in hypermedia learning systems: experts vs. novices. *Computers in Human Behavior*, 22, 251–266.
[4] Chen, S. Y., Liu, X., 2008. An Integrated Approach for Modeling Learning Patterns of Students in Web-based instruction: A Cognitive Style Perspective. ACM Transactions on Computer-Human Interaction. 15(1), Article 1.
[5] Clewley, N., Chen, S. Y. and Liu, X. (2010) Cognitive Styles and Search Engine Preferences: Field Dependence/Independence vs. Holism/Serialism. *Journal of Documentation.* 66(4), 585-603.
[6] De Jong, K. A., 1975. Analysis of the behavior of a class of genetic adaptive systems. Ph. D. Disseration, The University of Michigan, Ann Arbor, MI.
[7] Ford, N., & Miller, D. (1996). Gender differences in Internet perception and use. In: Electronic Library and Visual Information Research. Papers from the third ELVIRA conference. 30 April-1996 (pp. 87–202). London: ASLIB.
[8] Ford, N., Miller, D., & Moss, N. (2001). The role of individual differences in internet searching: An empirical study. Journal of the American Society for Information Science and Technology, 52(12), 1049–1066.
[9] Gary, P. (1990). Evaluating hypermedia systems. Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people, p. 387–390.
[10] Han, J., & Kamber, M. (2001). *Data mining, concepts and techniques: Cluster analysis.* San Francisco: Morgan Kaufmann.
[11] Hand, D. J., Mannila, H., & Smyth, P. (2001). *Principles of data mining*. Cambridge, MA: MIT Press.
[12] Hay, B., Wets, G., Vanhoof, K., 2001. Clustering navigation patterns on a Website using a sequence alignmentmethod. presented at the IJCAI 2001 Workshop Intelligent Techniques for Web Personalization, Seattle, WA.
[13] Last, D. A., O_Donnell, A. M., & Kelly, A. E. (2001). The effects of prior knowledge and goal strength on the use of hypermedia. *Journal of Educational Multimedia and Hypermedia*, 10(1), 3-25.
[14] Minetou, C. G., Chen, S. Y., and Liu, X. (2008) Investigation of the Use of Navigation Tools in Web-based Learning: A Data Mining Approach. *International Journal of Human-Computer Interaction*. 24(1), 48-67
[15] Mohageg, M. F. (1992). The influence of hypermedia linking structures on the efficiency of information retrieval. *Human Factors*, 34(3), 351-367.
[16] Moss, L., Atre, S., 2003. Business intelligence roadmap: The complete lifecycle for decision support applications. Addison-Wesley.
[17] Nielsen, J. & Tahir, M. (2002). *Homepage Usability：50 websites deconstructed*. New Riders Publishing.
[18] Sandelands, L.E., Buckner, G.C., 1989. Of art and work: aesthetic experience and the psychology of work feelings. In: Cummings, L.L., Staw, B.M. (Eds.). *Research in Organizational Behaviour*, 11. , pp. 105-131
[19] Tukey, J. W., 1977. Exploratory data analysis. Reading, MA: Addison-Wesley.
[20] Wang, M., Rees, S. J., Liao, S. Y., 2002. Building an online purchasing behavior analytical system with neural network. In Zanasi, Brebbia,&Melli (Eds.), DataMining III.WIT Press.