JOURNAL OF INFORMATION SCIENCE AND ENGINEERING 34, 1175-1186 (2018)
DOI: 10.6688/JISE.201809_34(5).0004

Applying Learning Analytics to Deconstruct User Engagement by Using Log Data of MOOCs

MING-CHI LIU[1], CHEN-HSIANG YU[1], JUNGPIN WU[2],
AN-CHI LIU[1] AND HSI-MIN CHEN[1,+]
[1]*Department of Information Engineering and Computer Science*
[2]*Department of Statistics*
Taichung, 407 Taiwan
E-mail: {mingcliu; chyu; cwu; acliu; hsiminc[+]}@fcu.edu.tw

Previous research noted that software requirement analysis should move beyond usability to understand and design for more engaging user experiences. This study analyzes system logs collected from the Massive Open Online Courses (MOOCs) to capture user engagement over time. We measured engagement through mapping event logs with three components of engagement. We further analyzed the helpfulness of engagement measurement in predicting grades. The results showed that there was a significant and moderate positive correlation between the behavioral, cognitive, and emotional engagement and quiz scores. A multiple linear regression analysis also showed that higher behavioral and cognitive engagement were related to higher quiz scores. Thus, this study performed three classification methods and found that the ANN method got the highest accuracy. This study applied sequential analysis to discover the difference between video watching behavior of the high and low quiz scores students. The results indicate that it is important to take into account all of the behavioral, cognitive, and emotional engagement in understanding user engagement during the system development process.

Keywords: system log analysis, user engagement, MOOCs, learning analytics, emotion

1. INTRODUCTION

User engagement is conceptualized as a need-based psychological state of users toward a system (how motivated they are) [1]. The state of user engagement can be observed by the behavior of user involvement and participation. Research has shown that user engagement has a positive effect on system success [1]. O'Brien and Toms [2] argued that **software requirement analysis** should move beyond usability to understand and design for more engaging user experiences. It is thus important to measure the engagement of users during the development process of a system.

Previous studies [3, 4] have shown that system log analysis might provide a way to capture user engagement over time. For example, Ramesh, Goldwasser, Huang, III and Getoor [5] measured the counts of posting and viewing to predict student engagement on a MOOC. Such work focused on the measurement of user engagement as behavioral participation (*e.g.*, the frequency of completing the tasks). However, the conceptualization of engagement should be defined as more than a sum of the individual behavioral component.

As noted by Fredricks, Blumenfeld and Paris [6], engagement is characterized as a multi-dimensional construct, referring to behavioral engagement, cognitive engagement,

Received September 9, 2017; revised November 6, 2017; accepted January 4, 2018.
Communicated by Don-Lin Yang.

and emotional engagement. Measuring engagement solely as the frequency of task participation may focus only on behavioral engagement and ignore the multifaceted nature of engagement [7].

Nevertheless, the challenge of using system logs to fully understand the user engagement lies in exploring the relationships between event logs and components of engagement. In this paper, the system logs collected from the Massive Open Online Courses (MOOCs) were used as the case study. We measured engagement through mapping event logs with three components of engagement. Further, we also analyzed the helpfulness of engagement measurement in predicting grades. This paper aims to stimulate a discussion on ways that the system log analysis can be used to better understand user engagement for the purpose of system design.

The organization of this paper is as follows. Section 2 describes the theoretical background which led to the prospective of learning analytics method. Section 3 details the experimental methodology. Section 4 shows the analysis results and discusses the findings. Lastly, Section 5 concludes this work and presents directions for further research.

2. COMPONENTS OF ENGAGEMENT

As defined by Trowler [8], "Student engagement is concerned with the interaction between the time, effort and other relevant resources invested by both students and their institutions intended to optimize the student experience and enhance the learning outcomes and development of students and the performance, and reputation of the institution."

As argued by Sinclair and Kalvala [9], log analysis about engagement in MOOCs overwhelmingly refers to student actions such as videos watched, quizzes answered and posts made on the forums. For example, Anderson, Huttenlocher, Kleinberg and Leskovec [10] selected six Coursera courses, including three machine learning courses and three probabilistic graphical models courses and analyzed the student learning behavior during the courses. Their findings showed that the pattern of student learning behaviors could be clustered into five groups: viewers, solvers, all-rounders, collectors, and bystanders. Moreover, Ferguson and Clow [11] selected four Coursera courses, including physical sciences, life sciences, arts and business and analyzed student learning behavior during the courses. They classified students into seven classification groups: samplers, strong starters, returners, mid-way dropouts, nearly there, late completers, and keen completers. Finally, Khalil and Ebner [12] take social aspects of information technology as the target course and classified students into four types of groups: dropout, perfect students, gaming the system and social, and learning participation. The aforementioned research focused on K-means clustering analysis, and did not consider student engagement. These measures represent the level of engagement on a single count variable, but do not reflect whether that the collected data can be interpreted as a benchmark for learning improvement.

This study uses the same concept to measure engagement. The justification for classifying different video log events to three components of engagement was based on Fredricks, Blumenfeld's theory [6] and Li and Baker's study [7]. Therefore, this section describes the video interaction events to identify the components of engagement (see Fig. 1).

When users interact with the video, the video system will generate five main events: play_video, pause_video, seek_video, speed_change_video, and stop_video. The play_video event will be generated when users start to play the video. When the video is normally played to the end, the seek_video event will be generated first and then is the stop_video event. When users pause a video, the video system will generate a pause_video event. When users fast-forward or rewind the video, the video system will produce the speed_change_video event.

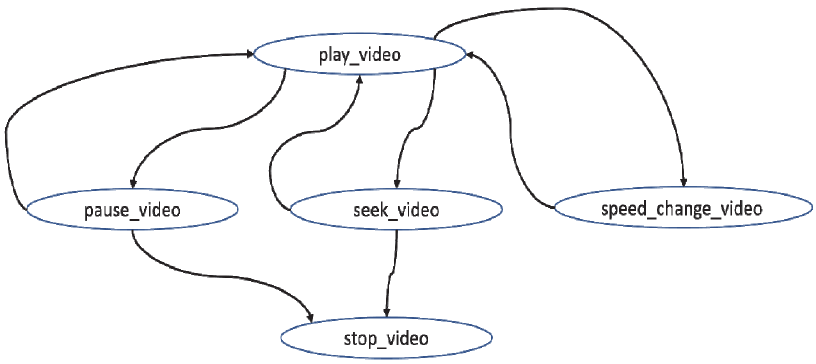

Fig. 1. The flow of video interaction events.

Behavioral engagement is referred to learning participation. Thus the stop video event is related to that learners have completed the video watching behavior. Cognitive engagement is referred to learning understanding, thus the pause and seek video events are related to that learners attempt to understand the unclear parts. Emotional engagement is referred to learning affection, thus change video speed event is related to not interest in the content of the video or unconsciously fast-forward/rewind the video.

2.1 Behavioral Engagement

We used one log event to indicate behavioral engagement. Stop video event: when the video player reaches the end of the video file and play automatically stops.

2.2 Cognitive Engagement

Cognitive engagement refers to the psychological investment in learning and relates to use self-directed strategies to promote one's understanding [6]. In this study, we measure cognitive engagement by two log events. Pause video event: when a user selects the video player's pause control. Seek video event: when a user selects a user interface control to go to a different point in the video file.

2.3 Emotional Engagement

Emotional engagement refers to student attitudes and student interest and values [6].

In this study, we measure emotional engagement by one log event. Speed change video event: when a user selects a different playing speed for the video.

3. METHODOLOGY

3.1 Sample

This study uses log-data from 155 MOOCs in OpenEdu platform. The courses ranged from 2014 to 2016. Eighty-four courses which provided quizzes and contained five video events (play video, pause video, stop video, seek video, speed change video) were considered in this study. 6433 students registered these 84 courses and log-data on 2697 students who attempted at least one quiz, and contained five video events after registration was used in this study.

3.2 Measurement

In this study, the behavioral engagement was measured by the times of videos students stopped each course. The cognitive engagement was measured by the number of pauses, and seeking in videos watched in each course. The emotional engagement was measured by the number of speed changes in videos watched in each course. In addition, we measured learning outcomes by students' total quiz scores (the sum of scores a student got on each quiz he/she attempted each course). Also, we generated an "is_passed" binary variable from calculating total quiz scores (if a student's total quiz score was greater than 60, is_passed was calculated as passed; otherwise, it was failed). 1125 students were labeled as passed, and 1572 students were labeled as not passed.

3.3 Analysis

We applied correlation to investigate whether there are relationships between three engagement components and total quiz scores. A K-means technique was employed to partition the students according to three engagement components. We also employed multiple linear regression techniques to predict student scores by video event variables. Furthermore, three classification methods (Support vector machine, Random forest, Artificial neural network) were performed and compared to predict "is_passed" variable. The 10-fold cross-validation is used to assess the accuracy and validity of classification models.

To further understand user engagement in MOOCs, this study employed Lag Sequential Analysis (LSA) to discover the difference of behavioral patterns of passed and failed students in MOOCs. LSA is used to test the statistical significance level of sequential correlation among video event variables [13]. The statistics of LSA involves a series of steps. The first step is to arrange the video watching events in chronological order. The second step is to conduct the following matrix calculations: (1) Sequential frequency transfer matrix; (2) Condition probability matrix; (3) Expected-value matrix. The third step is to calculate Z-scores using the calculated matrices. The sequential behavioral patterns with a Z-score higher than 1.96 ($p < 0.05$) were considered as significant. More importantly, this study only chooses the most popular video in each course to conduct LSA.

4. RESULTS

4.1 Correlation Between Video Events and Quiz Scores

As shown in Table 1, Spearman's rank correlations were calculated.

Table 1. Correlation between video events and quiz scores.

	Components	Events	*R*	*P*
Scores	Behavioral	Stop video	0.36	.00
	Cognitive	Pause video	0.42	.00
		Seek video	0.37	.00
	Emotional	Speed change video	0.17	.00

From Table 1, we confirmed significant and high overall positive relation between the behavioral engagement and quiz scores. Similarly, there was a significant and very high positive correlation between the cognitive engagement and quiz scores. Interestingly, we found that there was a significant and moderate positive correlation between the emotional engagement and quiz scores.

4.2 K-means Clustering Analysis

To better understand the relationship between three engagement components, we used K-Means to partition the students into three clusters according to their engagement variables. According to the results of practical experiment tests, the three groups will be more obvious to find the learner's behavior participation model. The cluster model revealed that Cluster 1 to 3 had 1257, 640, 800, and respondents, respectively. To draw the result of clustering, a principal component analysis was performed to reduce the dimensionality of features. As shown in Fig. 2, the result of K-Means clustering exists three clusters.

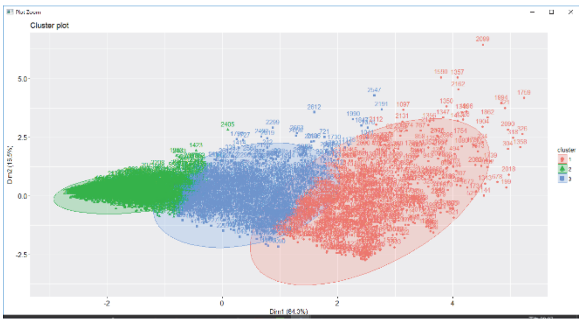

Fig. 2. The result of K-means clustering.

Table 2. The means of video events and quiz scores in three clusters.

Cluster	Stop	Pause	Seek	Speed change	Score
1	0.1	0.14	0.11	0.06	0.33
2	0.34	0.39	0.27	0.12	0.54
3	0.65	0.76	0.44	0.17	0.71

Table 2 indicates that Cluster 3 got the highest scores, the following was Cluster 2, and Cluster 1 got the worst scores. The frequency of video events, interestingly shows a similar pattern. Cluster 3 got the highest frequency in all video events; On the contrary, Cluster 1 got the lowest frequency in all video events. The findings may imply that high scoring students have a preference to watch videos at their own pace. They are frequently to stop, pause and seek the video play. However, the low score students are not.

4.3 Multiple Linear Regression Analysis

A multiple linear regression was undertaken to examine the variance in students' total quiz scores. Four predictors were loaded into the model using the Enter method. Table 3 shows that the model was able to explain 22.7% of the sample outcome variance (Adj.R^2 = .226), which was found to significantly predict the outcome, $F_{(4,2692)} = 197.767$, $p < .001$. Three of the predictor variables significantly contributed to the model. High frequencies of pausing video, seeking, and stopping video were related to higher quiz scores. The frequency of changing video speed did not contribute to variance. There was a medium effect size ($d = 0.29$).

Table 3. Multiple linear regression analysis of quiz scores.

	R^2	Adj.R^2	F	P	Constant
Model	.227	.226	197.767	<.001	.276
Predictor variable			Gradient	t	p
Stop video			.264	8.58	.000
Pause video			.143	4.08	.000
Seek video			.316	10.23	.000
Speed change video			.006	0.11	.912

4.4 Classification Analysis

Three classification methods (Support vector machine, Random forest, Artificial neural network) were performed and compared to predict is_passed variable. The Confusion matrices of three classification methods are shown in Tables 4-6. The elements in the confusion matrix indicate the correctly and incorrectly classified data for the passed and failed classes. These matrices are used to evaluate the performances of three classification methods. The overall classification precision is high (> 70%). The overall recall rate of the failed class is high (> 79%). However, the recall rates of the passed class for three classification method varied. The Random Forest method gets a lower recall rate of the passed class (56.44%) and the ANN method achieves highest one (66.93%). Moreover, the accuracy of SVM, Random Forest, and ANN are 73.79%,

Table 4. Confusion matrix of SVM results.

SVM	true failed	true passed	class precision
pred. failed	**1281**	416	75.49%
pred. passed	291	**709**	70.90%
class recall	81.49%	63.02%	

Table 5. Confusion matrix of random forest results.

Random Forest	true failed	true passed	class precision
pred. failed	**1356**	490	73.46%
pred. passed	216	**635**	74.62%
class recall	86.26%	56.44%	

Table 6. Confusion matrix of ANN results.

ANN	true failed	true passed	class precision
pred. failed	**1252**	372	77.09%
pred. passed	320	**753**	70.18%
class recall	79.64%	66.93%	

73.82%, and 74.34%, respectively. The ANN method gets the highest accuracy. In summary, the ANN method outperforms the SVM and Random Forest methods.

4.5 Sequential Patterns of Video Watching Behavior for the Past and Failed Users

Table 7 presented the Z-scores of all users. Tables 8 and 9 only presented the Z-scores of the passed users and the failed users. If the Z-score is more than 1.96, it illustrates p is less than 0.05. In other words, the video watching behavior from the row to the column is significant in sequence. Take Table 7, for example, the Z-score of 'play' row and 'seek' column is more than 1.96, and it indicates that the behavioral sequence from 'play' event to 'seek' event ('play' → 'seek') reaches continuity significantly. Based on the calculated Z-scores, the sequential patterns of the all users (Fig. 3), the passed users (Fig. 4) and the failed users (Fig. 5) are displayed. In these figures, each node represented a video event, and the nodes connected with solid lines and arrow suggested the sequential relationships between video events reached statistical significance.

In Fig. 2, the video watching behavior of all users reveals that these learners tended to play a video and seek the video (play → seek; seek → play), or pause the video (play → pause; pause → play). They may also play a video and change the video speed (play → speed), or pause the video after changing the video speed (speed → pause). Moreover, they may also stop a video, and pause the video (stop → pause; pause → stop), or play the video (stop → play). The overall pattern may exhibit that after playing the video, users tend to either (1) continually seek the video to go to a different point in the video file, or (2) pause the video or change the video speed, and then finally stop the video.

The Figs. 4 and 5 of the sequential patterns can be employed to further compare the differences between the passed- or the failed-users. The asterisks in Fig. 4 note the specific significant relationships observed for the passed users, which are not found in the failed users. Likewise, the asterisks in Fig. 5 are the significant relationships found in the failed users that are not significantly observed for the passed users.

As for the passed users (see Fig. 4), they tended to play a video through to the end, and the player stopped the video automatically (play → stop). Contrary to the passed users, the failed users tended to slow down or speed up a video through to the end, and then the player stopped the video automatically (speed → stop) (Fig. 5).

Table 7. Adjusted residuals table (Z-score) of video watching behavior for all users.

	play	seek	pause	Stop	speed
play	−110.85	24.38*	95.82*	−0.34	11.42*
seek	44.89*	41.65*	−60.44	−44.52	−30.08
pause	89.49*	−60.71	−60.78	59.74*	−19.20
stop	17.41*	−25.21	19.92*	−11.93	−7.39
speed	−24.94	−20.96	9.50*	1.37	94.48*

* $p < 0.05$

Table 8. Adjusted residuals table (Z-score) of video watching behavior for passed users.

	play	seek	pause	stop	speed
play	−77.91	25.89*	57.80*	2.02*	7.99*
seek	44.6*	13.62*	−39.72	−30.59	−19.47
pause	47.39*	−33.55	−38.11	40.41*	−13.04
stop	13.99*	−19.46	15.36*	−9.86	−5.96
speed	−17.24	−13.03	8.01*	−1.03	60.16*

* $p < 0.05$

Table 9. Adjusted residuals table (Z-score) of video watching behavior for failed users.

	play	seek	pause	stop	speed
play	−79.51	10.46*	76.91*	−2.95	8.1*
seek	21.38*	42.17*	−45.82	−31.99	−22.89
pause	77.59*	−51.27	−47.41	44.44*	−14.08
stop	9.78*	−15.45	13.09*	−7.24	−4.49
speed	−18.12	−16.38	5.64*	2.93*	72.99*

* $p < 0.05$

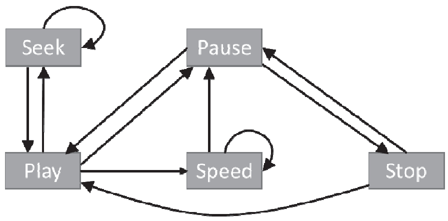

Fig. 3. The video watching behavior of all users.

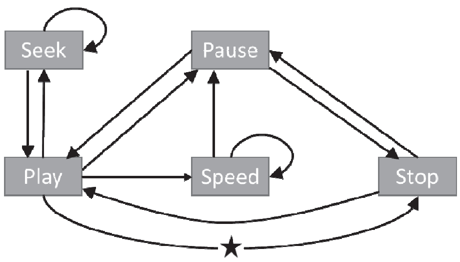

Fig. 4. The video watching behavior of the passed users.

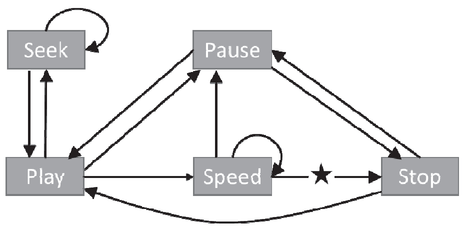

Fig. 5. The video watching behavior of the failed users.

5. DISCUSSION AND CONCLUSIONS

This study applied learning analytics to deconstruct user engagement by using log data of MOOCs. In other words, the engagement was first deconstructed into three components (behavioral engagement, cognitive engagement, and emotional engagement). And then, the importance of each component was evaluated by examining their helpfulness in predicting grades. The results of both correlation and clustering analysis of the three components and quiz scores shows that high scoring students had a preference to stop, pause, seek, and speed up/down the video play at their own pace. More specifically, they were more frequent in all of the three components than the low score students. For instructors, the findings may imply that they can use any of the three components to initially estimate student engagement. For system developers, the results indicate that it is important to take into account the behavioral, cognitive, and emotional engagement in help instructors understand user engagement during the system development process.

However, the results of both multiple linear regression and classification analysis indicated that only the behavioral and cognitive components were significantly contributed to the predicting model. In other words, high frequencies of pausing, seeking, and stopping video were precisely predicted higher quiz scores. The results echo Li and Baker [7] study showing that cognitive engagement has its unique contribution in predictting academic achievement.

Because we did not find the strong contribution of single emotional engagement to predict students' quiz scores, the sequential analysis of the video watching behavior was performed. The results showed that the passed users tended to play a video through to the end, and the player stopped the video automatically. Contrary to the passed users, the failed users tended to slow down or speed up a video through to the end, and then the player stopped the video automatically. This may be attributed to that emotional engagement should be further explored in more detail (*e.g.*, classify the speed change event into speedup and slowdown). Therefore, we suggest that platform developers can design functions such as the ranking of video watching to encourage students to activate their video learning.

One of the limitations of the research is that we report on a single platform (edX). Therefore, the findings in this research should be generalized with caution. Also, this research only considered students who attempted at least one quiz, and contained five video events after registration. Moreover, the score of quizzes that students did not take were zero. Therefore, this approach of prediction may be inappropriate for early dropout students. In addition, further research might conduct pre-test of learning to understand whether the students participating in the course had learned the results, examine the effect of degree of quizzes difficulty on the operations of the video, and analyze engagement according to other log records, such as chat room, discussion forum. The ethical issues of learning analytics such as data anonymization should also be concerned to conduct further research [14]. Most importantly, most of the current systems only focused on the measurement of user engagement as behavioral participation. We thus encourage further research to analyze user engagement from the perspective of multi-dimensional constructs for the purpose of better system design.

ACKNOWLEDGEMENT

This research is partially supported by the Ministry of Science and Technology, Taiwan, under Grant No. MOST 105-2511-S-006-015-MY2 and MOST-105-2511-S-035-006. Thanks to Dr. Nien-Lin Hsueh and OpenEdu team who develop the system and provide the open data.

REFERENCES

1. M. I. Hwang and R. G. Thorn, "The effect of user engagement on system success: A meta-analytical integration of research findings," *Information and Management*, Vol. 35, 1999, pp. 229-236.
2. H. L. O'Brien and E. G. Toms, "What is user engagement? A conceptual framework for defining user engagement with technology," *Journal of the American Society for Information Science and Technology*, Vol. 59, 2008, pp. 938-955.
3. K. Reed, "Users engage more with interface than materials at Welsh Newspapers Online Website," *Evidence Based Library and Information Practice*, Vol. 11, 2016, pp. 91-92.
4. L. Hakulinen, *et al.*, "The effect of achievement badges on students' behavior: An empirical study in a university-level computer science course," *International Journal of Emerging Technologies in Learning*, Vol. 10, 2015, pp. 18-29.
5. A. Ramesh, *et al.*, "Learning latent engagement patterns of students in online courses," in *Proceedings of the 28th AAAI Conference on Artificial Intelligence* 2014, pp. 1272-1278.
6. J. A. Fredricks, *et al.*, "School engagement: Potential of the concept, state of the evidence," *Review of Educational Research*, Vol. 74, 2004, pp. 59-109.
7. Q. Li and R. Baker, "Understanding engagement in MOOCs," in *Proceedings of the 9th International Conference on Educational Data Mining*, 2016, pp. 605-606.
8. V. Trowler, *Student Engagement Literature Review*, The Higher Education Academy, UK, 2010.
9. J. Sinclair and S. Kalvala, "Student engagement in massive open online courses," *International Journal of Learning Technology*, Vol. 11, 2016, pp. 218-237.
10. A. Anderson, *et al.*, "Engaging with massive online courses," in *Proceedings of the 23rd ACM International Conference on World Wide Web*, 2014, pp. 687-698.
11. R. Ferguson and D. Clow, "Examining engagement: analysing learner subpopulations in massive open online courses (MOOCs)," in *Proceedings of the 5th ACM International Conference on Learning Analytics and Knowledge*, 2015, pp. 51-58.
12. M. Khalil and M. Ebner, "Clustering patterns of engagement in massive open online courses (MOOCs): The use of learning analytics to reveal student categories," *Journal of Computing in Higher Education*, Vol. 29, 2017, pp. 114-132.
13. S. M. Wang, *et al.*, "Analyzing the knowledge construction and cognitive patterns of blog-based instructional activities using four frequent interactive strategies (problem solving, peer assessment, role playing and peer tutoring): A preliminary study," *Etr&D-Educational Technology Research and Development*, Vol. 65, 2017, pp. 301-323.

14. S. Slade and P. Prinsloo, "Learning analytics: ethical issues and dilemmas," *American Behavioral Scientist*, Vol. 57, 2013, pp. 1510-1529.

Ming-Chi Liu (劉明機) is an Assistant Professor at the Department of Information Engineering and Computer Science of Feng Chia University, Taiwan. His research interests include education technology, affecting learning, human/brain computer interface, nature language processing, and semantic analysis.

Chen-Hsiang Yu (余禎祥) is the Ph.D. Candidate at the Department of Information Engineering and Computer Science of Feng Chia University, Taiwan.

Jungpin Wu (吳榮彬) is an Associate Professor at the Department of Statistics of Feng Chia University, Taiwan. His research interests include sampling design, spatial statistics, experimental design, questionnaire design, and empirical process approach.

An-Chi Liu (劉安之) is the Chair Professor at the Department of Information Engineering and Computer Science of Feng Chia University, Taiwan. His research interests include network management, distributed system, and database system.

Hsi-Min Chen (陳錫民) is an Assistant Professor at the Department of Information Engineering and Computer Science of Feng Chia University, Taiwan. His research interests include software engineering, object-oriented technology, web technology, and distributed computing.