



## Research Paper

---

# Effect of online discussion forum in blended learning mode

Accepted 4<sup>th</sup> September, 2017

### ABSTRACT

The online discussion forums are popular platforms for information exchange and problem solving, which play an important role in blended learning mode, which is made up of SPOCs (Small Private Online Courses) and traditional classroom learning. However, very few studies have been published on the effect of online discussion forums in blended learning. In this study, we collected a comprehensive set of data from online discussion forum on the Experiment of Computer Composition course (a MOOC course). These data were classified based on the participants' interactive behavior and analyzed using the principal component analysis and regression analysis methods. Our results demonstrated that online discussion forums pose significant impact on blended learning and can provide valuable suggestions to improve learning outcome and progress.

Cuijiao Fu\*, Depei Qian, Qian Liu and  
Xiaopeng Gao

School of Computer Science and  
Engineering, Beihang University,  
Beijing, China.

\*Corresponding author. E-mail:  
fucuijiao@buaa.edu.cn

**Key words:** Online discussion, blended learning, factor analysis, regression, MOOCs.

---

### INTRODUCTION

In traditional classroom learning mode, teachers and students can communicate face to face, but this mode is restricted by time and space. Learning through MOOCs (massive open online courses) is not limited by time and space, but because of the involvement of a large number of participants, there is a great difference between their level of cognitive and receptivity. To a certain extent, it affects the evaluation of learning in MOOCs. SPOCs are special MOOCs for class in higher education (Martin, 2012), the number of participants is small and the level of the learner is consistent. SPOCs retain the advantage of MOOCs' as online video can be reused, and the online discussion forum is not restricted by time and space. Hence, it is a good choice to combine classroom learning and online learning. Broadly, blended learning is a learning activity that uses different methods and locations to achieve learning objectives (Howard et al., 2006). A narrower definition is the combination of traditional classroom learning and e-learning (Melton et al., 2009).

Blended learning mode composed of traditional classroom and SPOC learning, online discussion forums play an

important role. The forum is not only a platform for knowledge-sharing and problem-solving, but also highlight questions of participants that can be gathered, and then teacher especially explain them in the classroom. It is a useful complement to traditional learning. There are numerous researches about online discussion forum and they are limited in several ways. Although the role of the forum has been shown in e-learning (Figueira and Laranjeiro, 2008; Abel et al., 2010; Jacob and Sam, 2008; Brinton et al., 2014), it turns out that the number of participants as the course progresses gradually decreased in MOOCs forum (Brinton et al., 2014). The factors on the use of forum have been studied (Jacob and Sam, 2008; Seliaman, 2013), such as social influence and online social skills, but online and offline interaction is not considered. The existing research methods, such as Social Network Analysis methodology (Abel et al., 2010), content analysis technique (Nesterova and Azpeitia, 2010), business process model and mining techniques (Wang et al., 2014), network structure (Hecking et al., 2015), are effective methods; however,

they have little consideration on interaction online and offline in blended learning mode. In addition, Post in the forum is text, and it is inconvenient to collect statistic data, but some visualization diagrams (Figueira and Laranjeiro, 2008) or interface (Zhu et al., 2015) is appropriate for domain oriented. To the best of our knowledge, there is little research regarding the effect of online discussion forum on arranging learning content and evaluating learning performance and difficulty.

In this study, we designed some Python programs to collect data from the MOOC online forum, and classified the data according to the participants' interactive behaviors. Based on the data, we studied the relationship between the number of posts and the final score in the forum. The correlation coefficient of the final score and the participants' interactive behaviors were analyzed using principal component analysis and regression analysis methods.

The main contributions of this paper are:

- 1) The positive influence on the learning of online discussion forum was studied by applying the principal component analysis in the blended learning mode;
- 2) We proposed a new model of participants' behavior in online discussion forum and final score;
- 3) We studied the significant feedback in learning content and progress in online discussion forums in blended learning mode.

Our experimental data is the real data on Experiment of Computer Composition course in Beihang University. The experimental results showed that participants' behavior in the forum and final score has linear correlation, and the number of participants' questions and discussions in the forum is proportional to the final score. And hotspot can reflect the learning difficulty, and we verified the relation with the performance of each learning contents.

The rest of this study is organized as follows: Related works are summarized and thereafter, the research method is provided. This is followed by results and discussion and finally, conclusion of the study.

## RELATED STUDIES

Here, some notable previous studies are reviewed.

### Characteristics of blended learning mode

In traditional learning mode, teachers and students can communicate face to face; it is easy to be restricted by time and space. Thus MOOCs have this advantage. Video resources and online discussion forum can be visited at any time (Brecht, 2012). SPOCs have the similar structure with MOOCs (Fox, 2013). Blended learning consisted of

learning in traditional classroom and learning through SPOCs (Melton et al., 2009). In the blended mode, the students can learn video resources online by themselves, and they can use the online forum to discuss with teachers or others. The universal questions of the students can be gathered and then the teacher, especially explain them in the classroom (Jacob and Sam, 2008). Although the aforementioned provide the complementary advantage of blended learning, specific mode still needs to be studied.

### The role of online forum

The role of online discussion in improving the ability of problem-solving and the ability of critical thinking was discussed in a comparative study (Jacob and Sam, 2008). Some importance of the forum was outlined for e-learning (Abel et al., 2010). There is a relationship between Students' behavior in the forum and the final score, and teacher can predict the final score of students according to behavior in the forum (Romero et al., 2013; Marcus and Albrecht, 2015). The number and type of questions in the forum were used to modify the version of the blended model (Ferri et al., 2015). Online discussion has more participants than the in the classroom in the course 'linear algebra' (Lin et al., 2015). It has better performance of learning according to the social network of students in online forum forms discussion group (Junus et al., 2014). Although the role of the forum has been shown, it turns out that the number of participants as the course progresses gradually decreased in MOOCs forum and some discussion were unconcerned with the course (Brinton et al., 2014).

### Factors affecting the use of the Forum

The personalized access strategies were studied to help people who visit the forum to find relevant information (Abel et al., 2010). Factors, such as social influence and online social skills, are the main factors that affect the use of the forum and share knowledge (Seliaman, 2013). Some emotion factors and motivation was studied on knowledge sharing in the forum (Lin et al., 2015). Although some factors on the use of the Forum have been studied, interaction is considered online and offline.

### Existing research method about online discussion forum

Interactive behavior of participant was analyzed using Social Network Analysis method (Figueira and Laranjeiro, 2008; Abel et al., 2010; Wassermann and Faust, 1994). A content analysis technique was adopted to study the

characteristics of participation and draw a result from linear correlation analysis between numbers of participations in each category (Nesterova and Azpeitia, 2010). The business process model and mining techniques were employed to analyze the behavior of online knowledge sharing (Wang et al., 2014). The network structure was adopted to analyze the interactive pattern and feature of the social structure of online forum (Hecking et al., 2015). Although the aforementioned method performs well, they have little consideration for interaction online and offline in blended learning mode.

The asynchronous online interactive behavior was analyzed and described in the forum through interaction diagrams – Graph (Figueira and Laranjeiro, 2008). A model was put forward for domain oriented course and automatic generation of visualization of the execution state to tackle the shortcomings of the text in MOOC forum (Zhu et al., 2015). However, Post in the forum is text, and it is inconvenient to collect statistic data.

Hence, we are motivated to study the effect of the online discussion forum in blended learning mode.

## METHODOLOGY

Here, we first describe set data and thereafter, the analysis method. We analyzed the discussion forum posts of Beihang in MOOCs: Experiment of Computer Composition, a free MOOC on the edX platform. We selected this course because it was a domain oriented course adopting blended learning mode. We release video resource and design task each week in MOOC, It is approximately 48 h lecture in total.

A total of 390 people posted in the forum, students discussed with others and teachers and TAs, the discussion was active throughout the course.

### Data set

It is inconvenient to collect and analyze the data of participants because posts are mostly text in the forum. We design some Python programs to collect data from the online forum in our MOOCs, and mainly obtain the number and date of each student' posts. The data were classified as Comments, Questions Discussions, Votes, Drosed, Words, Points, etc. Questions are the posts of participant's proposing questions in the forum. Discussion involves the posts of participant discussing the questions. The Comments are the posts of commenting answer to the question. Votes are the number posts that agree with answers. Drosed is the situation of answer adopted by others. Points are given as a bonus according to performance in the forum. Score is the final score of this course, not including Points. Date is the day in which post is issued in the forum.

According to the result of the Python programs, the statistics data about Comments, Discussions, Questions, Votes, Droseds, Points, Words and Dates were obtained.

Figure 1 shows that it is the number of daily posts in the forum from the start to end of the course.

## Analysis of the method

The relationship between the final score and participants' interactive behaviors was analyzed using principal component analysis method and then the regression analysis methods were applied to the model participants' behavior. Finally, we compared the number and date of posts with the arrangement of learning content and students' score of each test.

## RESULTS AND DISCUSSION

Based on the dataset described previously, we analyzed the data by applying principal component analysis (PCA) and linear regression, and drew a series of interesting conclusions.

### Model of participants' behaviors and the final score

The participant's behaviors were divided into five types: Comments, Discussions, Questions, Votes, and Drosed. They were used as independent variables. The score was used as the dependent variable. We analyzed the relationship between the independent variable and the dependent variable using the regression analysis and PCA. According to the principal component analysis, the variance of Discussions accounts for 95.5% of the population variance. The results of the regression analysis showed that the top three factors are: Discussions, Questions, and Votes.

These data were information on students' questions and discussions about learning content and learning progress, and they were mainly positive or negative feedback of learning arrangements. Students who gave feedback information about learning arrangements and learning difficulty had better overall understand of the course, and it was the students' intuitive response to provide the current learning content and indirect feedback of the acceptability. This provides the basis for teachers to adjust the learning content and progress.

### The hot spots and the difficulties of learning are consistent

By comparing Figure 1 and Table 1, it can be observed that the number of discussions begin to increase from the next

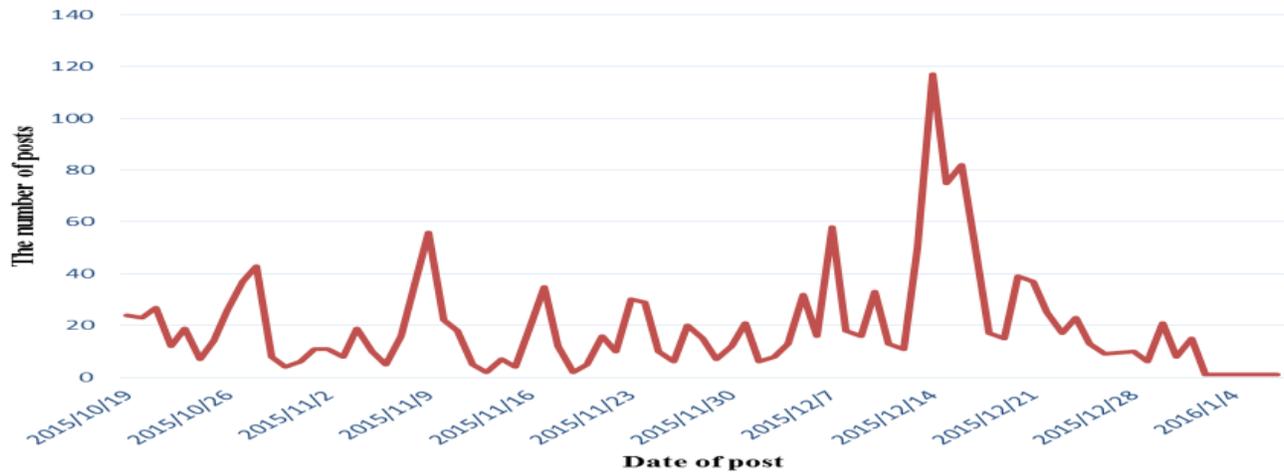


Figure 1: The number of daily posts in the forum.

Table 1: Schedule of design task released.

S/No.	Date	Is or is not hotspot
P0	11/10-17/10	-
P1	17/10-24/11	-
P2	25/11-31/10	The fourth hotspot on Nov. 26
P3	1/11-7/11	-
P4	8/11-14/11	The third hotspot on Nov. 9
P5	1/11-21/11	The fifth hotspot on Nov. 16
P6	22/11-28/11	The sixth hotspot on Nov. 23
P7	29/11-5/12	-
P8	6/12-13/12	The second hotspot on Dec. 7
P9	14/12-1/1	The first hotspot on Dec.15

day after the task was issued. When the students were given more questions about design, they discuss more. From Figures 1 and 2 and Table 1, the top 6 frequent discussions can be seen: P9, P8, P4, P2, P5, and P6. P9 and P8 were comprehensive and difficult designs, more discussions about them were normal. Students were mandated to complete the design of p0 to p7 that was basic task. The discussion mainly focused on P4, P2 and P5. The instruction format was used in P2, a design language, technology of debugging and simulation. P4 was a comprehensive design where the Verilog language was

first time used to design CPU. More discussion will appear when people encounter a new thing which is consistent with the law of cognition. Students first used the assembly language to program in P2, and there were more discussions about P2 because there was no practice about assembly language before p2. From Table 1, it can be seen that the proportion of students that obtain A level began to reduce from P2. According to these data, we adjust the plan of learning by adding a task regarding assembly language before p2. It was the first time the Verilog language is used to design pipelined CPU in P5,

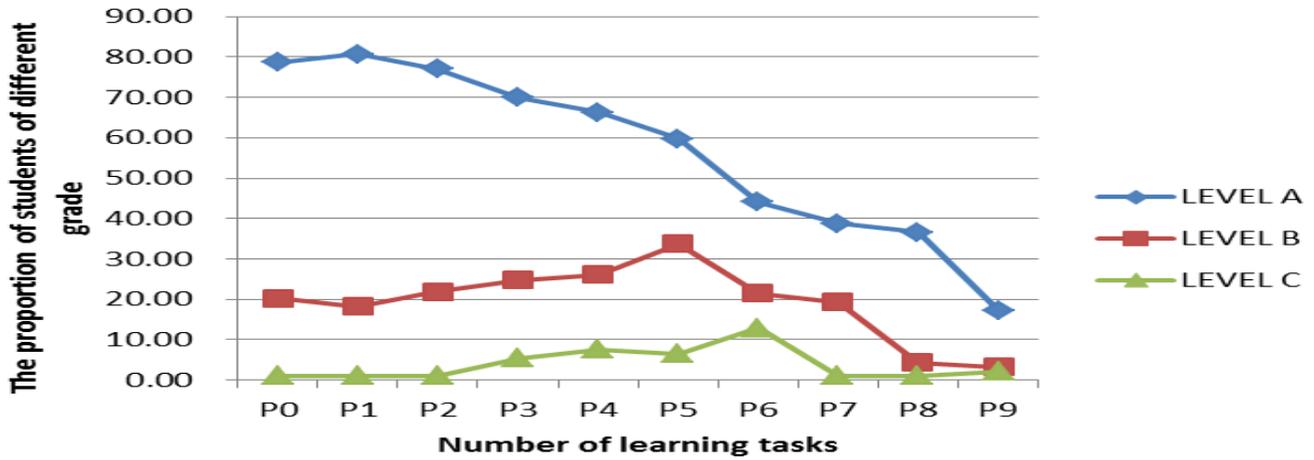


Figure 2: Score of each of the design tasks.

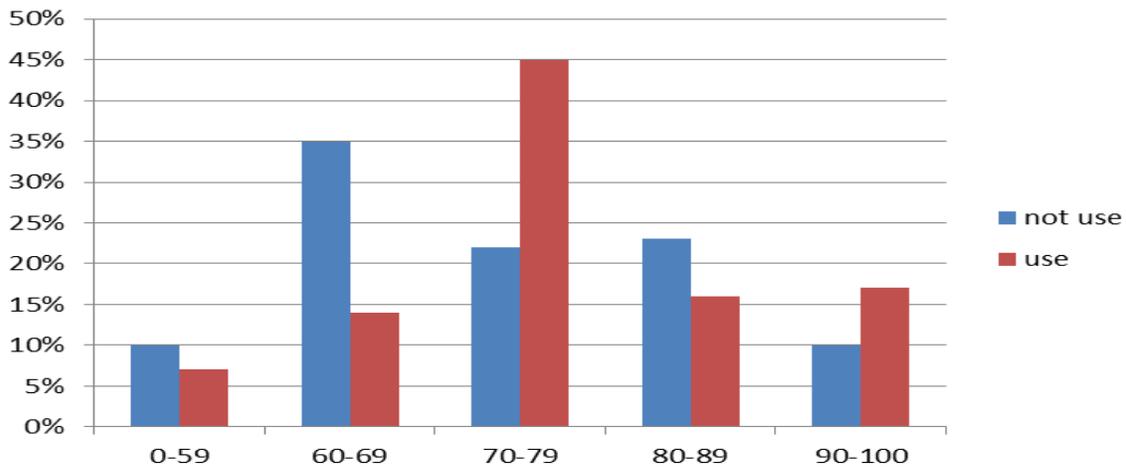


Figure 3: Comparison of student's performance with Use and non use of the forum.

and there were relatively more discussion about P5. From Figure 5, it can be concluded that there were more discussions when students, for the first time, engaged in the new pattern. More video resources were afforded online through MOOCs and the teacher gave further explanation in the classroom to help students according to the hotspot of the forum.

**Comparison of student's performance with the use and non-use of the forum**

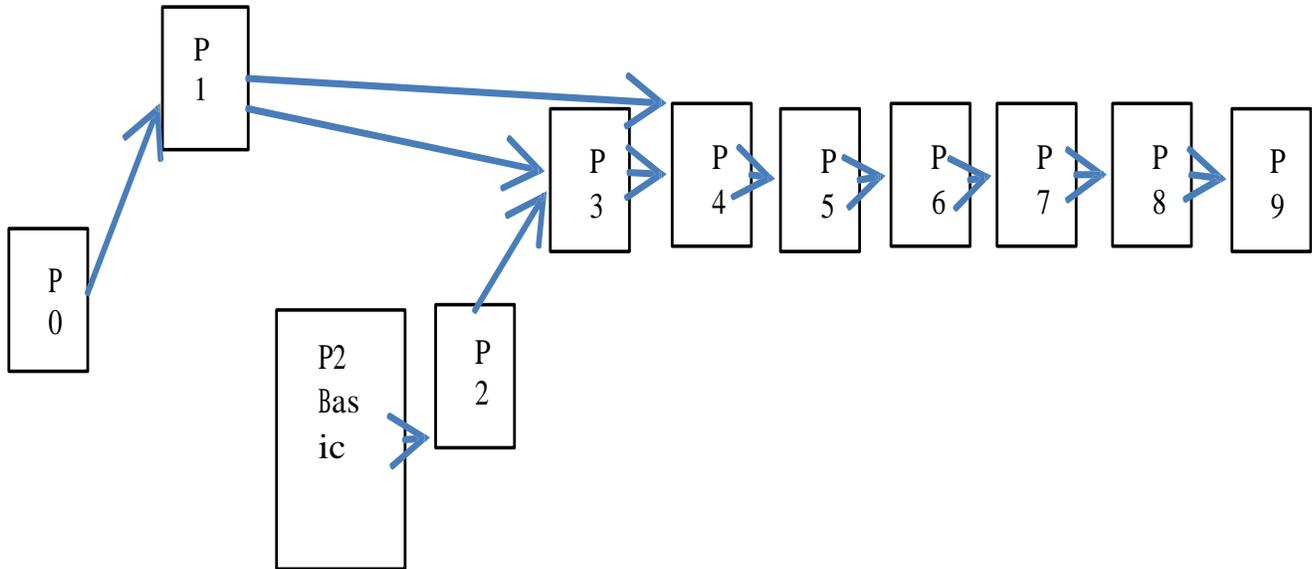
In Figure 3, the y-axis is the proportion of students, while the x-axis is the score segment. From the figure, it can be seen that the excellent rate significantly improved by 7%, the passing rate increased by 3%, the failure rate significantly declined, and the number of student to complete the very difficult experiment increased by 23%. The overall performance has greatly improved.

**Appropriate adjustment of learning schedule**

Figure 4 shows the relationship between the various parts of our curriculum. Figure 5 shows preparation as to change of learning of schedule. By adding a learning task associated with P2 before p2, the particular effect would be verified through relevant data of the course in 2017.

**Conclusion**

In this study, we design some Python programs to crawl the data of online discussion forum in our MOOCs of Computer Organization and Design, and classified the data according to the participants' interactive behaviors. The data were analyzed by applying principal component analysis and regression analysis. Our results showed that online discussion forums pose significant impacts on learning; the number of posts and final score were significantly related;



**Figure 5:** The relationship of each learning content adjusted.

and the questions of the posts and other factors in participants' behaviors are linearly correlated with the final score. In addition, online discussion forums could provide valuable suggestions for the adjustment of learning content and progress. Therefore, some incentive measures should be adopted to encourage students to participate actively in the forum. The forum provides a platform for students to learn from passive to active and provide feedback to teachers. It is a key link in mixed learning mode and an important way to link teachers and students. Through the forum, the students actively asked and answered the questions and such as students can guide the learning trend in forum and play a key role in learning. Therefore, future research on the impact factors of performance of students is proposed

## ACKNOWLEDGEMENTS

The authors are grateful to the Experiment of Computer Composition teaching team for their help in designing cases and performing researches. The financial support from China Scholarship Council is gratefully acknowledged. Anonymous reviewers are also acknowledged for their thoughtful feedbacks.

## REFERENCES

- Alan WG, Harry JW, Jiexun L, Weiguo F (2014). Mining Knowledge Sharing Processes in Online Discussion Forums, 47th Hawaii International Conference on System Science pp. 3898-3907.
- Álvaro Figueira, Joanne Laranjeiro (2008). Work in Progress – iGraphs for Characterization of Online Communities. 38th ASEE/IEEE Frontiers in Education Conference(FIE) pp.F4E-13 - F4E-14
- Bonnie F, Joyelle M, Ann W, David M (2015). A Feedback-Based Approach for Evolving a Blended Class Model for Large Enrollment, Multiple Section Circuits Courses.
- Christopher GB, Mung C, Shaili JL, Zhenming L, Felix MFW (2014). Learning about Social Learning in MOOCs: From Statistical Analysis to Generative Model, IEEE TRANSACTIONS ON LEARNING TECHNOLOGIES, VOL. 7: (4):346-359.
- Elena Nesterova ricardo Ulloa Azpeitia (2010). Assessment of learning activities in discussion forums online, IEEE EDUCON Education Engineering pp.487-492.
- Fabian Abel (2010). Ig Ibert Bittencourt, Evandro Costa, Nicola Henze, Daniel Krause, and Julita Vassileva. Recommendations in Online Discussion Forums for E-Learning Systems, IEEE TRANSACTIONS ON LEARNING TECHNOLOGIES 3:165-175.
- Fox A (2013). –From moocs to spocs, Communications of the ACM, 56:12:38–40.
- Fred G, Martin (2012). Will massive open online courses change how we teach?. Commun. ACM 55, 8 pp.26-28.
- H. David Brecht H (2012). –Learning from online video lectures, || J. Info. Tech. Edu. Innovations in Practice, Vol 11
- Howard L, Remenyi Z, Pap G (2006). –Adaptive Blended Learning Environments, || 9th International Conference on Engineering Education T3K-11.
- Joyce Z, Jeremy W, Mitchell G, Jeffery W, Renan Z, Philip JG (2015). Toward a Domain-Specific Visual Discussion Forum for Learning Computer Programming: An Empirical Study of a Popular MOOC Forum, IEEE Symposium on Visual Languages and Human-Centric Computing (VLHCC).Kasiyah J, Lia S, Heru S (2014). Social Cognitive, Teaching, and Metacognitive Presence in General and Focus Group Discussion: Case Study in Blended e-Learning Linear Algebra Class.
- Marcus K, Albrecht F (2015). Predicting Students' Success Based on Forum Activities in MOOCs, The 8th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Application Warsaw, Poland 925-928.
- Melton BF, Bland HW, Chopak-Foss J (2009). Achievement and Satisfaction in Blended Learning versus Traditional General Health Course Designs, || in International Journal for the Scholarship of Teaching and Learning 3:1-26.
- Mohamed ES (2013). Exploring the Adoption of Online Discussion Forums for knowledge Sharing and Social Relations among Virtual Communities.
- Romero C, López MI, Luna JM, Ventura S (2013). "Predicting students'

final performance from participation in on-line discussion forums  
Computers & Education, 68:458-472.  
Second European Network Intelligence Conference p.114-121.  
Seibu Mary Jacob, Hong Kian Sam (2008). Effect of Online Discussion  
Forums to Improve Mathematics Problem Solving and Critical  
Thinking – A Comparative Study 366-370.  
Tobias H, Irene-Angelica C, Ulrich HH (2015). Analysis of User Roles and  
the Emergence of Themes in Discussion Forums.  
Wassermann S, Faust K (1994). –Social Network Analysis: Methods and  
Applications. || Cambridge University Press, New York.  
Yu-Chun L, Yi C, Timothy K (2015). Shih and Min-Te Sun. Using Open edX  
Forum for Supporting Collaborative Learning 44th International  
Conference on Parallel Processing Workshops pp.20-26.

Yu-Chun L, Yi C, Timothy KS Min-Te S (2015). Using Open edX Forum for  
Supporting Collaborative Learning, 44th International Conference on  
Parallel Processing Workshops pp20-26.

**Cite this article as:**

Fu C, Qian D, Liu Q, Gao X (2017). Effect of online discussion forum  
in blended learning mode. Acad. J. Educ. Res. 5(10): 377-383

**Submit your manuscript at**

<http://www.academiapublishing.org/ajer>