Can Participation and Procrastination in Discussion Forums Predict Project Performance in Computer Science Courses?

Jaebong Yoo, Erin Shaw, Jihie Kim, Information Sciences Institute, University of Southern California
Email: jaebong@isi.edu, shaw@isi.edu, jihie@isi.edu

Abstract: This paper presents an analysis of project grades and discussion forum participation. In a study of 1,593 messages, it is shown that only 13% of the project grades can be explained by forum postings. However, by applying a new metric, Average Posting Time to Deadline, about 27% of the grade can be explained. The strength of the study lies in its use of targeted project-based forums and grades and the new metric

Introduction

Asynchronous discussion forums can encourage collaborative learning and promote discovery-oriented activities. Researchers have sought to understand the relationship between forum participation and achievement by discourse analysis. The most common way of evaluating the effect of forum participation on achievement is to correlate message incidences with course grades (Davies & Graff, 2005). Traditional discussion forum metrics used in quantitative studies include the number of the total messages, initial postings and replies, number of messages read, and response times from previous messages (Palmer, Holt, & Bray, 2008). The results indicate that the relationship between participation and achievement is inconsistent and may depend upon measuring methods. This paper explores the influence of participation on student achievement in a targeted context. First, participation is considered with respect to project-based forums that are used by students to seek answers to questions about the current project. Second, student achievement is considered with respect to project grades. Finally, traditional forum metrics are considered in conjunction with a new metric based message posting times relative to the project due date. This number is referred to as average posting time to deadline, or APTTD. The sections of the paper are laid out as follows. In the following section, the methodology of the study is explained; in the next section, the results of the study are presented; and in the final section there is a concluding discussion.

Methodology

The study took place in the context of an Operating Systems course in the Computer Science Department at the University of Southern California. The same instructor has taught it for the past 15 semesters. The course discussion board consists of "Projects", "Lectures", "General Nachos Questions", "TA Office Hours" and "Humor" forums. For this study, we consider student project grades and participation in corresponding project-related forums only, and not overall forum participation and final grade. In this respect, our approach is distinctly different from other approaches (e.g., Davies & Graff, 2005; Palmer et al., 2008). Participation on the discussion board is optional but can be used to influence a borderline grade at the instructor's discretion.

Sample

Student data, project grades, and online discussions were collected for three fall semesters of the course. The number of registered students in each semester was 85, 119, and 124, for the years 2006, 2007, and 2008, respectively, and of the 215 students who participated in discussion forums posted 1,593 messages: initial posts (693) and replies (900). Project grades were transformed into normalized scores (Z-score) to avoid potential errors due to variability such as time, instruction, teacher, score, and other factors. For three years, the discussion board was observed to inspect the on-line activities of the participants because the quantity of student posts is a common and usually straightforward measure to gather from learning management system (Bliss & Lawrence, 2009).

A New Metric

We assumed that the longer average posting time to deadline students have, the better performance they have. The average posting time to project p's deadline of student s is denoted $APTTD_{ps}$. The $APTTD_{ps}$ is defined as follow:

$$APTTD_{ps} = \left\{ \sum_{p=1}^{P} \sum_{m=1}^{M_{ps}} \frac{T(end_p) - T(message_{sm})}{T(end_p) - T(start_p)} \cdot \frac{1}{M_{ps}} \right.$$

where $T(end_p)$ is the deadline of project p, $T(start_p)$ is the start date of project p, and $T(message_{sm})$ is the time stamp when student s posts the *m-th* message on the project p forum of the discussion board.

Results

A Pearson correlation coefficient was computed to assess the relationship between the normalized project grades (Z-score) and the number of initial/reply/total, and APTTD. Table 1 summarizes the results. Overall, there was a highly significant, positive correlation between Z-score and APTTD. This indicates that students who posted messages earlier from the deadline tend to get better grade than others.

Table 1: Inter-correlations between variables among participants.

Correlations	Total posts	Initial posts (IP)	Reply posts (RP)	APTTD	Z-score
Total posts	1	.862** .000	.974** .000	.512** .000	.155* .023
Initial posts (IP)	.862** .000	1	.724** .000	.653** .000	.147* .041
Reply posts (RP)	.974** .000	.724** .000	1	.456** .000	.151* .031
APTTD	.512** .000	.653** .000	.456** .000	1	.307** .000
Z-score	.155* .023	.147* .041	.151* .031	.307** .000	1

$N = 215$; $^{*}p < .05$; $^{**}p < .01$; Pearson Correlation Sig. (2-tailed).

Multivariate linear regression analysis was conducted with Z-scores as the dependent variable. All other known variables were initially introduced as independent variables in Table 1. The result shows the coefficients of the regression model and their significance. The regression equation model is: $Z - score = -.210 - .096IP + 0.028RP + 1.01APTTD$. An analysis of variance test suggests that the regression model is significant, $F(3, 211) = 8.094$, $p < 0.001$, although the model predicts only 10.3% of the variation on Z-score ($R^2 = 0.103$). When the number of total message is larger than 12, the result indicated the predictor APTTD explained 27% of the variance ($R^2 = 0.27$, $F(3,30) = 3.60$, $p < 0.05$). It was found that the APTTD predicted the Z-score of high participating students 16.7% more than that of all participants (only considering the total posts, ($R^2 = 0.13$, $F(1,32) = 4.59$, $p < 0.05$). The APTTD is highly correlated to forum participation frequency, about 27% of the grade can be explained.

Conclusion

Our results indicate that *average posting time to deadline*, or APTTD, is statistically significant and highly correlated to discussion forum participation frequency. It is also observed that there is a significant, positive correlation between project grades and APTTD. For more accurate assessment, other latent factors should be considered in the future with a semantic-based analysis. Our results may provide hints on best instructional practices such as promoting earlier resolution of project-related questions on the discussion board and preventing rushes.

References

Bliss, C. and B. Lawrence (2009). "From posts to patterns: A metric to characterize discussion board activity in online courses." Journal of Asynchronous Learning Networks 13(2): 15~32.

Davies, J. and M. Graff (2005). "Performance in e learning: online participation and student grades." British Journal of Educational Technology 36(4): 657-663.

Palmer, S., D. Holt, & S. Bray. (2008). "Does the discussion help? The impact of a formally assessed online discussion on final student results." British Journal of Educational Technology 39(5): 847-858.

Acknowledgments

The authors would like to thank USC Senior Lecturer Michael Crowley for his assistance and the members of the PedTek team. This work was funded by the National Science Foundation, under a CCLI and CISE awards #0618859, #0917328.