Demographic Characteristics of Students Who Do or Do Not Post in an

Undergraduate Engineering Online Discussion Forum

Authors

Affiliations

1. Introduction

The current schooling norm is that students have access to help during the class hours and limited office hours. They are not in touch with their instructors or classmates outside of the class. The introduction of an online discussion forum for a class helps students connect with their peers, teaching assistants (TAs) and even instructors outside of class and office hours. Such a discussion forum provides a public space where ideas can be exchanged, questions can be shared, and information can be stored. Students can help each other clarify doubts with fewer time and space limits. These forums can then be considered to be a social network, where Social Network Analysis (SNA) and statistical techniques can be applied for studying the interaction of the students and the level of students’ participation.

In this work, we apply SNA and statistical techniques to an online discussion forum to determine how participation on the forum relates to the students’ demographic characteristics. This statistical analysis will help instructors better understand relations between students’ participations and their backgrounds.

1. Research Questions

The following research questions guide this work:

1. How do students’ demographic characteristics relate to discussion forum participation?
2. Among students who participate, how do students’ demographic characteristics relate to the number of their posts?

We expect that there is a strong relationship between students’ demographic characteristics and participation level. Among the demographic characteristics we investigate, we expect female students to have higher rates of participation in the discussion online forum compared with male students because (xxxx)We also expect that international students feel less comfortable discussing their questions in such a public, online space, so we expect them to be underrepresented in forum participation. Since the course material is closest to mechanical engineering students, we expect more ME students to be active in online discussion than students from other majors.

Regarding to the second question, we expect that the density of female students of higher number of posts should be over male students. Our hypothesis is based on previous work, more details will be discussed in the following literature.

1. Related Work
2. What are Online Discussion Forums?

An online discussion forum is a virtual environment where students can communicate with their peers and instructors without temporal or geographical barriers [3][9]. Recently, academic online discussion forums have become more popular among undergraduate courses due to their asynchronous nature [10]. Research studies have shown that students participate in meaningful discussions in online discussion forums. Other studies suggest that students’ engagement in discussion forums has a strong relationship with their academic performance. For example, Minichiello and Hailey (2013) have suggested that engaging in the online learning forums is highly correlated to students’ performance in first-year calculus[3]. Cheng et al..[1] suggest that the implementation of an online discussion forum benefited students because they demonstrate that online collaborative learning reflected improved academic outcomes.

The level of engagement in discussion forums can be influenced by teachers’ involvement and the possibility of receiving course credits (Garrison et al., 1999, 2001, 2003; Guzdial & Turns, 2000; Taradi & Taradi, 2004). However, the studies that looked at external motivators for student engagement with the forum focused more on the text of the students’ post; they might have overlooked the effect of other forms of participation.

Most studies regarding students’ engagement in online discussion forums focus on the motivational influences on students’ academic behavior. Only limited studies have focused on the relationship of students’ demographic characteristics and their level of engagement. For example, in a case study of a social online discussion forum, Yeh et al..[3] studied the participation of users based on their gender, nationality, and age. They found that more than half of the users were female (53%), and half of the users were between the ages of 16-30. Ke et al..[4] complete a study on how students’ ethnicity affects their level of participation in an online learning forum and concluded that students’ minority status correlated with lower satisfaction with the web-based, distance-learning class.

Most of the demographic related studies on academic forums have examined public learning or social forums influence, instead of a discussion forum only focusing on a class. In our study, all participants have the chance to meet with peers and instructors on campus. Therefore, it is more interesting to study if students’ demographic characteristics make a difference in the level of engagement. Here the discussion forum was incorporated as part of the class website for a sophomore-level dynamics and vibrations class that incorporated active, blended, and collaborative learning strategies. Students used the forum to communicate with their classmates regarding homework questions, example problems, or other course content.

1. Social Network Analysis (SNA)

Social Network Analysis (SNA) is the use of network theory to analyze social networks. This turns the network to a visualized social network diagram, where nodes represent the individuals and edges depicting the relationship between individuals. In this study nodes are represented as circles and edges as lines. Statistical computation is aiming at the design of algorithm for implementing statistical methods. It is the application of software open-source to statistics nowadays. In this study, we especially use bar graph to compare results from different variables.

Previous studies showed that SNA is a common tool to evaluate students’ online discussion performance. He[2] uses SNA software to visualize students’ online discussion participation networks. That paper evaluates individual student’s online discussion performance based on SNA pagerank and in-degree out-degree centralities. Here we use SNA to quantize the number of posts each person posts online.

1. Research Method

In this section, we depict the sample data and the features we used for analysis.

* 1. Data Collection

This mechanical engineering(ME) course forum assigns each post a unique id. As long as students write a message online, their posts are recorded under the student’s id. Then, part of the post related information is extracted into a csv file by a software tool. This includes the person id, which person this post is communicated with, and post time. In on our discussion forum situation, the instructor always starts a new thread by posting homework problems or lecture notes, and students post their doubts or opinions under instructor’s posts. Since we are only interested in students’ online social behavior, the posts started with an instructor has been removed.

//Thread and posts

* 1. SNA

There are several SNA methods that one can consider when analyzing a social network. Particularly in this paper for the context of evaluating student’s online discussion performances, we will use out-degree centrality to count the number of messages sent by a student. If a student acquires high numbers in out-degree, it indicates that he or she is more active in sending messages to others. There were 954 unique students participated in our course’s online discussion board and total number of messages was 1861 in three academic semesters from 2015-2016. These included students’ posts and replies, but initial posts from instructors and teaching assistants were not counted.

* 1. User Profiles

We collected all unique user profiles from surveys. We retrieved the following user profiles data:

* + - * User Personal data: Gender, Nationality, Ethnicity, Major, and ADCI(Advanced demographic C)

Gender is divided by Female and Male. Due to this study happened in the USA, nationality is divided by international students and Americans(citizens). By ethnicity, Americans are categorized by White American, Asian American, Hispanic or Latino, African American, Unknown, and 2 or more races. In this case, we did not dig into international students’ ethnicity, as international students only occupy a very limit amount of total. As mentioned previously in online discussion forum session, this class is about dynamics and vibration, which is required for mechanical engineering(ME) students, selective for Agriculture Engineering(AAE), Nuclear Engineering(NE). Students from Multidiscipline Engineering Study, Biomedical Engineering can also take this class for credits, so I summarize them as one Other category.

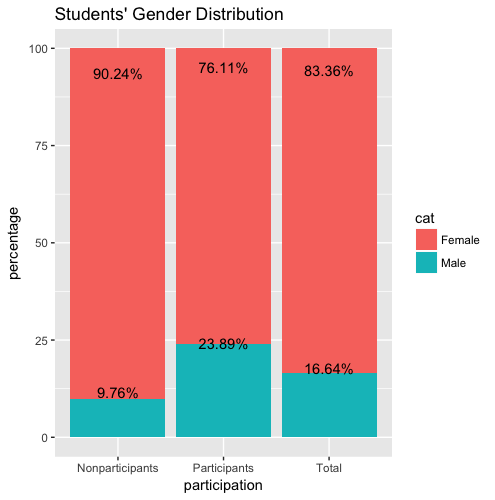
Models based on demographic attributes are important since it helps to determine the connectivity based on social attributes[7]. We got all user personal data based on students’ answers to a survey. It is imported as a raw data. Then, a filter is applied based on consented status. We totally got 739 consented students, among which 360 are participants, and 379 are nonparticipants. Three data frames of user personal data plus out-degrees are generated via R. One is of consented participants, one is of non-participants, and one is of total students. Each data frame contains belonging students’ demographic information. Manual validation check is applied. For example, we manually counted the number of students who post below 5 times, and it matches the data R calculated. Statistical graphics are generated for each perspective of selected demographic characteristics. Because we are also interested in finding if there is a demographics pattern among how frequently a person posts online, later we focus on analyzing participants. Density plot of gender versus number of posts is conducted.

Last but not at least, Chi-Square goodness of fit test is applied. This test allows us to test whether the observed proportions for a categorical variable differ from hypothesized proportions[5].

1. Results and Analysis

First, we give an overview of sample data we collected, including gender distribution, ethnicity distribution, major distribution, and ADCI distribution. We compare the result with “Chi-Square goodness of fit” test. Then we looked at the gender and students’ performance distribution among the number of posts.

1. Participations vs. Nonparticipations
   1. Gender Distribution



*figure 1 Student gender distribution among participants, nonparticipants and total.*

From the distribution graph above, we can see that 23.89% of discussion participants are female students. The percentage value is less than that of male students due to the unbalanced ratio of this engineering class. From the third column, we can see that the whole class only have 16.64% of female students. More than half of the female students get involved in this online discussion forum more or less. There are 76.1% of participants are men but 83.36% of total class students are male students. Therefore, female students are relatively more active in participating in online discussion forum. The gender distribution is similar to the result of many previous studies[3][6], where they declared that women are more active than men on social network media and online discussion blog. It also matches the Donovan’s observation[8].

* 1. Ethnicity Distribution

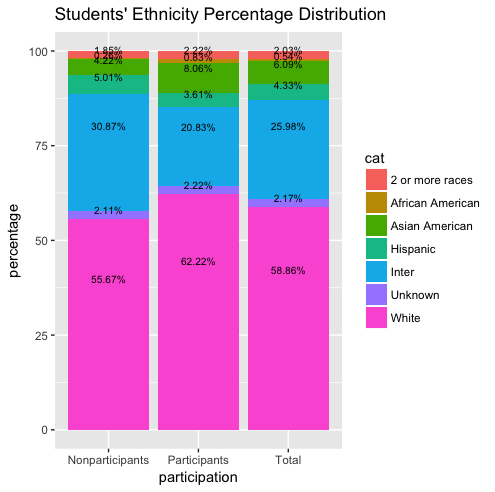


figure 2.a

figure 2.b figure 2.c

*figure 2 Students’ Ethnicity Distribution. Figure 2.a, the distribution is of stacked histogram.*

Figure. 2 shows the nationality and ethnicity distribution of the sample data. Our total data shows that the white American people dominate the population, followed by International students, Hispanic, and Asian Americans. There are 58.86% of the class students are white American. Among all participants, 62.22% are the whites, and 55.67% of nonparticipants are white people. From these data, we can see that white Americans over-represented. Next, from column three, we can see that 25.98% of the class students are internationals, but this ratio decreases to 20% in participants column. As a result, international students look under-represented in this forum. Moreover, there are 4% of the class students are Hispanic or Latino and 6% are Asian Americans. The ratio of Hispanic students in participants column decreases to 3.6% but the ratio of Asian Americans increases to 8%. Then we can say that Hispanics less involved but Asian Americans are more involved in this online forum. Therefore, White Americans and Asian Americans are over-represented but international students and Hispanics are under-represented. This matches one of the previous work[3], where they declared that minority status has lower satisfaction with the web-based distance study.

* 1. Major Distribution

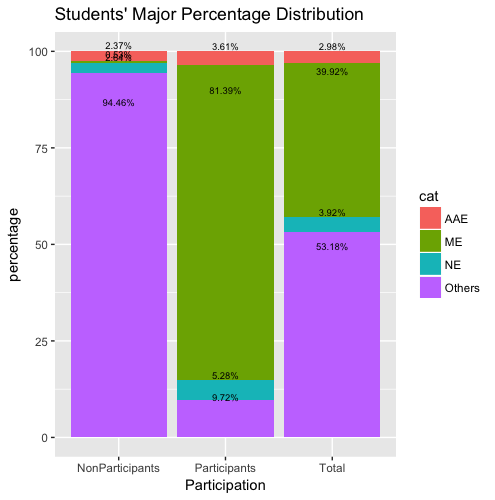


figure 3.a

*figure 3. Major Distribution among participants, nonparticipants and all students. ME represents Mechanical Engineering, AAE represents Agriculture Engineering, NE represents Nuclear Engineering.*

*Figure 3.a is the stacked histogram of major distribution. Figure 3.b, 3.c, 3.d are pie chars of major distribution for participants, nonparticipants and all students respectively.*

Figure 3 shows the major structure of the sample data. This online discussion forum is designed for a dynamic course, so all students are directly from mechanical engineering(ME), Agriculture Engineering(AAE), Nuclear Engineering, and other engineering departments, where students can take this course as a selective. Among those, ME students are the majorities, after are NE, AAE and others. From the graph, we can see that there are total 39.92% ME students, but 81.39% of participants are MEs. We can see that ME students are very active involved. There are totally 3.92% NE students, and 5.28% are among participants. Therefore, NE students also interact with their peers actively. There are 2.98% of total students from Agriculture Engineering, and they take 3.61% off total participants. We can see that AAE students are over-represented. Last there are 53.18% of students from other engineering department, who take this course only for credits. The ratio of participants in participants’ group decreases to 9.72%, and the ratio of nonparticipants increases to 94.46%, so students taking this course only for credits are very under-represented. Therefore, we can conclude that students from ME, NE and AAE for whom this course is mandatory or selective are more involved in this online discussion forum, but students from other engineering department are less involved.

1. Among Participants:

1. Gender Distribution

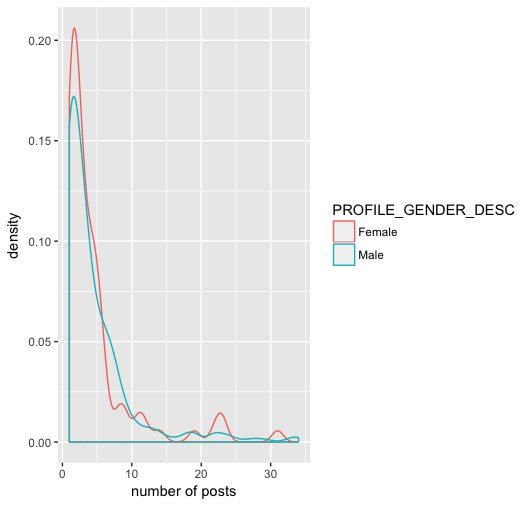


figure 4 Gender Distribution among the number of posts

The figure above shows the gender density trend among all participants. We can obviously see that women have a higher peak value at lower number of posts (0 - 5) and medium-high posts (20 - 25). Here we didn’t count one student, who posts 150 times over one semester, in order to keep the trend line more appropriate to describe the situation.

1. Discussion

From the results we got above, we are able to conclude that gender, ethnicity and majors relate to discussion forum participation. , which matches to the many previous studies and our hypothesis. and Hispanic . In this specific course in USA, international students are minorities. This observation corresponds to our original expectation and previous studies, where Ke’s group demonstrate that students’ minority status correlated with lower satisfaction with the web-based, distance-learning class. According to our major study, students from mechanical engineering who take the course as a requirement or students from nuclear and agriculture engineering who take the course as a selective are more active than those who take the course just for credits. Furthermore, ME students are the most engaged group in this class, which matches our hypothesis in the beginning. Last but not at least, female students have a very high density of students who post occasionally (under 5 posts) compared to their male counterparts, and have a slightly high density among students who post very often, which also match our expectation.

Limitation:

Due to the fact that we only design our research based on consented sample data, it doesn’t contain all students enrolling the class. However, consented students take up 84.7% of total students, where consented participants are 96% of all participants.

1. Conclusion

In this paper, social network analysis combined with statistical graphics and validation check have been used to understand demographics of students who do or do not participate in online discussion forums. The results will help instructors have a better understanding of whom may benefit most from online discussion and whom do not feel comfortable using web-based discussion forums.

1. Future Work

This work extends our knowledge of who uses online collaboration tools, and future work will analyze the content of posts and explore the influence of forum participations on grades via a regression model.

1. Reference
   1. P.Suraj and V. S. K. Roshni, “Social Network Analysis in Student Online Discussion Forums”, *2015 IEEE Recent Advances in Intelligent Computational System(RAICS)*, 2015.
   2. P. He, “Evaluating Students Online Discussion Performance by Using Social Network Analysis”, 2012.
   3. C. Yeh, Y. Lin, T. Lin, “Demographics of Social Network Users- a Case Study on Plurk”, 2012.
   4. D. Andrews, J. Preece, M. Turoff, “A Conceptual Framework for Demographic Groups Resistant to Online Community Interaction”, 2001.
   5. Engaging students for success in calculus with online learning forums
   6. “Analysis on the General profile of users on Plurk.com”
   7. “Demographic and Structural Characteristics to Rationalize Link Formation in Online Social Networks”
   8. “Characterizing user behavior and information propagation on a social multimedia network”
   9. S.K. Taradi, M. Taradi, "Expanding the traditional physiology class with asynchronous online discussions and collaborative projects", Advances in Physiology Education, 28 (2) (2004), pp. 73-78
   10. L.M. HarasimOn-line education: a new domain R. Mason, A. Kaye (Eds.), Mindweave: Communication, computers, and distance education, Pergamon, Elmsford, New York(1989)
   11. D.R. Garrison, T. Anderson, W. ArcherCritical inquiry in a text-based environment: computer conferencing in higher education, The Internet and Higher Education, 2 (2–3) (1999), pp. 87-105

[1] C. K. Cheng, D. E. Paré, L.-M. Collimore, and S. Joordens, “Assessing the effectiveness of a voluntary online discussion forum on improving students’ course performance,” *Comput. Educ.*, vol. 56, no. 1, pp. 253–261, Jan. 2011.

[2] “1-s2.0-S1096751600000166-main.pdf.” .

[3] A. Minichiello and C. Hailey, “Engaging students for success in calculus with online learning forums,” in *2013 IEEE Frontiers in Education Conference (FIE)*, 2013, pp. 1465–1467.

[4] F. Ke and D. Kwak, “Online learning across ethnicity and age: A study on learning interaction participation, perception, and learning satisfaction,” *Comput. Educ.*, vol. 61, no. Supplement C, pp. 43–51, Feb. 2013.

[5] “What statistical analysis should I use? Statistical analyses using SPSS,” *IDRE Stats*. .