Profiling MOOC course returners:
How does student behavior change between two course enrollments?

Vitomir Kovanović[1], Srećko Joksimović[2], Dragan Gašević[1,2],
James Owers[1], Anne-Marie Scott[3], and Amy Woodgate[3]

The University *of* Edinburgh

1: School of Informatics 2: Moray House School of Education 3: Information Services

v.kovanovic@ed.ac.uk

INTRODUCTION

Massive Open Online Courses represent a fertile ground for examining student behavior. However, due to their openness MOOC attract a diverse body of students, for the most part, unknown to the course instructors. However, a certain number of students enroll in the same course multiple times, and there are records of their previous learning activities which might provide some useful information to course organizers before the start of the course.

In this study we examined how student behavior changes between subsequent course offerings. We identified profiles of returning students and also interesting changes in their behavior between two enrollments to the same course.

RESEARCH QUESTIONS

RQ1: What are common behavioral profiles of students who enroll MOOCs multiple times?

This question is the first step in our analysis and it allows for examining whether there are any particular forms of MOOC engagement by the students who enroll in the same courses multiple times.

RQ2: How do students change their behavior between subsequent offerings of the same course?

This follow-up question is a natural extension to our first question and focuses on student self-regulation of learning. Do students change their behavior between two offerings or they simply continue with the same form of participation as they did the first time?

METHOD

Dataset

The data for this study comes from the 28 offerings of the 11 different MOOCs offered by the University of Edinburgh on the Coursera platform (Table 1). In our analysis, we examined only data about students' first and second enrollment. That is, we did not analyze students who enrolled only once, and we also excluded any subsequent (i.e., third or fourth) enrollments. In total, we had 52,050 student course records.

Analysis procedure

We conducted a cluster analysis of the 52,050 enrollment records using the variables listed in Table 2. As overall student activity in each course was slightly different, we first performed unitization with zero minimum (i.e., x-min/range) per each course offering, along with a z-score normalization on the whole corpus. This ensured that 1) specifics of each course (and course offering) were taken into account for scaling each variable, and 2) all variables were on the same scale to ensure equal importance of variables. As there are many students who just enrolled and never accessed courses, we removed them from our cluster analysis and assigned them a predefined "Enroll Only" cluster.

We performed K-means clustering using Lloyd's algorithm (with ten restarts and a maximum of 300 iterations) for values of K between 2 and 10 and the evaluated the percentage of variance explained by the different clustering solutions. We also examined a transition graph between students' first and second enrollment to see what the most common cluster transitions are.

Table 1. *MOOCs analyzed in the study.*

Course	Offers
Artificial intelligence planning	1, 2
Animal behavior and welfare	1, 2
AstroTech: The science and technology behind astronomical discovery	1, 2
Astrobiology and the search for extraterrestrial life	1, 2
The Clinical psychology of children and young people	1, 2
Critical thinking in global challenges	
E-learning and digital cultures	1, 2, 3
EDIVET: Do you have what it takes to be a veterinarian?	1, 2
Equine Nutrition	1, 2, 3
Introduction to philosophy	1, 2, 3, 4
Warhol	1, 2

Table 3. *Identified clusters.*

Cluster	Students	%
Enroll Only (E)	22,932	44.1
Low Engagement (LE)	21,776	41.8
Videos & Quizzes (VQ)	2,120	4.1
Videos (V)	5,128	9.9
Social (S)	94	0.2

Table 2. *Variables used for clustering.*

Variable	Description
Days	No. of days active
Submissions	No. of sub. assignments
Wiki	No. of wiki page views
Discussions	No. of discussion views
Posts	No. of posts written
Quizzes	No. of quizzes attempted
Quizzes unique	No. of diff. quizzes attempted
Videos unique	No. of diff. videos watched
Videos	No. of videos watched

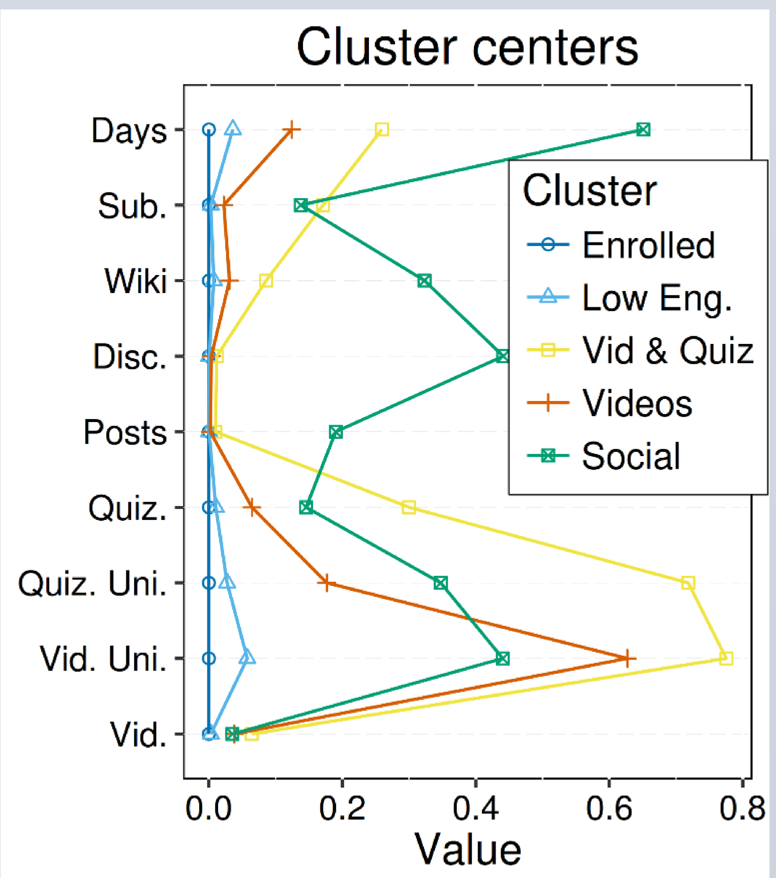

Figure 1. *Cluster profiles.*

RESULTS

Clustering

We identified five clusters of the behavior of returning students. The cluster centers are shown in Figure 1 while their relative sizes are shown in Table 3. These results are aligned with the previous work on online courses [3] and MOOCs [1] that showed similar disproportions between highly active and inactive students. The identified clusters (Figure 1) reveal that the largest part (85% of all enrolled students) have no or have very little course activities. Around 10% of the students focused primarily on viewing video lectures, while 4.1% of students were highly engaged and, besides watching videos, also utilized quizzes and engaged in homework assignments. Finally, less than 1% of student put an emphasis on online discussions, while being less engaged with video lectures. This cluster of students also stayed longest active in courses.

Cluster transitions

To investigate how student behavior changes between subsequent course enrollments, we constructed a directed state transition graph (Figure 2) which shows what percentage of first enrollment cluster members transferred to other clusters (or remained within the same cluster). The majority of students from all the clusters except the "Social" cluster either just enrolled in a course or had very low level of engagement. A certain number of students who utilized both video lectures and quizzes during their first enrollment either retained the same level of engagement or focused primarily on video lectures in the second course enrollment.

These two patterns are likely driven by the goal of obtaining course certificate or brushing up on a particular course topic. Finally, the most interesting finding is related to the students from the "Social" cluster who had the highest level of participation in online discussions and also most days spent in the course. While a certain number of students became disengaged in the next offer of the course, a large chunk of them (28%) kept their level of participation, signaling the goal of engaging with other learners rather than the prescribed course content.

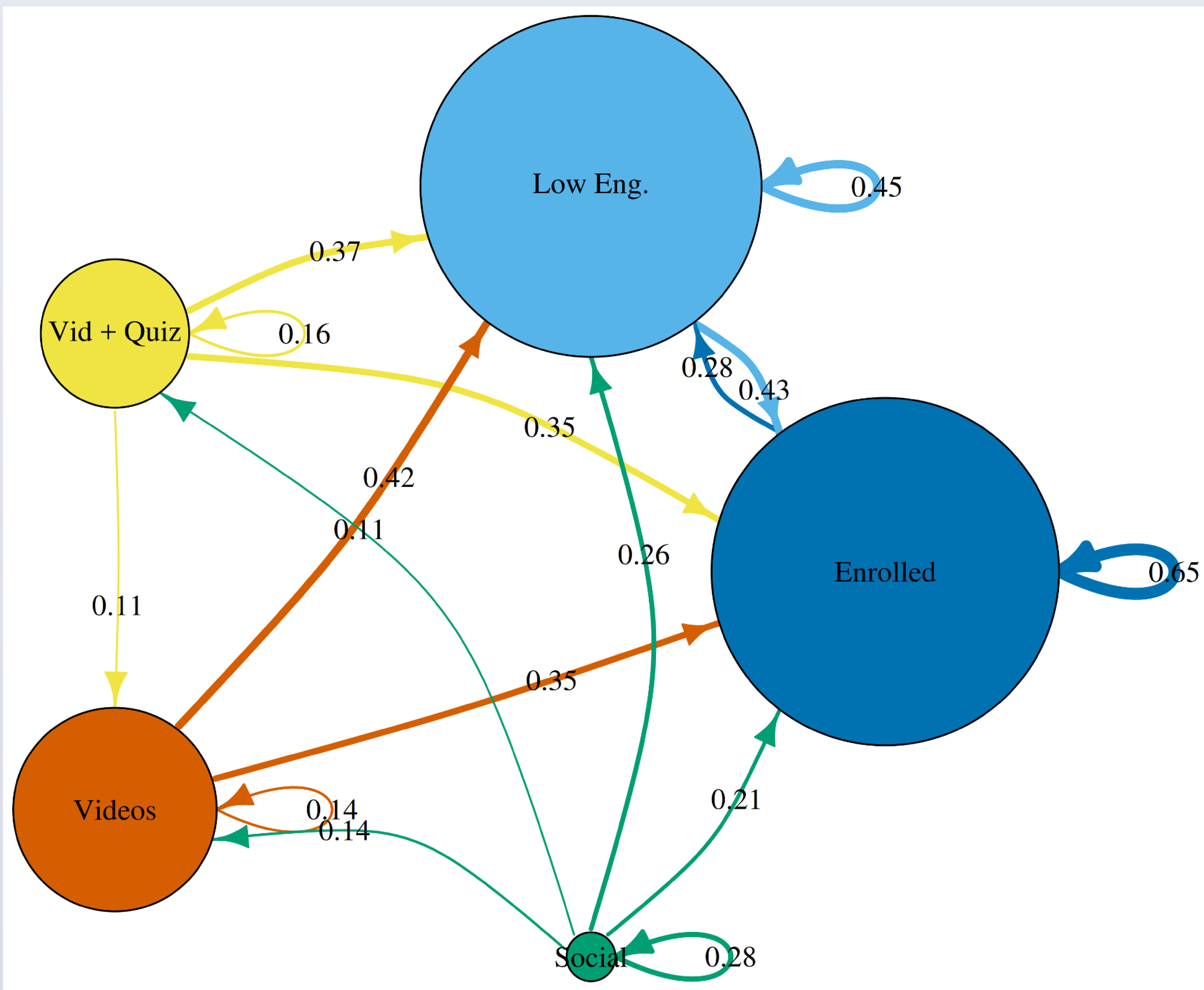

Figure 2. *Cluster transition graph. Sizes of nodes represent the number of students while edge labels represent percentage of source cluster transitioning to the destination cluster.*

IMPLICATIONS

First, as a majority of students who were not active (or had low levels of activity) in their first enrollment were likely to stay inactive, course instructors might consider targeting those particular students with a certain set of instructional interventions which would increase their levels of participation. Similarly, students who exhibited high levels of activity in the first offer might be targeted with interventions that would encourage them to participate more in the discussions, or with interventions related to particularly challenging course content (as indicated by their quiz and assignment scores in the previous enrollment). Finally, through identification of socially engaged students, instructors might identify suitable community teaching assistants which could be then better supported by the instructional team.

REFERENCES

1. Doug Clow. 2013. MOOCs and the Funnel of Participation. *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, ACM, 185–189.

2. R.F.a Kizilcec and E.B Schneider. 2015. Motivation as a lens to understand online learners: Toward data-driven design with the OLEI scale. *ACM Transactions on Computer-Human Interaction* 22, 2.

3. Vitomir Kovanović, Dragan Gašević, Srećko Joksimović, Marek Hatala, and Olusola Adesope. 2015. Analytics of communities of inquiry: Effects of learning technology use on cognitive presence in asynchronous online discussions. *The Internet and Higher Education* 27: 74–89.