

Mapping Divergence and Similarities in MOOC Learning Paths

[in-progress, submitted to a conference]

Saman Rizvi *

Bart Rienties, Jekaterina Rogaten

[firstname.lastname@open.ac.uk]

Institute of Educational Technology (IET), The Open University, UK

Introduction

- MOOCs; 81+ million learners, 9400+ courses, 800 + universities/
Institutions
- Most critical challenge: Learners' retention
- Learners leave behavioral traces behind, stored in voluminous
system logs; an asset !
- Log-based behavioral modeling in educational domain
 - Intelligent Tutoring System (ITS)
 - Learning Management System (LMS)

Introduction

- “Academic success” in MOOCs?
 - Partly hidden in learners’ journeys through their respective learning activities, and in interactions with a variety of learning resources
 - *Processual*, and guided by learners’ intentions
- Comprehensive log data exploration is needed to understand learners’ behavioral patterns and their temporal learning choices

Introduction

- Initial exploratory analysis: Three distinct clicking patterns
- Markers, Partial-Markers, and Non-Markers
- Markers represented learners who marked all their activities as completed

Theoretical Framework

- The Open University Learning Design Initiative (OULDI)
- Categorizes all learning tasks into seven learning activity types;
 - Assimilative,
 - Finding information,
 - Communication,
 - Productive,
 - Interactive,
 - Adaptive,
 - Assessment
- A well-articulated abstraction of all learning activities

Research Aim (s)

- To understand how learners perform in Massive Open Online Courses (MOOCs)?
- Whether the performance varies with the MOOC design (or discipline)?
- What are the temporal dynamics?
 - Progression
 - Engagement-duration

Research Question

- **RQ1.** In terms of activity access frequency, and activity learning time, how and to what extent does participatory behavior vary with the learning design in the four MOOCs?

- **RQ2.** How and to what extent do access frequencies and temporal learning paths differ between Markers, Partial-Markers, and Non-Markers in the four MOOCs?

Data Selection

Table 1. Four MOOCs description

Course	Discipline	Total Learners	Markers	Partial-Markers	Non-Markers
MOOC 1	Nature & Environment and Science, Eng. & Math	2086	449	832	805
MOOC 2	Tech & Coding and Business & Management	981	114	426	441
MOOC 3	Business & Management	1927	291	805	831
MOOC 4	Languages & Cultures and Study Skills	11763	843	6390	4530

Methods

➤ Educational Process Mining (EPM)

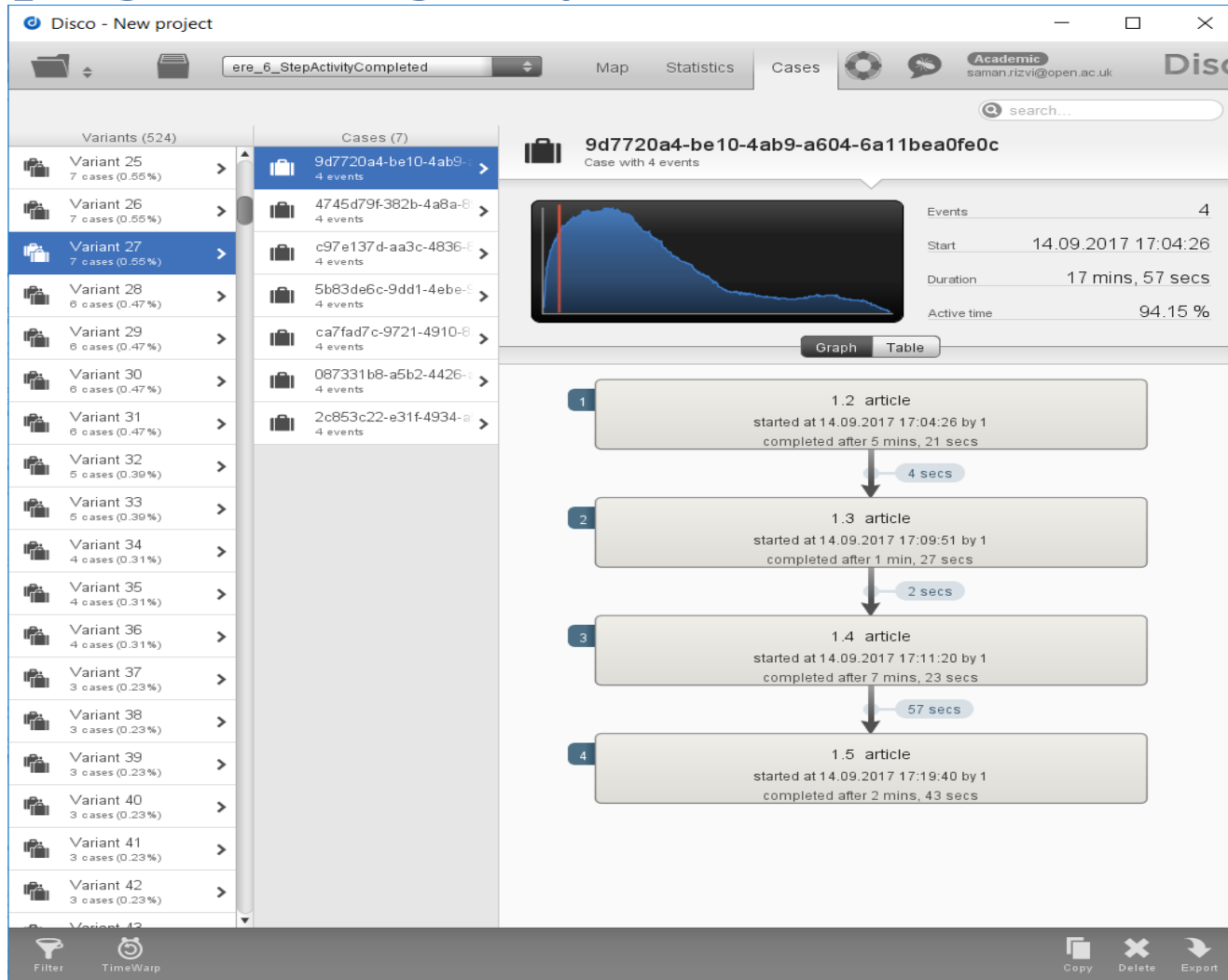
Table 2. Relative frequency of access and median duration

	Activity Distribution	Activities Accessed			
MOOC1	Article = 44 (64.70%) Discussion=12 (17.65%) Video=8(11.76%) Quiz=3(4.41%) Test= 1(1.5%)	Activity	▲ Frequency	Relative frequency	Median duration
		Assimilative_Article	27,349	64.71 %	2 mins, 1 sec
		Communication_Discuss	7,323	17.33 %	1 min, 33 secs
		Assimilative_Video	5,731	13.56 %	2 mins, 36 secs
		Assessment_Quiz	1,796	4.25 %	3 mins, 12 secs
		Assessment_Test	63	0.15 %	23 mins, 38 secs
MOOC2	Article = 52 (52.52%) Video = 24 (26.09%) Discussion = 7 (7.61%) Quiz = 5 (5.43%) Test = 4 (4.35%)	Activity	▲ Frequency	Relative frequency	Median duration
		Assimilative_Article	8,272	55.4 %	1 min, 30 secs
		Assimilative_Video	4,409	29.53 %	2 mins, 1 sec
		Communication_Discuss	1,313	8.79 %	53 secs
		Assessment_Quiz	829	5.55 %	2 mins, 41 secs
		Assessment_Test	109	0.73 %	8 mins, 51 secs
MOOC3	Article = 46 (54.12%) Video = 18 (21.18%) Discussion= 13 (15.29%) Quiz = 3 (3.53%) Test = 4 (4.70%) Audio = 1 (1.18%)	Activity	▲ Frequency	Relative frequency	Median duration
		Assimilative_Article	15,055	50.34 %	1 min, 9 secs
		Assimilative_Video	8,291	27.72 %	2 mins
		Communication_Discuss	5,172	17.29 %	46 secs
		Assessment_Quiz	975	3.26 %	1 min, 3 secs
		Assimilative_Audio	311	1.04 %	38 secs
		Assessment_Test	105	0.35 %	3 mins, 36 secs
MOOC4	Article = 41 (38.32%) Video = 24 (22.42%) Discussion = 16 (14.95%) Audio = 3 (2.8%) Quiz = 13 (12.15%) Test = 6 (5.61%) Assignment = 1 (0.93%) Review = 1 (0.93%) Reflection = 1 (0.93%)	Activity	▲ Frequency	Relative frequency	Median duration
		Assimilative_Video	64,408	36.32 %	2 mins, 29 secs
		Assimilative_Article	54,471	30.72 %	1 min, 21 secs
		Communication_Discuss	30,622	17.27 %	1 min, 14 secs
		Assessment_Quiz	22,377	12.62 %	3 mins, 24 secs
		Assimilative_Audio	3,223	1.82 %	1 min, 57 secs
		Productive_Assignment	728	0.41 %	0 millis
		Assimilative_Review	636	0.36 %	0 millis
		Adaptive_Reflection	618	0.35 %	7 secs
		Assessment_Test	260	0.15 %	6 mins, 54 secs

Table 3. Frequency of access for all three categories.

MOOC	Learning Activities Accessed
MOOC1 Article = 44 (64.70%) Discussion =12 (17.65%) Quiz=3(4.41%) Test= 1(1.5%) Video=8(11.76%)	Markers
	Partial-Markers
	Non-Markers
MOOC2 Article = 52 (52.52%) Discussion = 7 (7.61%) Quiz = 5 (5.43%) Test = 4 (4.35%) Video = 24 (26.09%)	Markers
	Partial-Markers
	Non-Markers

Mapping Learning Trajectories



List of the 524 types of learning sessions obtained from the log. The type 27 shows 4 end-to-end interactions (events), with the time associated with the duration of the session (variant 27: learning path of a subgroup of 7 learners).

Results

Table 4. Top Variants (Large subgroups following a learning trajectory).

M O O C	Category	Cases in top Variants	Events	First Activity	Last Activity
M O O C 1	Markers	147 (32.7%)	67	1.1 Video	4.18 Article (missed 4.16)
		43 (9.7%)	16	1.1 Video	1.16 Article
		27 (6.0%)	1	1.1 Video	1.1 Video
	Partial- Markers	70 (8.4%)	2	1.1 Video	1.2 Article
		40 (4.8%)	3	1.1 Video	1.3 Article
		29 (3.5%)	4	1.1 Video	1.4 Article
	Non- Markers	543 (67.5%)	1	1.1 Video	1.1 Video
		68 (8.5%)	2	1.1 Video	1.2 Article
		18 (2.2%)	3	1.1 Video	1.3 Article

Conclusion

- Learners remained more interested in video based Assimilative activities but not in article based assimilative activities
- Duration of interest: median learning time
 - Reading-based assimilative activities: Ranged between 1 minute 9 seconds to 2 minutes 1 second
 - Communication Activities: Ranged between 46 seconds to 1 minute 33 seconds

Conclusion

- If the analyses were performed without first grouping learners into respective categories, the findings would have remained strongly biased towards the majority group
- In this case, the majority group is Non-Markers making up to between 38.5% (MOOC 4) and 45% (MOOC 2) of overall learners' population
- Top variant in Markers; Learners completed almost all activities
- Overall, the top three variants in Markers category comprised of 'Auditing' or 'Completing' learners, also referred to as either 'strong starters', or 'keen completers' in other MOOC literature
- Top variant in Non-Markers; accessed only first activity, never resumed their learning, i.e., 'Samplers' (visiting first few activities only)

General Conclusion

Fall in love with the process, and the results will follow.

[Eric Thomas]

Implications / Future Work

- Using Process Mining in combination with other analytics techniques, such as
 - manual or automated clustering during preprocessing, (to produces context-aware behavioral models)
 - Clustering, natural grouping in data
- Utilize not only the log data but also other contextual or interaction information, typically not captured in event log data; such as
 - learners' demographics,
 - discussion text or
 - learning outcomes

Thank You

Q / A

Suggestions

Saman Rizvi *

Bart Rienties, Jekaterina Rogaten

[firstname.lastname@open.ac.uk]

Institute of Educational Technology (IET), The Open University, UK