Profiling MOOCs from Viewing Perspective

Kang WEI and Jian-ping LI

College of Liberal Arts and Sciences, National University of Defense Technology, Changsha, China

Keywords: Viewing perspective, MOOCs.

Abstract. Course assessments inextricably connect to learning behaviors, which are different from traditional courses to Massive open on line courses (MOOCs). The emerged high course dropout and low exam participation make viewing be the mostly presentive behavior of MOOC learning, and make us call for new assessment indexes. Based on the viewing behavior data provided by course platform iCourse, we provide a method to recognize potential all-rounders, and indexes to measure course attractions and order correlations between teaching and learning. For example, we treat the video label of viewing events as a random variable, adopt its information to measure viewing width, and then use the geometric mean of the information and the relative length of viewing time comparing with the length of videos to measure course attractions. The index can measure the diminishing marginal utility for learners as well as the information increment by viewing new videos. However, the unclear relationship between viewing and learning makes our results could not be used to assess course quality positively. Nevertheless, these can be used to detect which aspects of a course need to improve.

Introduction

MOOCs are products of the combination of the Internet and education [1]. Their transmission advantages inherited from Internet enlarge the scale of traditional teaching, and break the limitation of learners’ time and their space. MOOCs have been regarded as a "particle accelerator for learning"[2], and also be treated as a way to reduce the disequilibrium in education resource[3]. The differences between the courses taught in classroom and MOOCs exist in teaching pattern, admission conditions, dropout rate, registrant type, learning purpose, and so on [4-6]. Hence it is unsuitable to using assessment indexes of traditional course to profile MOOCs, such as completion rate and excellence rate [7].

Course assessments are essentially to determine the degree to which the teaching plan actually reaches its goal[8], and so inextricably connect to learning behaviors. Based on the actions of online learning, learner behaviors and performance can be revealed through learner participation[9], learner behavior patterns, and the relationships between participation and performance[10]. Facing the emerged high course dropout and so low exam participation, we need new assessment indexes for MOOCs. Compared with the participation of exams and quizzes viewing is more common and presentive behavior of online learning. Hence assessing from viewing perspective can involve many learners, which makes it be a promising direction.

This paper is a short survey of the works in References [11-14]. Firstly, we review the recognition of potential all-rounders in Reference [12], which helps course providers focus on analyzing the dropout factors of these learners and inspire their participation [15,16]. The motivations of learners studying MOOCs are not limited to pass examinations, which could be knowing specific concepts, comprehending some contents, and so on. The determination of a potential all-rounder could be viewed as a unit with a failure mode controlled by fatigue-stress nature. The live of such a unit follows a lognormal distribution. Hence we provided a method to recognize potential all-rounders as those learners whose the length of viewing time follows a distribution, namely lognormal.

Secondly, we address MOOC attractions or learner attentions by synthetically measuring viewing widths and lengths of learners. For a learner, we continuously measure his viewing width through the information calculated based on his viewing time distribution and then use the geometric mean of
the information and his relative length of viewing time comparing with the length of videos. The index profiles the diminishing marginal utility of learning, and the increment of the information received by viewing new videos. Specific indexes based on the information and the corresponding entropy, such as Shannon evenness, average information of videos, can measure the attraction balance over videos, and then the persistence of learner attentions.

Thirdly, we address order correlations between teaching and learning. The order of learning could be measured by the order of viewing videos, and the order of teaching can expressed by the label of videos. For each video, we design a weighted average of video labels to measure its adjacent video labels in the sense of viewing actions. Then, we measure the order correlation by the Pearson’s correlation coefficients between the weighted average label and video label. The order correlation has less meaning to courses of humanity, but would be critical to courses of sciences.

Viewing Behavior Data

The platform ICourse provides the data of viewing behavior of 8 courses in 2017 semester. The courses are respectively selected from the fields of natural science, social science, humanities, and engineering. Those courses have many learners. The data contain the length of video’s time. The data also contain the label, the length of viewing time, and the start time of viewing video for each learner. The data are collected before 10/11/2017. Each course is completed or close to complete.

<table>
<thead>
<tr>
<th>Course</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calculus</td>
<td>2.955</td>
<td>129</td>
<td>8.08</td>
<td>1.00</td>
<td>0.19</td>
<td>2</td>
</tr>
<tr>
<td>Game theory</td>
<td>4.764</td>
<td>38</td>
<td>7.14</td>
<td>2.24</td>
<td>0.43</td>
<td>66</td>
</tr>
<tr>
<td>Finance</td>
<td>6.380</td>
<td>63</td>
<td>5.37</td>
<td>1.31</td>
<td>0.33</td>
<td>2</td>
</tr>
<tr>
<td>Psychology</td>
<td>3.827</td>
<td>26</td>
<td>5.01</td>
<td>0.91</td>
<td>0.20</td>
<td>59</td>
</tr>
<tr>
<td>Spoken English</td>
<td>11.719</td>
<td>46</td>
<td>3.03</td>
<td>0.32</td>
<td>0.11</td>
<td>7</td>
</tr>
<tr>
<td>Etiquette</td>
<td>3.846</td>
<td>41</td>
<td>7.79</td>
<td>1.27</td>
<td>0.21</td>
<td>22</td>
</tr>
<tr>
<td>C Language</td>
<td>17.541</td>
<td>81</td>
<td>12.5</td>
<td>1.54</td>
<td>0.14</td>
<td>39</td>
</tr>
<tr>
<td>Python</td>
<td>13.417</td>
<td>53</td>
<td>10.3</td>
<td>0.90</td>
<td>0.09</td>
<td>28</td>
</tr>
</tbody>
</table>

Index \(a\) is the number of learners, \(b\) is the number of videos, \(c\) is average viewed video number of learners, \(d\) is average viewing time length of learners, \(e\) is average time length of videos, and \(f\) is the number of all-rounders.

Some statistical indexes of viewing behaviors are shown in Table 1, which are used to measure the relationship between viewing behaviors and video lengths. Suppose learners \(\{L_1, \ldots, L_m\}\) view \(n\) videos \(\{V_1, \ldots, V_n\}\). For each learner \(L_s (s = 1, \ldots, m)\), denote the set of labels of he viewed videos by \(S^V_s\). For video \(V_i (i = 1, \ldots, n)\), denote the set of numbers of learners who viewed it by \(S^L_i\). Denote the length of video \(V_i\) by \(l_i\), the time length of learner \(L_s\) that view \(V_i\) by \(t^s_i\).

Calculate the average relative length of viewing time (unit: video), and the average quantity of videos viewed. We assume that the learner priory views the whole video, the rate of the two averages \(C_1 = (\sum_{i=1}^{n} \sum_{s=1}^{m} l^s_i / l_i) / \sum_{i=1}^{m} |S^V_i|\) measures the viewing completion rate of learners. The viewing rate can also be measured by \(C_2 = 1/n \sum_{i=1}^{n} \sum_{s \in S^L_i} \min(t^s_i / l_i, 1)/|S^L_i|\) at the same hypothesis. Both \(C_1\) and \(C_2\) negatively correlate to the average length of videos. Long videos cannot capture the attentions of learners all the time. Therefore, the contents of long videos must be designed carefully to attract learner attentions.

Viewing Pattern Recognition

The interest spans of learners to a course are different. Some interest in the whole courses, and some only interest in some contents. Hence the viewing behaviors of learners can be sketchy classified as
two viewing patterns: segment and all-round, and so learners as segment-learners and all-rounders. Discussing the factors of dropout and engagement for segment-learners has limited insight, but is meaningful for potential all-rounders. Hence it is meaningful that to what extent one could be an all-rounder or a segment-learner, which can be recognized by following method.

For each course, the number of all-rounders in the sense of viewing all videos is very small. However, even the learners, who determine to complete a course, may do not view all videos. They can be treated as potential all-rounders. For such learners, their determination could be viewed as a unit with a failure mode with fatigue-stress nature. Their live length follow a distribution called lognormal. The determination live length of a learner can be treated as the length of his viewing time. An algorithm is given to recognize these learners (Table2).

Table 2. A recognition algorithm of potential all-rounders.

| Input: viewing time lengths $t_s$ and viewing video numbers $n_s$ of learners $L_s$ ($s = 1, ..., m$). |
| For each possible $k$ do: |
| Do KS test for learners $L_s$ satisfying $n_s > k$ with the hypothesis of following the distribution lognormal. |
| Break if the test cannot reject the hypothesis at significance level 5%. |
| Output: the $k$ when break happens. |

The learners $L_s$ satisfying $n_s$ larger than the current $k$ are recognized as potential all-rounders.

In each course, viewing some videos, the length of viewing time for those learners follows a lognormal. The other learners could be regarded as segment-learners, because they only viewed a few videos. The viewing length of a segment-learner approximately follows a power law with an exponent cut-off. Some indexes of potential all-rounders and segment-learners are listed in Table3.

Table 3. Some indexes of the empirical data.

<table>
<thead>
<tr>
<th>Course</th>
<th>Pattern</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calculus</td>
<td>A</td>
<td>569</td>
<td>28.12</td>
<td>3.848</td>
<td>4.078</td>
<td>23.06</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>2,386</td>
<td>3.302</td>
<td>0.319</td>
<td>1.083</td>
<td>2.385</td>
</tr>
<tr>
<td>Game theory</td>
<td>A</td>
<td>1,522</td>
<td>16.86</td>
<td>5.919</td>
<td>3.578</td>
<td>14.13</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>3,242</td>
<td>2.581</td>
<td>0.51</td>
<td>0.872</td>
<td>1.689</td>
</tr>
<tr>
<td>Finance</td>
<td>A</td>
<td>1,057</td>
<td>21.54</td>
<td>5.501</td>
<td>3.82</td>
<td>16.2</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>5,323</td>
<td>2.157</td>
<td>0.478</td>
<td>0.602</td>
<td>1.276</td>
</tr>
<tr>
<td>Psychology</td>
<td>A</td>
<td>799</td>
<td>15.3</td>
<td>3.1</td>
<td>3.531</td>
<td>13.3</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>3,028</td>
<td>2.294</td>
<td>0.336</td>
<td>0.648</td>
<td>1.643</td>
</tr>
<tr>
<td>Spoken English</td>
<td>A</td>
<td>583</td>
<td>19.61</td>
<td>2.426</td>
<td>3.724</td>
<td>18.45</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>11,136</td>
<td>2.164</td>
<td>0.211</td>
<td>0.636</td>
<td>1.67</td>
</tr>
<tr>
<td>Etiquette</td>
<td>A</td>
<td>2084</td>
<td>12.95</td>
<td>2.213</td>
<td>3.084</td>
<td>10.8</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>1,762</td>
<td>1.683</td>
<td>0.157</td>
<td>0.469</td>
<td>1.035</td>
</tr>
<tr>
<td>C Language</td>
<td>A</td>
<td>2,367</td>
<td>46.57</td>
<td>6.609</td>
<td>5.161</td>
<td>43.24</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>15,174</td>
<td>7.147</td>
<td>0.75</td>
<td>1.827</td>
<td>5.833</td>
</tr>
<tr>
<td>Python</td>
<td>A</td>
<td>2,549</td>
<td>28.76</td>
<td>2.748</td>
<td>4.475</td>
<td>28.75</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>10,868</td>
<td>5.6</td>
<td>0.461</td>
<td>1.791</td>
<td>5.243</td>
</tr>
</tbody>
</table>

Pattern A and B are potential all-rounders and segment-learners respectively. Index $a$ is the number of learners, $b$ is the average viewed video number of learners, $c$ is the average viewing time length of learners, $d$ is the average entropy of learners, and $e$ is the geometric mean (2).

**MOOC Attraction Measurements**

When a learner views a course, we could regard the label of the video he chose to view is a random variable. Ignoring the order of course contents, a learner views more videos and in a more uniformly way, then the uncertainty of viewing which video is higher. The entropy is a measurement of uncertainty. Denote the label of the video in a viewing event of $L_s$ by $X_s$. The probability of choosing $V_i$ is $p(X_s = i) = t_i / \sum_{j=1}^{n} t_j$. Therefore, the entropy of $X_s$ is
Entropy (1) measures viewing width in a continuous way. If a learner views a video in a short time, his entropy will changes a little. Hence entropy decreases the shortcoming of the measurement by the number of viewed videos.

Entropy only depends on viewing time distribution. In general, the attraction of a course correlates positively to the viewing time length of the learner. Hence we can integrate width and length of viewing videos synthetically as an index. If the time-lengths of videos are equal, the unit of the relative viewing time length

\[
\frac{\sum_{i=1}^{n} t_i}{l_i}
\]

and the unit of the information \(2^{H(X_s)}\) are the same, which is the length of a video. Hence the geometric mean of them can be used to address the viewing width and length synthetically

\[
I(X_s) = \sqrt{2^{H(X_s)} \sum_{i=1}^{n} \frac{t_i}{l_i}}.\tag{2}
\]

The geometric mean (2) give a measurement of the MOOC attraction to individuals, especially to segment-learners. Since the lengths of videos are assumed to be equal, its unit is the video length. Hence the value of \(I(X_s)\) could be explained as the number of effectively viewed videos by learner \(L_s\). The reasonability is shown as follows.

Learner \(L_s\) viewed video \(V_1\) with length \(t_1 = l_1\), then \(H(X_s) = 0\) and \(I(X_s) = 1\). If \(V_1\) and \(V_i\) with length \(t_1 = l_1\), \(t_2 = l_2\), then \(H(X_s) = 1\) and \(I(X_s) = 2\). Therefore, \(\frac{\partial^2 I(X_s)}{\partial (t_i^2)} < 0\) describes the diminishing marginal utility of learning. The formula (2) describes the increment of information in the process of viewing new videos, because

\[
(p_1 + p_2) \log(p_1 + p_2) - (p_1 \log p_1 + p_2 \log p_2) > 0.
\]

The eight courses are selected from 4 disciplines. Python and Spoken English can capture many learners. However, Calculus hardly attract learners. In order to compare the attractions of courses of different disciplines, we could use the average geometric mean over all learners

\[
\frac{\sum_{i=1}^{m} I(X_s)}{m},
\]

without the heterogeneity of course learner numbers. It shows although viewing time length and viewing video number are unsuitable to measure MOOC attraction to individual (their shortcomings are mentioned in Section 1), their average over learners could measure MOOC attractions at certain degrees.

Now we analyze the balance of MOOC attractions over videos. For a course \(\{V_1, V_2, \ldots, V_n\}\), we denote \(l_i\) as the viewing time length of all its learners on \(V_i (i = 1, \ldots, n)\), and let \(P_i = l_i / \sum_{j=1}^{n} l_j\), then the entropy

\[
H(X) = -\sum_{i=1}^{n} P_i \log P_i
\]

describes the balance of a course’s attraction. However, courses have different video numbers. The entropy also depends on the number of videos. In order to compare this balance, we should delete the heterogeneity of video numbers. This can be done by averaging information over videos \(2^{H(X)} / n\), or by Shanon eveness \(H(X) / \log_2 n\).

The two balance indexes remove the number heterogeneity of learners and that of videos. The relatively low indexes of Spoken English due to one third viewing time attracting by only one video. If those balance indexes of a course are low, it means the course cannot attract learners persistently, and so its providers could improve some videos’ content. The operator called Rao-Sting [12]

\[
\Delta = \sum_{i,j(i \neq j)} d_{ij}^\alpha (p_i p_j)\beta \quad \text{(where } \alpha = \beta = d_{ij} = 1 \text{ for all } i \text{ and } j)\]

can describe the viewing balance, but it does not take into account the difference of course video numbers.
Correlation between Viewing and Teaching

The teaching order is a basic problem in pedagogy. Internet education changes teaching center from knowledge to learners. The level of the order correlation between teaching and learning affects the level of a course. The order of learning could be described by the order of viewing video at unsupervised situations. The order of teaching can be described by the label of videos. If nearly viewed videos have labels in sequence, the order of learning is consistent with the order of teaching.

A method is provided to measure the correlation of orders. For learner $L_i$, the viewing correlation between any two videos $V_i$ and $V_j$ of he viewed videos $S_i^V$ is we measured by $\omega_{ij} = f(|\tau_i^j - \tau_j^i|)$, where $f(\cdot)$ is a decreasing function, $\tau_i^j$ and $\tau_j^i$ are the viewing start times of $V_i$ and $V_j$ respectively. A small value of $|\tau_i^j - \tau_j^i|$ means that a viewing order exits between $V_i$ and $V_j$. For video $V_i$, a weighted summation is calculated as follows:

$$v(i) = \frac{\sum_{s \in S_i^V} \sum_{j \in S_i^V \cup S_j^V} \omega_{ij}}{\sum_{s \in S_i^V} \sum_{j \in S_i^V \cup S_j^V} \omega_{ij}}$$

(3)

The correlation between the weighted summation (3) and the video label measures the correlation between the order of teaching and the order of viewing. Let $\omega_b^i = \min(24/|s_i^j - s_j^i|, 1)$, and calculate three correlation coefficients, e.g. Pearson correlation coefficient, for the empirical courses. If a mathematic course has a low order correlation means that the teaching order needs to improve.

Conclusions and Discussions

The emergence of MOOC assessment topics is due to the differences between MOOCs and traditional courses. Learner-centered and autonomous learning, as features of MOOC teaching, fit the sprit of constructivism, in which learning behaviors play an important role in course assessments. Hence we use viewing behaviors, the presentive actions of MOOC learning, to assess course attractions and order correlations between teaching and learning, which help to find aspects of a course need to improve. Recognizing the segment-learners and potential all-rounders helps course providers focus attentions of on those potentially all-rounders and encourage they finish the whole courses. The reasonability of our results are validated against with the viewing behavior data provided by ICourse.

Acknowledgments

This work is supported by Hunan Education Science Planning Project (Grant No. XJK19BGD001).

References


[7] Wilkowski J, Deutsch A, Russell DM (2014, March) Student skill and goal achievement in the mapping with google MOOC. In Proceedings of the first ACM conference on Learning@ scale conference (pp. 3-10). ACM.


