Uncovering Collective Online Learning Habits

Zhi-hao XIAO* and Jian-ping LI

College of Science, National University of Defense Technology, Changsha, China

*Corresponding author

Keywords: Online learning habits, Learning behavior, Dipartite network, Power-law.

Abstract. In this paper, we analyzed the learners’ behavior data of the large-scale online course (MOOC). We used the data of learners’ learning behavior to establish the learner-video dipartite network. The two sets of the network are the learners and video collections. In this study, we found the pattern that learners watch videos obeys power-law distribution. In order to study the collective online learning habits, we performed a distribution fit on the learner's video behavior and calculated the parameters, and found that learners’ online learning behavior obeys the power-law distribution. Based on the above research, we found the collective online learning habits, which provides a theoretical basis for online course builders to master learners' learning habits.

Introduction

As a new form of education in the information age, large-scale online courses have profoundly changed the educational model and brought a new atmosphere for the development of education. Online courses use the advantages of the internet to break the boundaries of time and space and the walls of colleges. People can use the internet to learn high-quality courses and accept higher education at any time and any place. With the rise of large-scale online courses, learners are increasingly focusing on using internet resources for learning. We analyzed the learners’ behavior data of the large-scale online course, and used the data of learners’ learning behavior to establish a learner-video dipartite network. The two sets of the network are the learners’ set and the videos’ set. In the study, we found the pattern that learners watch videos obeys power-law distribution. The power-law comes from the analysis of the frequency of English words in the 1920s. The number of words commonly used is very small. Many words are not used frequently. Linguists found that the frequency of use of words and its priority of using are a constant power inverse relationship. Power-law actually has two popular explanations: one is the "long tail" theory, only a few large portals are concerned by many people, most portals are rarely followed, and this creates a long tail; the other is the Matthew effect, the rich get richer and the poor get poorer. Power-law distribution is a very common distribution in nature. Through recognizing learners’ learning behavior patterns of large-scale online course, we found that the number of videos watched by learners obey the power-law distribution, this rule is found in 22 courses offered on the iCourse. Our research provides a theoretical basis for mastering the rules of collective online learning.

Network Modeling

Learner-video network is a natural dipartite network. In the face of a large number of online course learning behavior data, our first thought is to establish a dipartite network. We used Geiph to build a learner-video dipartite network (Figure 1).

Here, we only show the learner-video dipartite network of the International oral English course, where the learners’ set has 3,324 learner nodes and the videos’ set has 37 video nodes. As shown in the figure, the red node in the network represents the learner node, and the blue node represents the video node. Since there are too many learner nodes, if they are arranged in an intuitive dipartite graph mode, the node arrangement will be too dense to be seen clearly, so this arrangement mode is adopted. We have a learner-video dipartite network, and we are very curious about what is the degree of node distribution?
The Degree Distribution of the Learners’ Nodes

In order to satisfy our curiosity, we study the degree distribution of the nodes in the learner-video dipartite network. Since there are few video nodes, we only studied the degree distribution of the learners’ nodes. As shown in Figure 2, we studied the learner node degree distribution of the four courses which are International oral English, Finance, Python language programming, and Advanced mathematics.

The size of the learner node degree represents the number of videos the learner watches. For example, the degree of the learner A node is 7, which means that the learner A has watched 7 videos. By visually observing the degree distribution of learner nodes in the dipartite network, we found that the majority of learners watched the video in a small number, and only a small number of
learners watched many videos. In other words, most of the learners are just segment-learners, however, a small number of learners participate in the whole course.

**Power Law Distribution Fitting**

By visually observing degree distribution of the learners’ nodes in the dipartite network, we infer that it obeys the power-law distribution. Let us test whether degree distribution of the learners’ nodes obeys the power-law distribution.

The power-law distribution formula is:

\[ Y = aX^{-b}, \]

Taking the base 10 logarithm of both sides of the formula:

\[ \log Y = \log aX^{-b} = \log a - \log X. \]

After taking logarithm of X, Y, the power-law distribution formula becomes a linear equation, so we use this method to test whether the learner node in the learner-video dipartite network obeys the power law distribution.

![Figure 3. Power-law distribution test.](image)

By fitting, we get the parameters a, b (table 1) in \( Y = aX^{-b}. \)

| Course                             | a       | b                |
|------------------------------------|---------|------------------|
| International oral English         | 9       | 1.68059851       |
| Finance                            | 9       | 1.72852902       |
| Python language programming        | 9       | 1.50640996       |
| Advanced mathematics               | 9       | 1.60948715       |

**Discussion and Conclusion**

In this study, we use the method of modeling network to analyze the learners’ learning behavior data of large-scale online courses. Through data processing, we establish a learner-video dipartite network to analyze the degree distribution of learner nodes. We find that the degree distribution of the learner node obeys the power-law distribution. That is to say, the number of videos watched by the learner obeys a power-law distribution. Based on this, we believe that the cost of clicks on the
internet is very low, learners are curious about the content of the course, so learners began to watch
the course video, but a large part of the learners have no patience, can’t insist on the whole process
of learning, the reaction on the power-law distribution is that a large part of the learner nodes’
degree is very small; a small part of learners are passionate about the content of the course, and
have the patience and perseverance to learn the whole course, the reaction in the power-law
distribution is that a small part of the learner nodes’ degree is very large.

In summary, we find that learner learning behavior obeys the power-law distribution in
large-scale online course, pay attention to experimental argumentation, and reveal the collective
online learning law. This theory provides a theoretical basis for website builders and course
instructors to master learners' learning behavior (most learners do not have the perseverance to
accomplish whole-course learning), and it is extremely urgent to increase the attractiveness of the
course.

References

[1] Yong Luo, Guochang Zhou, Jianping Li. Study on MOOC scoring algorithm based on Chinese
University MOOC learning behavior data[J]. Heliyon, 4(2018), e00960.

[2] Clauset, A., Shalizi, C. R., & Newman, M. E. J. Power-law distributions in empirical data.
SIAM Review, 51(2009), 661–703.

[3] R Lambiotte, M Ausloos. Uncovering collective listening habits and music genres in bipartite
networks[J]. Physical Review E, 72(2005), 066107.

[4] Tal Soffer, Tali Kahan, Eynat Livne. E-assessment of online academic courses via students' activities and perceptions[J]. Studies in Educational Evaluation, 54(2017), 83-93.

[5] S.A. Chapman, S. Goodman, J. Jawitz, A. Deacon. A strategy for monitoring and evaluating
massive open online courses[J]. Evaluation and Program Planning, 57(2016), 55-63.

[6] Justin M. Weinhardt, Traci Sitzmann. Revolutionizing training and education? Three questions regarding massive open online courses (MOOCs)[J]. Human Resource Management Review, doi.org/10.1016/j.hrmar.2018.06.004

[7] Anat Cohen, Orit Baruth. Personality, learning, and satisfaction in fully online academic
courses[J]. Computers in Human Behavior, 72(2017), 1-12.