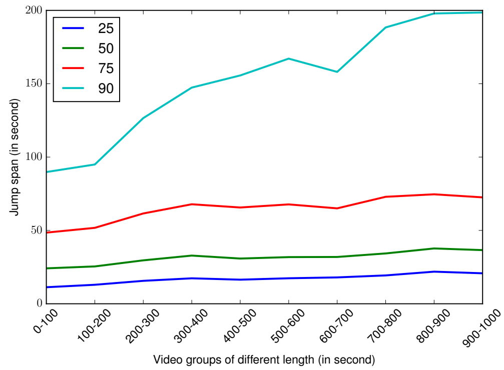

Figure 2: The distribution of different percentiles for jump span in different video length groups. Y-axis: position of jump span percentiles. X-axis: video groups of different length.

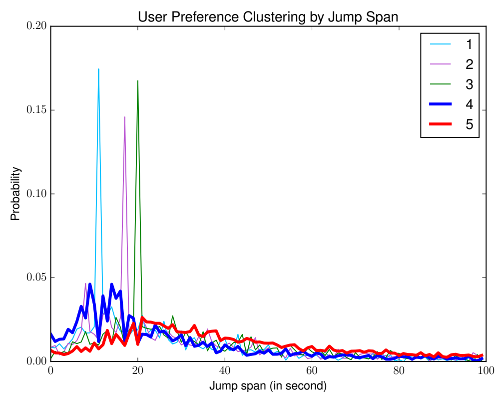

Figure 3: Cluster users into five categories by their complete-jump records. Y-axis: the probability of different jump spans. X-axis: jump span (in second). Category 1, 2, 3 represent users that have obvious preference, while category 4 and 5 represent users that almost have no preference.

general performance and user preferences. The investigation about the differences of general performances has two levels of granularity: (1) By course. After calculating the average start position and the span of complete-jumps, we found that users in non-science courses tend to rewind to the first half of a video, while users in science courses tend to minimize the rewind and only jump back from previous part of the video. (2) By video. Figure 2 shows the correlation between video length and jump span. In the investigation of user preferences, we categorize users into different types based on their jump span records leveraging k-means clustering. Figure 3 shows the result of user clustering.

3. METHOD

Based on the observations above, we extract a number of features which are adapted into a predicting model, i.e., Factorization Machine (FM). For each tuple (u, v, p_s, p_e), we define a set of features and construct a data instance $\mathbf{x}_i$, and compute the suggestion score by:

Table 1: Ranking performance of our method based on FM model and baseline method based on frequency with the measurement of hits@n

Course	Method	n = 1	n = 2	n = 3	n = 5
Science	Baseline	33.21	53.21	66.15	81.99
	FM	**37.05**	**60.40**	**76.04**	**89.59**
Non-science	Baseline	39.26	62.61	76.64	91.30
	FM	**42.25**	**72.42**	**88.43**	**96.05**

$$\hat{y}(\mathbf{x}_i) = w_0 + \sum_{j=1}^{d} w_j x_{i,j} + \sum_{j=1}^{d-1} \sum_{j'=j+1}^{d} x_{i,j} x_{i,j'} \langle \mathbf{p}_j, \mathbf{p}_{j'} \rangle \quad (1)$$

where $y(\mathbf{x}_i) \in [0, 1]$ indicates the likelihood of user u jumps to the corresponding position of $\mathbf{x}_i$; $\mathbf{p}_j, \mathbf{p}_{j'}$ are two k-dimensional latent vectors and $\langle \mathbf{p}_j, \mathbf{p}_{j'} \rangle$ models the interactions between variables $x_{i,j}, x_{i,j'}$ with the dot product of two latent vectors.

4. RESULTS AND DISCUSSION

We use the predicting result of FM to compare our automatic suggestion method based on machine learning model with the baseline method based on frequency through a ranking experiment. We use hits@n to measure the suggestion of the true end positions of all complete-jumps. Table 1 shows the result of the ranking experiment. We can see that our method based on machine learning model clearly outperforms the method based on frequency both in science courses and non-science courses.

In summary, we studied an interesting problem of automated navigation suggestions in MOOCs. We use a large collection of data from the courses of Xuetangx.com, providing investigating on jump-back behaviors from different perspectives. We found several interesting patterns and revealed the main factors that influence users' navigation behavior. Based on the discoveries, we developed a methodology aiming to understand the user intention and to suggest the best positions for a jump-back. Our experiments validate the effectiveness of the proposed method. We are also applying the method to a real online system and expect to have the function online very soon.

5. REFERENCES

[1] K. Chorianopoulos. Collective intelligence within web video. *Human-centric Computing and Information Sciences*, 3(1):1, 2013.

[2] K. Chorianopoulos, M. N. Giannakos, and N. Chrisochoides. Open system for video learning analytics. In *Proceedings of the first ACM conference on Learning@ scale conference*, pages 153–154. ACM, 2014.

[3] R. F. Kizilcec, C. Piech, and E. Schneider. Deconstructing disengagement: analyzing learner subpopulations in massive open online courses. In *LAK'13*, pages 170–179, 2013.

[4] D. T. Seaton, Y. Bergner, I. Chuang, P. Mitros, and D. E. Pritchard. Who does what in a massive open online course? *Communications of the ACM*, 57(4):58–65, 2014.