Methodology of measure of similarity in student video sequence of interactions.

Boniface Mbouzao
Polytechnique Montréal
boniface.mbouzao@polymtl.ca

Michel C. Desmarais
Polytechnique Montréal
michel.desmarais@polymtl.ca

Ian Shrier
McGill University
ian.shrier@mcgill.ca

ABSTRACT

Massive online Open Courses (MOOCs) make extensive use of videos. Students interact with them by pausing, seeking forward or backward, replaying segments, etc. We can reasonably assume that students have different patterns of video interactions, but it remains hard to compare student video interactions. Some methods were developed, such as Markov Chain and Edit Distance. However, these methods have caveats as we show with prototypical examples. This paper proposes a new methodology of comparing video sequences of interaction based both on time spent in each state and the succession of states by computing the distance between the transition matrices of the video interaction sequences. Results show the proposed methodology can better characterize video interaction in a task to discriminate which student is interacting with a video, or which video a student is interacting with.

Keywords

MOOC, Distance matrix, Edit Distance, Markov Chain, Optimal Matching Distance

1. INTRODUCTION

In online learning contexts, learner engagement is often measured by their interaction with video. The simplest measure is the total amount of time spent on video listening that can be used as an engagement measure [6]. But the availability of detailed interactions with a video allows more sophisticated measures, and comparison between video interactions.

Two common methods used to find the similarity between video interactions are the Markov Chain and Edit distance measures. The main limitation of using Markov Chain to compare video interactions sequences is that state transition probabilities do not take into account the time between states. Many sequences can have the same transitions probability matrix but represent different styles and length.

Boniface Mbouzao, Michel Desmarais and Ian Shrier "Methodology to measure of similarity in student video sequence of interactions." In: *Proceedings of The 13th International Conference on Educational Data Mining (EDM 2020)*, Anna N. Rafferty, Jacob Whitehill, Violetta Cavalli-Sforza, and Cristobal Romero (eds.) 2020, pp. 467 - 472

By contrast, the Edit distance approach to comparing video interaction sequences may take time into account if the sequences of events are mapped to a time scale and represented as activity segments, such as in [4]. However, large offset, such as a pause, in similar activity sequences will create large Edit distances that will shadow the similarity.

A methodology that can simultaneously take into account the time and transitions between activities could help the analysis of video interaction. It could help the analysis of the MOOCs and online teaching systems learning in video intensive environments, and could help to extract meaningful patterns of video interactions. It has often been used to classify students to identify students at risk (see for eg. [14, 8, 2]).

2. BACKGROUND

Among the different techniques to analyze video clickstream, some focus on extracting patterns, or motif, between events [3, 17, 16]. Descriptive statistics such as the video proportion played are also commonly used (see for eg. [15]). However, our focus is on measuring distance, or conversely similarity between video interaction patterns, and what are the most useful representations for that purpose.

We review the basics of the two families of methods and representations used in measuring video interaction similarity in more details and discuss their issues, before describing previous work with each approach, and then describe and evaluate the proposed method.

First, we describe the event data and a common transformation of events into activity sequences.

2.1 Events and activity sequences

Data on student interaction with videos relies on the notion of events associated to timestamps, such as "play" at 0:00:00 and "pause" at 0:00:10. There are five basic video interaction events: (1) *load*, (2) *play*, (3) *pause*, (4) *seek* and (5) *stop*.

The student can be considered in a state of listening to a video between 0 sec. and 10 sec., and in pause state thereafter. For example, suppose we have two students interactions:

Interaction sequences:

1: Play (4 seconds) then Pause (4 seconds) and then Play (4 seconds), 2: Pause (2 seconds) then Play (8 seconds) and then Pause (2 seconds).

Each student spent 12 seconds in total interaction with video. We can transform those two patterns of interaction into a sequence of activity states of 1 second intervals:

Activity sequences:

```
1:  Pl-Pl-Pl-Pl-Pa-Pa-Pa-Pa-Pl-Pl-Pl-Pl
2:  Pa-Pa-Pl-Pl-Pl-Pl-Pl-Pl-Pl-Pa-Pa-Pa
(Pl=Play and Pa=Pause)
```

We will name this type of sequence an *activity sequence*, where a polling interval is defined and the activity corresponds to the last event that occurred. Activity sequence encoding has been used in a few studies of student interaction patterns with a learning system [4, 1].

We now turn to how these sequences can be represented.

2.2 Markov Chain representation

A Markov chain is specified by a set of states and transitions between states. The process starts in one of the state s_i, then moves to another s_{i+1} with a probability of $p_{i,i+1}$. The Markov property stipulates that the transition probability is independent of states prior to s_i.

Considering the video interaction events as states, a student interaction can be represented as a Markov state transition matrix, where cells contain frequencies of transitions in the sequence, normalized such that row sums are 1, and thus represent transition probabilities.

For example, the two *interaction sequences* in the section above would result in the following event sequences:
Event sequences:

```
1:  Pl-Pa-Pl
2:  Pa-Pl-Pa
```

Contrary to *activity sequences*, *event sequences* do not carry the notion of a polling at regular time interval and ignore the time stamps on events. These event sequences would in turn result in a Markov Chain that is common to both:

$$\mathbf{M}_{\text{seq1.1}} = \mathbf{M}_{\text{seq2.1}} = \begin{array}{c c} & \begin{array}{cc} \textit{play} & \textit{pause} \end{array} \\ \begin{array}{c} \textit{play} \\ \textit{pause} \end{array} & \begin{pmatrix} 0/1 & 1/1 \\ 1/1 & 0/1 \end{pmatrix} \end{array}$$

A measure of distance between sequences can be computed from the two Markov matrices, such as the Frobenius norm of the cell-wise difference between the matrices. More on this below.

The limit of using Markov Chain to compare video event sequences lies in the fact that transitions probabilities can be the same for very different sequences. This issue is evident in the two sequences above that end up having the same Markov transition matrix. While it can be alleviated by having a start end state, it is clear that the loss of state duration information will lead to a loss of valuable information.

However, Markov Chains are efficient at capturing transition patterns and have been used with some success for clustering [12, 11], for creating student profiles of interactions [10, 5], and for simulated students [7].

2.3 Sequence Edit Distance method

The sequence Edit Distance method relies on measures found with word distances, where alphabet similarity between words is the basis of calculating similarity.

Edit Distance (ED), generates distances that represent the minimal cost in terms of insertions, deletions and substitutions for transforming one sequence to another. The cost of each deletion, insertion or insertion is 1 by default. This algorithm was originally proposed by Levenshtein [9] and is most common when computing distances between words [13]. For video listening sequences, the principle is the same but the alphabet is represented by the activity. For example, the ED measure for activity sequences 1 and 2 above yields a distance of 9 over a maximum of 12.

A notable property of the ED measure is that sequences of different lengths will necessarily have a non null distance, and therefore potentially miss regularities in interaction patterns of different length sequences. On the contrary, a Markov Chain representation is not sensitive to sequence length, or to the number of transitions for that matter (since the row sums are all normalized to 1), whilst its capacity to capture interaction patterns in sequences of different length.

3. PROPOSED METHOD, TMED

The proposed method, named TMED, is a combination of the two techniques: the Markov Chain and the ED measure. The combination of results give a full similarity between each pair of student sequences of interactions benefiting of advantages from both techniques.

3.1 Transition matrix

The video transition matrix of a student s for a video is expressed as:

$$\mathbf{M}_s = \begin{array}{c c} & \begin{array}{ccccc} \textit{load} & \textit{play} & \textit{pause} & \textit{seek} & \textit{stop} \end{array} \\ \begin{array}{c} \textit{load} \\ \textit{play} \\ \textit{pause} \\ \textit{seek} \\ \textit{stop} \end{array} & \begin{pmatrix} m_{1.1} & m_{1.2} & m_{1.3} & m_{1.4} & m_{1.5} \\ m_{2.1} & m_{2.2} & m_{2.3} & m_{2.4} & m_{2.5} \\ m_{3.1} & m_{3.2} & m_{3.3} & m_{3.4} & m_{3.5} \\ m_{4.1} & m_{4.2} & m_{4.3} & m_{4.4} & m_{4.5} \\ m_{5.1} & m_{5.2} & m_{5.3} & m_{5.4} & m_{5.5} \end{pmatrix} \end{array}$$

where $m_{j,k}$ is the *number of transitions* from event j to event k in an *activity sequence* obtained from an *interaction sequence*. And $\mathbf{M}_s$ is the transition matrix of student s interacting with a video. Contrary to a Markov Chain, rows do not necessarily sum to 1. In the case where no event occurs and the student remains in the same state for awhile (playing video or pausing video, for eg.) the increase of the matrix element m_i is the maximum number of transitions possibles within the time spent in that state counting the transition from one state to the same state.

3.2 Distance between two transition matrices

The distance between two student transition matrix is expressed as:

$$\begin{aligned} d(\mathbf{M}_{s1}, \mathbf{M}_{s2}) &= \|\mathbf{M}_{s1} - \mathbf{M}_{s2}\|_F \\ &= \sqrt{\sum_{i=1}^{5} \sum_{j=1}^{5} (m_{s1,j} - m_{s2,j})^2} \end{aligned}$$

An important question is what is the polling interval to choose. This interval will determine the total number of transitions in M_s. The choice is determined by the minimal interval required to avoid skipping events while transforming the event sequence to the activity sequence. In our case, this interval is set to 3 per second and it applies to all video interactions. The total number of transitions, $T_{s,i}$ in a given interaction matrix $\mathbf{M}$ for sequence s and video i is therefore:

$$T_{s,i} = L_{s,i} * N \quad (1)$$

where $L_{s,i}$ is the length of the interaction time and N is the polling interval.

The similarity between two interaction video sequences based on transition matrices with a video is then expressed as:

$$S_{mat}(\mathbf{M}_{s1}, \mathbf{M}_{s2}) = 1 - Dis(\mathbf{M}_{s1}, \mathbf{M}_{s2}) \quad (2)$$

$$Dis(\mathbf{M}_{s1}, \mathbf{M}_{s2}) = \frac{d(\mathbf{M}_{s1}, \mathbf{M}_{s2})}{T_{s1} + T_{s2}} \quad (3)$$

Where $S_{mat}(\mathbf{M}_{s1}, \mathbf{M}_{s2})$ is the similarity level between sequence of interaction of student $s1$ and student $s2$ of video i using matrix of interactions and $Dis(\mathbf{M}_{s1}, \mathbf{M}_{s2})$ is the dissimilarity between them. $d(\mathbf{M}_{s1}, \mathbf{M}_{s2})$ is the distance among them. T_{s1} and T_{s2} are the number of transitions of student $s1$ and student $s2$ sequence of the video i. If $S_{mat}(\mathbf{M}_{s1}, \mathbf{M}_{s2})$ is 0 then the two sequences are completely dissimilar and when it is 1 then they are completely similar. Between 0 and 1 shows the percentage of similarity between the two sequences of transitions.

3.3 Edit Distance measure (ED)

For each pair of sequences, we compute the ED distance to obtain the distance matrix and from there compute the level of similarity among them. The level of similarity between two sequences is computed using ED distance as:

$$S_{om}(seq_{s1}, seq_{s2}) = 1 - \frac{dist_{om}(seq_{s1}, seq_{s2})}{max(T_{s1}, T_{s2})} \quad (4)$$

Where $S_{om}(seq_{s1}, seq_{s2})$ is the similarity level between sequence of student $s1$ and sequence of student $s2$ of video i and $dist_{om}(seq_{s1}, seq_{s2})$ is the ED distance between the two sequences and T_{s1} and T_{s2} are the numbers of transition of the sequence of each student given in equation (1). $max(T_{s1}, T_{s1})$ is the maximum between the number of transitions of the two student sequences of interactions.

3.4 TMED

The last step of this proposed methodology is to combine the two techniques by taking for each pair of sequences the proper level of similarity among the levels given by each technique. This is meant to take into account the for complementary of those techniques: one can find styles and give good similarity for sequences of different lengths and the other gives regularity among sequences and gives good similarity among sequences from the same range length. The final similarity level is then given by:

$$S(seq_{s1}, seq_{s2}) = Select(S_{om}(seq_{s1}, seq_{s2}), S_{mat}(\mathbf{M}_{s1}, \mathbf{M}_{s2})) \quad (5)$$

Where $S(seq_{s1}, seq_{s2})$ is the level of similarity between sequence of interaction $s1$ and $s2$, $S_{om}(seq_{s1}, seq_{s2})$ similarity level between the two sequences based on ED distance as expressed in equation (4) and $S_{mat}(\mathbf{M}_{s1}, \mathbf{M}_{s2})$ similarity level between the two sequences based on sequence matrix as expressed in equation (2).

The function $Select()$ selects S_{mat} similarity if one of the two sequences is less than the half-length of the other, and selects the maximum level of similarity between the proposed method and the ED method otherwise.

One takes the maximum between ED similarity and matrix similarity to avoid the ED drawback of finding dissimilarity between sequences of same range of length but some mismatch between states as illustrated in section 4 below. The flow of the proposed method is illustrated in Figure 1 from the sequences to the computation of their similarity level.

4. VALIDATION

To validate the proposed method, we compare its capacity of finding the level of similarity between sequences with existing methods, namely the Markov Chain technique as used by Klingler et al.[8] and the ED based method used for clustering the same kind of sequences of interactions.

4.1 Prototypical cases

We first test the approach over prototypical cases where the patterns are obvious to the eye. For this purpose we take two main cases: sequences of same lengths of transitions and sequences of different length of transitions. For the same sequence length interactions, we considered a cyclic same sequence of transitions as illustrated in Figure 2a. The cycle of transitions is: `Lo-Pl-Pa-Pl-Pa-Se-Pl-St`. The cycle of transition can start anywhere and finish by St for any of the sequence.

The expected level of similarity should be close 100% as it is the same sequence following a cycle. The result based on ED distance cannot find that level of similarity as shown in Figure 2b compared to the Markov based method in Figure 2c (with some exceptions which do not reach the 100% similarity as expected, but close enough to be considered as such) and the proposed method in Figure 2d (finds perfect match of style by 100% similarity in each case). For these cyclic sequences, the proposed method and the Markov based similarity methods are performing better than ED based method in finding similarity between two cyclic same sequences of interactions.

The second validation of the proposed method is to compare it to a Markov based method for different length sequences given known similarities. For this purpose, we considered four sequences of same transitions levels as shown in Figure 3a. In this case, the percentage of transition between states is the same, but the time spent in each state is different from one sequence to another. The expected level of similarity depends here on the lengths of each sequence as the succession of states are the same for all four sequences. We should have then as result a progressive increase in level of similarity from the shortest sequence to the longest.

The result from the Markov Chain based method as in Figure 3a could not find the different levels of similarity as the percentage of transition between the states is preserved with

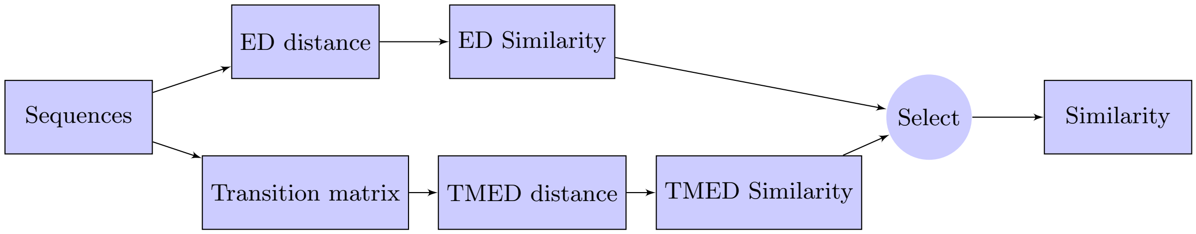


Figure 1: Flow of the proposed method to compute similarity between students' video sequences. "Select" is the selection process between the two technique similarity.

different sequences lengths. The proposed method performs better as shown in Figure 3b because it is based on the number of transitions rather than probability of transition as Markov Chain is.

4.2 Real dataset

The experiment on a real set of video interaction logs aims to test and compare the ability of the proposed method to recognize (1) the student behind an interaction log (data contains 4800 students), and (2) the video behind an interaction log. While this task is of no practical use, since both the video and student associated with an interaction log are already known in general, it provides a ground truth dataset to assess the discrimination power of each approach.

We choose three well-known classifiers such as support vector machine (SVM), boosted tree (GBM) and K-nearest neighbor (KNN) for each method of representation of sequence of interactions to predict first the student and then video to which sequence of interaction belongs. If a specific representation of student sequence of interaction is predictable in terms of which video and student that interact with the video, that means that the representation is able to better distinguish different types of interaction and even showing the specificity of a video in the way that students interact with it.

For the first part of the experiment where we predict student to which the sequence representation belongs, the algorithm arbitrarily selects one sequence of each student to predict among the nine (9) same student sequences representation and trains on the eight (8) others student sequences representation. The matrix distance used for Markov chain sequence representation and the proposed TMED representation is the one described above in section 2.2. In these two cases the dimension of the representation of each sequence is 25, that represent the 25 elements of transition matrix of each sequence representation as described in section 3.3. For the OM sequence representation, the matrix distance used is the one described in section **??** above. For the prediction 80% of the data is use for training and 20% for prediction. Each experiment is repeated 400 times using different set of students to predict (from 3 to 15 students). The data set is organized in such that all the student sequence present in the data set selected, the training set has 8 of their sequence representation and one in the testing set in each prediction run.

In the second part of the experiment, we used the same representations of student footage but instead of predicting the student, we predicted the video the student interacted with. We used the same training (80% of the data) and test (20% of the data) sets, making sure that in the data we had the same number of students interacting with each video. Since each student has nine (9) sequences of interaction representation, the number of predicted classes (video 1 to video 9) in each data set considered is the same regardless of the number of students considered. For this reason, balanced precision was included in the results to avoid the effect of having more students. Again, in this case, at each prediction run, the algorithm ensures that each student sequence representation in the data set considered is the same as its sequence representations in the test set in each run.

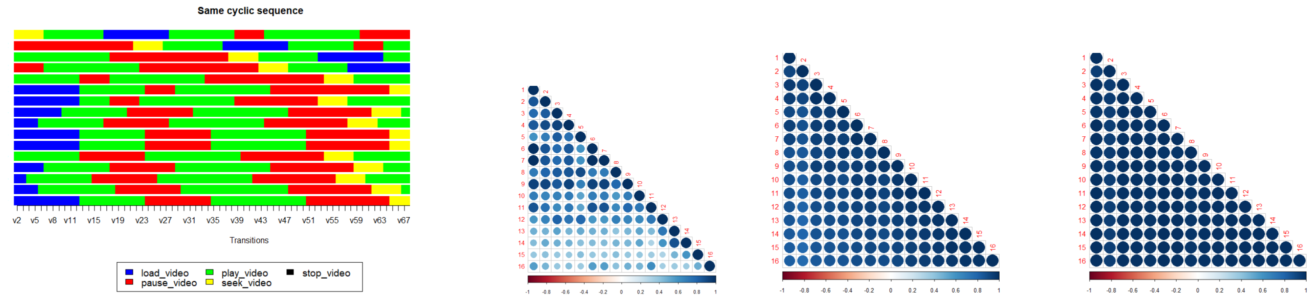
4.3 Real data results

The results show that the proposed TMED method through the level of similarity. Through the tests of validation on prototypical data, the proposed method yields better results than the other two existing methods as one can see through Figures 2 and 3. For the same sequence represented as a cyclic sequence of interaction with various ways of representation show in Figure 2 (a) the expected degree of similarity 100% but only the proposed method give us the closest results to the expected one as shown in Figure 2. One can also see in this figure that the Markov chain based similarity is the second-best estimation of similarity after the proposed method based one.

When we consider a same sequence of states with different lengths of time as shown in Figure 2 (a), the expected results of similarity is a progressive increase of level of similarity according to the length of the sequence. The classic Markov chain based method could not find that the length of sequences are different whereas TMED method is able to find it well (Figure 3 (c)).

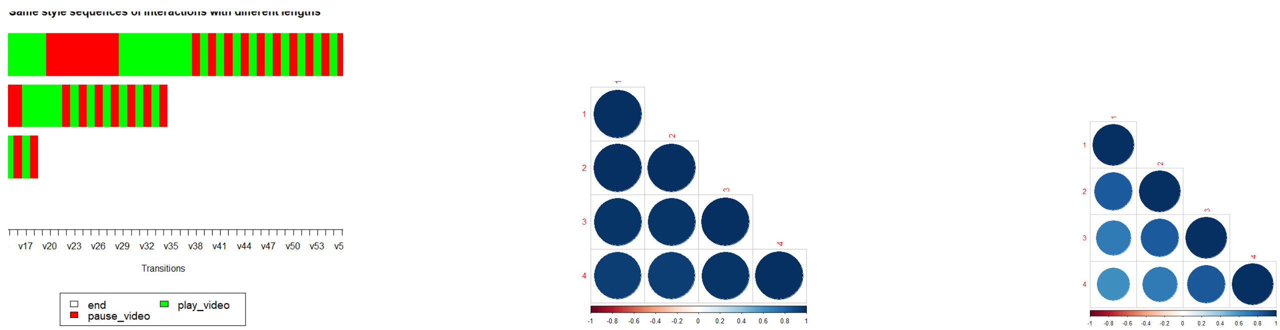
The experiment over the real data tasks tests the capacity of each method of representation of video interaction to identify each sequence of interaction in terms of student and video sequences. Results show better accuracy for TMED than the other ones (table 1). The performance parameters on student prediction using SVM, GBM and KNN on predicting five (5) students and twelve (12) students with nine (9) records of each student (where eight (8) records are for training and predicting one record of each student).

For predicting video, the complete results for forty-five (45) records from five (5) different students and hundred and eight (108) records from twelve (12) students in predicting the nine (9) videos are shown in table 2. They demonstrate



(a) Same cyclic sequence of transitions (b) ED (c) Markov (d) Proposed

Figure 2: Result of similarity: (a) The cycle starts and follows the same pattern of transition to close the cycle (b) Similarity based on Edit Distance (ED) cannot recognize the similarity of cyclic sequences. (c) Similarity based on Markov Chain can recognize the similarity, with some exceptions that not reach 100%. (d) Proposed TMED similarity can recognize cyclic sequences.



(a) Same cyclic sequence (b) Markov (c) Proposed

Figure 3: Similarity results from the sequence in (a): (b) similarity based on Markov Chain cannot recognize the duration in each state.(c) proposed TMED similarity can recognize the fact that those sequences are same but the level of similarity is based on the time spent in each state.

Predictions:	45 records, 5 target students								
Approach:	SVM			GBM			KNN		
Method:	ED	MC	TMED	ED	MC	TMED	ED	MC	TMED
Accuracy	0.60	0.00	**0.80**	0.40	0.00	**1.00**	0.20	0.22	**1.00**
F_1	0.75	0.00	**0.89**	0.57	0.00	**1.00**	0.33	0.36	**1.00**
Predictions:	108 records, 12 target students								
Accuracy	0.58	0.18	**0.67**	**0.42**	0.36	**0.42**	0.11	0.00	**0.40**
F_1	0.73	0.20	**0.78**	0.59	0.50	**0.63**	0.20	0.00	**0.67**

Table 1: Results of Twenty fold cross validation 400 runs of student prediction of 5 and 12 students using three different methods of representation of student interaction with videos showing that the proposed representation technique is performing better than others.

that the proposed method is also better on recognizing both video and student than the two other methods of presentation of student interaction with video.

These results suggest that the proposed method has a better way of representing a student video interaction with videos and so can be used for comparing two different interactions with video.

5. CONCLUSION

The proposed methodology aims to fill out a methodological gap on representing and comparing video sequences of interaction methods. The proposed method overcomes the drawbacks of the previous methods based on Markov Chain and sequence of interactions known as Edit Distance (ED). The main contribution of this proposed method is the fact that it takes into account the time spent in each state and the general style of succession of states. This offers a new tool to researchers who want to compared video viewers interaction and find eventually video style of interaction.

Predictions:	45 records, 9 target videos								
Approach:	SVM			GBM			KNN		
Method:	ED	MC	TMED	ED	MC	TMED	ED	MC	TMED
Accuracy	0.11	0.33	**0.56**	0.33	**0.56**	**0.56**	0.22	0.22	**0.33**
F_1	0.20	0.50	**0.72**	0.50	0.36	**0.72**	0.36	0.36	**0.50**
Predictions:	108 records, 9 target videos								
Accuracy	0.22	0.11	**0.56**	0.11	0.33	**0.56**	**0.22**	0.11	**0.22**
F_1	0.36	0.20	**0.61**	0.20	0.50	**0.61**	**0.36**	0.20	**0.36**

Table 2: Results of Twenty fold cross validation 400 runs of video prediction using three different methods of representation of student interaction with videos, ED (Edit Distance), MC (Markov Chain), TMED.

TMED combines two styles of representation of video sequence of interaction and computes the similarity based on the advantage of each style of representation. The ED based similarity is generally good on same range length of interaction sequences and the matrix of interaction based representation does better on sequences of different range of length.

The proposed method is also able to better represent a sequence of interaction when doing classification tasks as the results show. In fact, proposed method has a better performance in predicting student sequence of interaction and prediction video when having a representation of a video sequence of interaction.

References

[1] Yoav Bergner, Zhan Shu, and Alina von Davier. Visualization and confirmatory clustering of sequence data from a simulation-based assessment task. In *Educational Data Mining 2014*, 2014.

[2] Mina Shirvani Boroujeni and Pierre Dillenbourg. Discovery and temporal analysis of latent study patterns in MOOC interaction sequences. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, pages 206–215. ACM, 2018.

[3] Christopher G Brinton, Swapna Buccapatnam, Mung Chiang, and H Vincent Poor. Mining MOOC clickstreams: Video-watching behavior vs. in-video quiz performance. *IEEE Transactions on Signal Processing*, 64(14):3677–3692, 2016.

[4] Michel Desmarais and François Lemieux. Clustering and visualizing study state sequences. In *Educational Data Mining 2013*, 2013.

[5] Louis Faucon, Lukasz Kidzinski, and Pierre Dillenbourg. Semi-markov model for simulating MOOC students. *International Educational Data Mining Society*, 2016.

[6] Philip J Guo, Juho Kim, and Rob Rubin. How video production affects student engagement: An empirical study of MOOC videos. In *Proceedings of the first ACM conference on Learning@ scale conference*, pages 41–50. ACM, 2014.

[7] Robert Hlavatỳ and Ludmila Dömeová. Students' progress throughout examination process as a markov chain. *International Education Studies*, 7(12):20–29, 2014.

[8] Severin Klingler, Tanja Käser, Barbara Solenthaler, and Markus H Gross. Temporally coherent clustering of student data. In *EDM*, pages 102–109, 2016.

[9] Vladimir I Levenshtein. Binary codes capable of correcting deletions, insertions, and reversals. In *Soviet physics doklady*, volume 10, pages 707–710, 1966.

[10] Alice Marques and Orlando Belo. Discovering student web usage profiles using markov chains. *Electronic Journal of e-Learning*, 9(1):63–74, 2011.

[11] Sylvain Mongy, Fatma Bouali, and Chabane Djeraba. Analyzing user's behavior on a video database. In *Multimedia data mining and knowledge discovery*, pages 458–471. Springer, 2007.

[12] Sylvain Mongy, Chabane Djeraba, and Dan A Simovici. On clustering users' behaviors in video sessions. In *DMIN*, pages 99–103, 2007.

[13] Gonzalo Navarro. A guided tour to approximate string matching. *ACM computing surveys (CSUR)*, 33(1):31–88, 2001.

[14] Nirmal Patel, Collin Sellman, and Derek Lomas. Mining frequent learning pathways from a large educational dataset. *arXiv preprint arXiv:1705.11125*, 2017.

[15] Tanmay Sinha. " your click decides your fate": Leveraging clickstream patterns from mooc videos to infer students' information processing & attrition behavior. *arXiv preprint arXiv:1407.7143*, 2014.

[16] Tanmay Sinha, Patrick Jermann, Nan Li, and Pierre Dillenbourg. Your click decides your fate: Inferring information processing and attrition behavior from mooc video clickstream interactions. *arXiv preprint arXiv:1407.7131*, 2014.

[17] Tanmay Sinha, Nan Li, Patrick Jermann, and Pierre Dillenbourg. Capturing "attrition intensifying" structural traits from didactic interaction sequences of MOOC learners. *arXiv preprint arXiv:1409.5887*, 2014.