



Unfolding Sentimental and Behavioral Tendencies of Learners' Concerned Topics From Course Reviews in a MOOC

Journal of Educational Computing

Research

0(0) 1–27

© The Author(s) 2018

Reprints and permissions:

sagepub.com/journalsPermissions.nav

DOI: 10.1177/0735633118757181

journals.sagepub.com/home/jec



Sannyuya Liu¹, Xian Peng¹, Hercy N. H. Cheng¹,
Zhi Liu¹, Jianwen Sun¹, and Chongyang Yang¹

Abstract

Course reviews, which is designed as an interactive feedback channel in Massive Open Online Courses, has promoted the generation of large-scale text comments. These data, which contain not only learners' concerns, opinions and feelings toward courses, instructors, and platforms but also learners' interactions (e.g., post, reply), are generally subjective and extremely valuable for online instruction. The purpose of this study is to automatically reveal these potential information from 50 online courses by an improved unified topic model Behavior-Sentiment Topic Mixture, which is validated and effective for detecting frequent topics learners discuss most, topics-oriented sentimental tendency as well as how learners interact with these topics. The results show that learners focus more on the topics about course-related content with positive sentiment, as well as the topics about course logistics and video production with negative sentiment. Moreover, the distributions of behaviors associated with these topics have some differences.

¹National Engineering Research Center for E-Learning, Central China Normal University, Wuhan, PR China

Corresponding Author:

Xian Peng, National Engineering Research Center for E-Learning, Central China Normal University, No. 152, Luoyu Road, Wuhan 430079, PR China.

Email: px87374006@126.com

Keywords

behavioral and sentimental analytics, topic modeling, learning analytics, behavior-sentiment topic mixture

Introduction

Massive Open Online Courses (MOOCs) provides new sources of data and opportunities for large-scale experiments that can advance teaching and learning (Reich, 2015). As the advent of learning analytics and educational data mining (Siemens & Baker, 2012), most previous works were mainly conducted from structured data by the quantitative analysis in MOOCs (Halawa, Greene, & Mitchell, 2014; Kizilcec, Piech, & Schneider, 2013; Wen & Rosé, 2014), for example, the statistics of learners' clickstream activities and exam grades, for the purpose of discovering individual preferences, engagement patterns, and revealing the relationship between diverse interactive behaviors, and final performance in MOOCs (Guo, Kim, & Rubin, 2014; Taylor, Veeramachaneni, & O'Reilly, 2014). Nowadays, a large amount of unstructured textual data is produced by the application of interactive technologies in MOOCs, such as discussion forums, course reviews, or chat rooms. By means of these tools, learners are usually willing to express themselves when they engage in the process of learning activities. Therefore, the qualitative analysis of online textual conversations produced by learners can be adopted to obtain in-depth understanding of their implicit psychological features and subtle behaviors. Among various interactive tools in MOOCs, this study is particularly interested in course reviews, which is thought to be more appropriate for our study. Because unlike in forums where learners have more objective technical discussions related to the professional knowledge for specific courses (Brinton et al., 2014; Klisc, McGill, & Hobbs, 2017), course reviews contain more personal concerns and subjective opinions directly related to content of courses, style of instructors, construction of platform, and so forth.

Currently, owing to the ever-growing scale of learners and learners-generated reviews, it is quite time and labor consuming for both learners and instructors to manually read through all reviews for courses. Consequently, this has brought great challenges to observe and track the focused topics hidden in learners-generated reviews. In addition, sentimental detection for different topics is as important as what learners really care about. While learners are encouraged to rate a course after completion, these courses' marks are typically similar such that it is challenging to accurately identify and assess their emotional tendencies. Moreover, merely making an evaluation at the level of domains (e.g., courses and products) is apparently insufficient (El-Halees, 2011; Leong, Lee, & Mak,

2012; Wan, 2011). For example, when a learner hesitates to choose an appropriate course, he or she may want to know not only the topics that other learners frequently talk about (e.g., course content and video production) but also their attitudes toward these topics, such as discontent over the difficulty of course. In this case, it is imperative to discern the latent topics involved in learners-generated reviews and their associated sentiments toward various topics (Ramesh, Goldwasser, Huang, Daumé, & Getoor, 2014; Wen, Yang, & Rose, 2014; Xu, Zhang, & Wang, 2015).

Additionally, it is also critical to understand explicit behavioral tendencies of learners' interaction with the texts in social learning environments. Qiu, Zhu, and Jiang (2013) applied Latent Dirichlet Allocation (LDA) to a Twitter data set, but they found that the interest of topics of some users was hardly distinguished because of their almost identical topic distributions. When examining users' behavioral patterns, they found a significant difference. Furthermore, some of the users were more likely to publish original posts while others mostly tended to engage in interactions such as reply, retweet, and mention. More importantly, Zhao, Cheng, Hong, and Chi (2015) claimed that users typically presented different types of behaviors around different topics in social media. Moreover, by constructing learners' online social profiles characterized from their behavioral signals, more personalized topic recommendations could be accurately provided for different types of behaviors. In a sense, the behavior types (e.g., post a review or respond to a review) should be regarded as an important feature in connecting with learners' evaluated topics.

Therefore, in order to detect learners' focused topics, topics-oriented sentimental tendencies, as well as how learners interact with these topics, this study proposes a machine learning method called Behavior-Sentiment Topic Mixture (BSTM) to promote effective practices in social learning environments, which is an improved unified topic model. In BSTM, both sentiment and behavior features are taken into account on the exploitation of latent semantic topical space of textual reviews. As a result, for learners, they could quickly look at the detailed various aspects of course evaluation (El-Halees, 2011), thus making an optimal choice of which courses are more appropriate; for educators, they could identify learners' psychological problems and get prepared for timely intervention in the learning process (Leong et al., 2012); for administrators, they could refine the function and design of platform to improve learners' interactive experiences (Brinton et al., 2014).

Related Work

In this section, abundant work related to our research is introduced from two following aspects, text mining in E-learning context and sentimental detection and behavioral analysis based on topic model.

Text Mining in E-Learning Context

In the field of E-learning, text mining is regarded as one of the core technologies promoting the application of learning analytics (Leong et al., 2012; Simsek, Shum, Liddo, Ferguson, & Sándor, 2014; Yu & Luna, 2013). Understanding textual contents generated from online learners can uncover their concerns, attitudes, and even behavior motivation. Rather than focusing on the learners' summative assessment in the online education community, the applications of text mining in E-learning are more emphasized to assess the learners' learning process. Such applications have covered a wide range of topics including mining of learners' opinions (El-Halees, 2011; Leong et al., 2012), assessment of learners' knowledge ability (Klein, Kyrilov, & Tokman, 2011), warning of learners' behavioral risks (Ehrenreich, Underwood, & Ackerman, 2014), visualization of learners' profiles (Simsek et al., 2014), and so forth. We note that among these topics, the mining of learners' opinion is more relevant to our study. For example, Yu and Luna (2013) employed the text mining techniques to detect and extract semantic knowledge from comments in an E-learning system. The result demonstrated that the key critical topics and praiseful topics from learners were identified.

Starting from 1994, many prior works have paid attention to the sentimental states of learners in the learning environments (Altrabsheh, Coccoa, & Fallahkhair, 2015; O'regan, 2003; Sylwester, 1994). Compared with recognition of physiological features (Woolf et al., 2009), facial expressions (Grafsgaard, Wiggins, Boyer, Wiebe, & Lester, 2013), and speeches signals (Bahreini, Nadolski, & Westera, 2016), text-oriented sentimental analysis is the least invasive for online learners' interaction. Along this way, there are two main lines of researches in the field of sentiment analysis applied in E-learning settings. On the one hand, some studies first adopted opinions mining methods to discern learners' sentimental polarities derived from textual interactive activities toward courses, and then provided dynamic feedbacks for instructors to adjust their teaching methodologies and for learners to enhance their self-awareness (El-Halees, 2011; Kontogiannis, Valsamidis, Kazanidis, & Karakos, 2014; Leong et al., 2012; Rodriguez, Ortigosa & Carro, 2012; Yu & Luna, 2013). On the other hand, some studies tended to explore the relationship between the sentimental tendencies of learners and dropout rate (Adamopoulos, 2013; Ramesh, Goldwasser, Huang, Daumé, & Getoor, 2013; Ramesh et al., 2014; Wen et al., 2014). Interestingly, a novel view was indicated by Wen et al. (2014) that students with positive sentiment were not consistent with an enjoying course experiences, and students with negative sentiment were not consistent with a frustrating course experiences.

Given the current works on the exploration of learners' behavioral interaction and participation of text-based content in the online learning community (Anderson et al., 2012, 2014; Ehrenreich et al., 2014), most of them have

simply utilized a series of statistical indicators, such as the number of posts, to discover learners' behavioral differences (Romero, López, Luna, & Ventura, 2013), thus revealing their behavioral regularities and even predicting their academic success (Tobarra, Robles-Gómez, Ros, Hernández, & Caminero, 2014). While these statistical indicators, to some extent, can be beneficial to explain learners' behavioral preferences and profiles, it is not sufficient to get a deeper understanding of behavioral analysis. This is because these behavioral indicators do not map any semantic information-hidden textual contents.

Sentimental Detection and Behavioral Analysis Based on Topic Model

Since the growing amounts of electronic archives or corpus has brought administrators huge burdens in categorizing and querying, it is almost impossible for them to accomplish this task just depending on human intervention. To tackle this problem, Blei, Ng, and Jordan (2003) proposed the LDA model for the first time, which was known as a popular unsupervised topic model that could automatically identify potential thematic meaning in large-scale textual sets. In subsequent years, extending on the basic LDA, many variants considering the dimension of time (Fu, Yang, Huang, & Cui, 2015), topical structure (Grimmer, 2010), and metadata information (Ramesh et al., 2014) had been applied in diverse fields. However, these researchers often offered an automatic process for detecting potential textual topics without considering the feature of sentiment, especially in the context of online education. In recent years, the studies on topic modeling are presented by combining topic and sentiment collectively in a document. From the view of unified model generated, it has shown to be superior to a single topical language model. For example, Mei, Ling, Wondra, Su, and Zhai (2007) first introduced a novel probabilistic model called Topic Sentiment Mixture (TSM), in which sentiment is considered as a topic independent language model and each word came from either certain topic or sentiment. Ramesh et al. (2013, 2014) proposed a probabilistic soft logic model, which embeds substantive linguistic features derived from seeded topic analysis, in order to extract topics and the corresponding polarities of sentiment of forum posts in MOOCs. These models do not establish a clear mapping relation between topics and sentiments and therefore cannot be explicitly used to separate topics and sentiments. Jo and Oh (2011) extended Sentence-LDA to Aspect and Sentiment Unification Model (ASUM), which could discover pairs of sentiments and topics. ASUM posited that each document was a mixture over positive and negative sentiments, and each sentiment drew the probabilistic distribution over the specific evaluative topics. By contrast, BSTM focuses on discerning a series of underlying topics in each review and these topics' corresponding sentimental polarities.

As a well-known social scientist Erving Goffman (1978) said, people created different images to other audiences in their everyday life and engaged in

conversations with different topics. In other words, they usually interacted with different topics by using different types of behaviors. Zhao et al. (2015) employed a matrix factorization technique to build a topic recommender, for improving user's topical interest profiles according to their various behavioral signals. Their results showed that this technique was capable of providing more personalized topic-based content recommendations by separating a user's profiles into several behavioral profiles. Although this study does not adopt the method of topic modeling, it clearly indicates that user behavior types should be utilized as an important contextual feature. A similar work to our study, Qiu et al. (2013) introduced an LDA-based behavior-topic model combining user topic interests and behavioral patterns in online social network settings. Nevertheless, these studies usually ignore the distinct feature of sentiment derived from textual content which has not been well embedded in the existing topic models. By contrast, sentiment information on evaluating the detailed topics of reviews and behavior information on expressing these specific topics are all considered to guide the generative process of BSTM. Furthermore, because of the discrepancy of context, the existing models may not be applicable to the analysis of online learners' course reviews in MOOCs settings.

Research Questions

Many researchers have made great progress in discerning the collection of hidden topics and the corresponding sentiments toward these topics effectively in the field of business intelligence (Jo & Oh, 2011; Xu et al., 2015) and social network (Fu et al., 2015; Mei et al., 2007). Until now, only a few studies have incorporated behavior feature derived from their interactions with the textual contents into topic modeling in social media platforms (Qiu et al., 2013; Zhao et al., 2015). However, to the best of our knowledge, there are no studies simultaneously incorporating sentiment and behavior features into topic modeling, in order to automatically discover sentimental and behavioral tendencies of learners' concerned topics in social learning setting. Therefore, this study puts emphasis on the automatic analysis of course reviews from the three dimensions including topic, sentiment, and behavior detection in a MOOC platform. Additionally, in view of the optimization of experimental results, the effectiveness of our proposed BSTM needs to be validated. More specifically, this study targets at four research questions described as follows:

1. Compared with the standard models, can BSTM achieve better performance? And how to determine the suitable number of topics to ensure the best performance of BSTM?
2. What are the most concerned topics mined from course reviews among learners in a Chinese MOOC?

3. What are the associated sentimental tendencies toward learners' concerned topics?
4. How do learners interact with their concerned topics?

BSTM Analysis Model

To answer the above four questions, a unified topic model called BSTM model is proposed in this section, which will be utilized in exploiting implicit topics, emotion information, as well as explicit interactions inferred from online learners' course reviews in a MOOC.

Study Preparation

- Seed words: Seed words are commonly consisted of basic features represented by coherent concepts that make up a topic and are closely related to courses. Some of them are selected from the course syllabus and collected from non-standard Web words generated from learners, for example, “*学霸*/super scholar” and “*潜水*/lurk.”
- Sentiment words: Sentiment words, which contain subjective emotions, are usually used to express individual opinions, attitudes, and views toward diverse aspects. In our study, a Chinese sentiment lexicon is set up, including 9,594 positive words and 12,884 negative words, that is, a total of 22,478 sentiment words.
- Behavioral categories: In the Chinese MOOC platform called *Guokr Web*, learners are able to post reviews, reply to reviews, like reviews, and even show their earned certificates. Note that showing earned certificate is a special behavior and is considered as an effective incentive to encourage learners to increase their learning participation. For learners, it may show a kind of self-affirmation.
- Review structure: A review consisted of several sentences, each of which is assumed to be a specified topic, a corresponding sentiment label, and a dominant behavior. Each sentence is composed of several words, and each one either belongs to topic-specific words or sentiment words.

Model Description

BSTM is an extension of Sentence-LDA (Jo & Oh, 2011) that imposes a constraint that all words in a single sentence are generated from one topic on a sentence level, instead of a document level for each word corresponding to one topic. It assumes that when generating one sentence of a review, we first choose

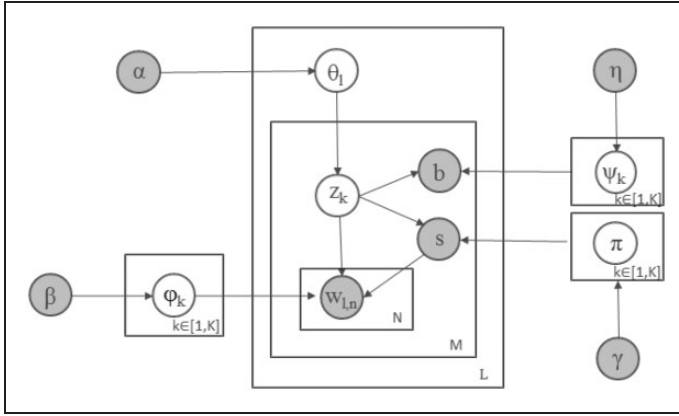


Figure 1. Behavior-sentiment topic mixture model (BSTM).

a distribution of a mixture of several topics. Then, a topic is randomly sampled from the distribution and this topic follows a Bernoulli distribution of sentiment. Finally, each word of one sentence is randomly sampled from the chosen pair of topic-sentiment. Thus, more fine-grained topics with the distribution of sentiments and behaviors can be directly identified. In other words, our model is able to not only uncover what learners say but also present how they say as well. Figure 1 illustrates the graphical representations of BSTM.

To facilitate the formal description in Figure 1, the solid circle node and the hollow circle node represents the known observed variables and the unknown hidden variables, respectively. The arrows between different nodes reveal closer dependences. For instance, a topic variable z_k is related to the behavior variable b and the sentiment variable s . The set of notations of BSTM are listed in Table 1.

Here, BSTM model will be described in detail. First, we assume that a total number of L learners participate in interacting with course reviews. Each review is composed of several sentences and each sentence corresponds to only one topic. Let r be subreviews of m -th sentence, then each review can be denoted as $R = \{r_1, r_2, \dots, r_m\} (1 \leq m \leq M)$. Therefore, the topic distribution of learners $n = 1, \dots, N_{lm}$ is represented as $l = \{z_1, z_2, \dots, z_k\}$. Second, we make an assumption that there are a total number of K topics and each topic has a probability distribution over words (ϕ_{kw}) denoted as $z_k = \{w_1, w_2, \dots, w_n\} (1 \leq n \leq V)$. In addition, each topic is closely coupled with behavior b and sentiment s ; each topic draws a multinomial distribution over behavior ψ_{kb} and draws a Bernoulli distribution over sentiment π_{ks} . Thus, the probability distributions over words for various pairs of topics-sentiments as well as topics-behaviors can be collectively represented. When a word w is sampled, we also assume that each word either comes from topic-specific words or sentiment words. At last, we assume

Table 1. Descriptions of Notations of BSTM.

Notations	Description	Notations	Description
L	The number of learners	α	Dirichlet prior of learner-topic
M_l	The number of sentences by learner l	β	Dirichlet prior of topic-word
N_{lm}	The number of words in l 's n -th sentence	η	Dirichlet prior of topic-behavior
K	The number of topics	γ	Dirichlet prior of topic-sentiment
V	The vocabulary size	θ	learner-topic distribution
z	A topic label	φ	topic-word distribution
b	A behavior includes $\{PO, LI, RE, SC\}$	ψ	topic-behavior distribution
s	A sentiment label	π	topic-sentiment distribution
w	A word label		

that θ_{lk} , φ_{kw} , ψ_{kb} , and π_{ks} have Dirichlet priors α , β , η , and γ separately. Next, we will show the elaborate generative process of BSTM model in Figure 2.

Parameter Estimation

In BSTM model, w is an observed known variable; Dirichlet prior parameters α , β , γ , and η are generally assigned based on our previous researches. The rest latent variables θ , φ , ψ , and π need to be estimated by the above known variables. The Gibbs Sample method proposed by Griffiths and coworkers (Steyvers & Griffiths, 2007) has shown to be effective to work out the appropriate parameter estimation problem. A joint conditional probability distribution incorporating behavior and sentiment information is given by

$$\begin{aligned}
 p(z_i = k | z_{-i}, w, b, s) \propto & \frac{C_{lk}^{LK} + \alpha}{\sum_{k=1}^K C_{lk}^{LK} + K \cdot \alpha} \frac{C_{kw}^{KW} + \beta}{\sum_{w=1}^V C_{kw}^{KW} + V \cdot \beta} \frac{C_{mkb}^{MKB} + \eta}{\sum_{b=1}^B C_{mkb}^{MKB} + B \cdot \eta} \\
 & \times \frac{C_{mkj}^{MKS} + \gamma}{\sum_{s=1}^S C_{mkj}^{MKS} + S \cdot \gamma}
 \end{aligned} \tag{1}$$

where z_i represents the assigned topic k on one sentence m_i and z_{-i} represents the assigned topic k on other sentences other than the current sentence m_i . C_{lk}^{LK} denotes the total number of sentences associated with topic k for one single learner l . C_{kw}^{KW} denotes the total number of words associated with topic k . C_{mkb}^{MKB} denotes the total number of behaviors that co-occur with topic k in

Step 1: Initialize model and the related experimental parameters K, α, β, η and γ

Step 2: For each topic $z_k, k=1, \dots, K$

Draw multiple words from topics: $\varphi_k \sim Dir(\beta)$

Draw multiple behaviors from topics: $\psi_{kb} \sim Dir(\eta)$

Draw multiple sentiments from topics: $\pi_{ks} \sim Dir(\gamma)$

Step 3: For every learner $l=1,2,\dots,L$

Draw multiple topics: $\theta_l \sim Dir(\alpha)$

— For l 's the m -th sentence, $m=1,\dots,M$

Draw a topic $z_{lm}, z_{lm} \sim Multi(\theta_{lk})$

Draw a dominant behavior $b_{lm}, b_{lm} \sim \psi_{z_{lm}}$

Draw a positive or negative sentiment $s_{lm}, s_{lm} \sim \psi_{z_{lm}}$

— For each word $n, n=1,\dots,N_{lm}$

Draw a word $w_{lmn}, w_{lmn} \sim Multi(\varphi_{z_{lm}})$

Step 4: Conducting Gibbs Sampling method and updating the related model parameters

Step 5: Repeating until the optimization of the model

Figure 2. The generative process for all learners' reviews in BSTM.

sentence m . C_{mkj}^{MKS} denotes the total number of sentiments that co-occur with topic k in sentence m . The related hidden parameters θ, φ, ψ , and π are updated. Third, repeating the previous step until the model is end and reaches a convergence state. As a result, the approximate probability distribution of topic k for one learner l is defined by

$$\theta_{lk} = \frac{C_{lk}^{LK} + \alpha}{\sum_{k=1}^K C_{lk}^{LK} + K \cdot \alpha} \quad (2)$$

The approximate probability distribution of word w for topic k is

$$\varphi_{kw} = \frac{C_{kw}^{KW} + \beta}{\sum_{w=1}^V C_{kw}^{KW} + V \cdot \beta} \quad (3)$$

The approximate probability distribution of behavior b for topic k in sentence m is

$$\psi_{mkb} = \frac{C_{mkb}^{MKB} + \eta}{\sum_{b=1}^B C_{mkb}^{MKB} + B \cdot \eta} \quad (4)$$

The approximate probability distribution of sentiment j for topic k sentence m is

$$\eta_{mkj} = \frac{C_{mkj}^{MKS} + \gamma}{\sum_{s=1}^S C_{mkj}^{MKS} + S \cdot \gamma} \quad (5)$$

In this way, the variables learner-topic θ , topic-word φ , topic-behavior ψ , and topic-sentiment π are calculated, respectively, which enables us to observe the detailed evaluation information related to courses.

Method

This section describes the basic information of the subjects in this study, as well as the process of data collection and data analysis.

Participants

From January 2014 to June 2016, the dataset was retrieved from a MOOC platform, namely *Guokr Web* (mooc.guokr.com), which is one of the most popular MOOCs learning platforms in China. A total number of 6,556 unique learners enrolled in the top 50 rated courses (e.g., language, literature, finance, science, engineering, and management) in the platform. Their age is generally between 12 and 50, and most of them come from developed provinces or cities in China. In addition, the average number of participated courses for them is more than one. Besides, these learners posted at least one comment during or after the course. Note that this platform allows the users and organizations to analyze the online MOOC data just for the aim of research and management without learners' identification. In addition, our study does not use and reserve learners' personal information, such as names.

Data Collection

In this study, the method of web crawler was adopted to obtain learners’ comments (including publishing time, text contents, behavior label, etc.) from course reviews module in *Guokr Web*. There are a total number of 12,524 reviews from 50 courses in the platform. We note that almost all of these reviews are written in Chinese. Table 2 shows some basic statistics from these reviews.

Besides, four types of behaviors are defined. The coding scheme (as in Table 3) of online interactive behavior is adopted to ensure the convenience and consistency for the further analysis. In general, each review is assigned one label. A particular case is that when a participant posted course reviews, he or she showed his or her earned course certificate. In such case, this review is assigned two labels *#PO* and *#SC* together. Note that the sampling cases are generated after each review has been split into several sentences. Then, each sentence is turned into a set of tokens consisting of nonredundant words by utilizing the Chinese word segmentation system of Chinese academy of sciences ICTCLAS (Zhang, Liu, Cheng, Zhang, & Yu, 2003). Finally, stop words (e.g., “和/and,” “他/he,” and “这样/so”), low frequency words (based on our experiences, if the occurrence number of a word is less than 5), noise words (e.g., starting with http, www), and improper characters (e.g., emoticons) are all filtered out.

Table 2. The Properties of the Data Sets.

Number of reviews	Number of sentences	Number of words	Avg. number of sentences per review	Avg. number of words per review	Number of positive words	Number of negative words
12,524	50,096	770,977	4.00	61.56	74,125	18,501

Table 3. The Coding Scheme of Learners’ Interactive Behaviors.

Behaviors	Codes	Description
Post	PO	Post an original review
Like	LI	Give someone’s review a thumbs up
Reply	RE	Reply to someone’s review
Show earned certificate	SC	Decide whether to show earned certificate when someone posts a review

Data Analysis

To detect sentimental and behavioral tendencies of learners' concerned topics that were hidden in the 12,524 course reviews, an improved unified topic model BSTM was developed and a two-stage content analysis was conducted. In the first phase, we used a quantitative analysis method to evaluate the effectiveness of the model for the first research question (see Research Questions section). In the second phase, the model generation technique was adopted to detect the potential topics learners were willing to evaluate in course reviews, topics-sentiments distributions, as well as topics-behaviors distributions for the second to fourth research questions (see Research Questions section). In addition, we also adopted descriptive statistics to compute the average probability of typical topics, between sentimental value of topics and behavioral probability correlation coefficient, and so on. The significant alpha was set to .05.

Results

This section presents not only the evaluation results for BSTM but also the results of the frequent topics learners discuss most, topics-oriented sentimental tendencies as well as how learners interact with these topics.

Evaluation Measures for BSTM

In this part, our proposed topic model BSTM is evaluated from two aspects. First, to validate the effectiveness of BSTM, a comparison with other two standard models (LDA and T-LDA) was made. Second, to ensure the best performance of BSTM, the number of topics was determined.

Comparison of model effectiveness. It is acknowledged that LDA is expected to obtain the distributions of topics over words which are easy to be distinguished from each other. In this case, BSTM aims to capture different topics, which are closely related to specific dominate behaviors and sentiments. To validate the effectiveness of BSTM integrating with the behavior and sentiment features, an empirical study on topical extraction is performed. The Kullback–Leibler (KL; Mei et al., 2007) divergence is generally used to calculate the similarity between two pairs of probability distributions. The smaller the KL score, the more distinct the different topics. In our study, we use it to measure the similarity of distributions of topics from T-LDA (Zhao et al., 2011), BSTM, and LDA over behaviors and sentiments. Here consider the computation of distribution of topic-behavior as an example. The equation of KL is as follows:

$$KL(Q_{k,b}||P_{k,b}) = \sum_i^{|B|} Q_{k,b}(B_i) \log Q_{k,b}(B_i)/P_{k,b}(B_i) \quad (6)$$

Since the KL divergence is asymmetrical, in this work, the average KL divergence is adjusted between them for a symmetric measurement by the following formula:

$$KL(Q_{k,b}||P_{k,b}) = (KL(Q_{k,b}||P_{k,b}) + KL(P_{k,b}||Q_{k,b}))/2 \quad (7)$$

In BSTM, $P_{k,b}$ and $Q_{k,b}$ are equal to $\varphi_{k,b}$. Note that T-LDA makes an assumption that each comment containing a series of words is only associated with one topic, and LDA constrains that each word within one review is assigned with one topic. In this case, the computation of co-occurrence of behaviors within one topic is different for T-LDA, BSTM, and LDA.

The purpose of this study focuses on the combination of behavior and sentiment features with topic analysis. Compared with T-LDA and LDA, BSTM obviously has a lower KL divergence score as presented in Figure 3(a) and 3(b), which indicates that the distributions of behaviors toward different topics are more effectively distinguishable and separable from each other, as in the case with the distributions of topics-sentiments. Thus, embedding behavior and sentiment features into topic modeling are verified for our experiments.

Determining number of topics. To get a better model, we utilize a parameter turning way to repeatedly conduct tests. The Dirichlet prior parameters α , β , γ , and η are empirically assigned with 0.1, 0.01, 0.1, and 0.1, respectively. Besides, the number of topics is known as a critical factor. Our study bases on the computation of Perplexity, Similarity, Entropy (PSE) parameter, which is a comprehensive quantitative evaluation index including perplexity (Fu et al., 2015),

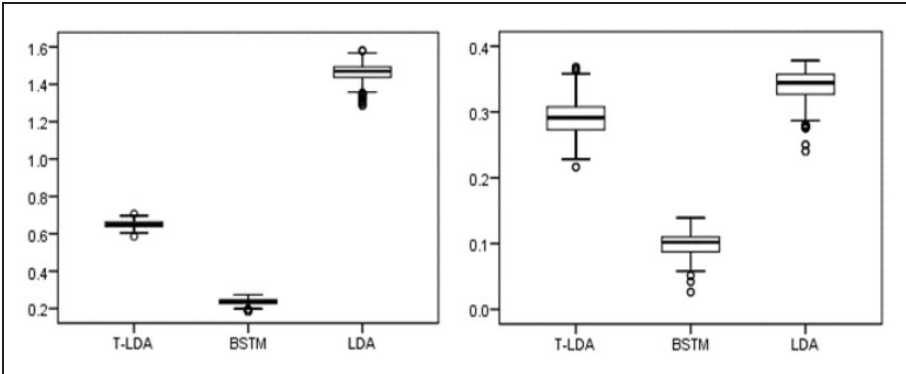


Figure 3. Comparison for T-LDA, BSTM and LDA, with levels of KL. (a) The dissimilarity of topics corresponding to behavior distribution and (b) The dissimilarity of topics corresponding to sentiment distribution.

similarity (Mei et al., 2007), and entropy (Qiu et al., 2013), to determine the suitable number of topics, thus effectively improving the performance of BSTM. The detailed formal description about the related parameters is presented as follows. PSE is a rational comprehensive assessing index, with a smaller value being preferred and is given by

$$PSE = Perplexity \cdot Similarity \cdot Entropy \tag{8}$$

The experiment is performed gradually to determine the number of topics. First, for each test, the total number of iterations is set to 500 and the initial number of topics is set to 20. Then, 20 topics are added in each experiment until the value of PSE is smaller and more stable.

As shown in Figure 4, with the increase of numbers of iterations, the fitting rate of model becomes fast. All of the curves reach a steady state after 400 times sampling. It can also be clearly observed that when the number of topics is equal to 100, the values of PSE are no less than the curves with different topic numbers. This means that there is no slight decline when the number of topics is set to 80 and the performance of our model is difficult to be improved. Hence, 80 topics are chosen for further investigation in our study.

Discovery of Typical Topics From Online Learners’ Course Reviews

Table 4 illustrates the top five topics (all above the average probability 0.0125) learners discuss most and the top 10 words related to these topics with the

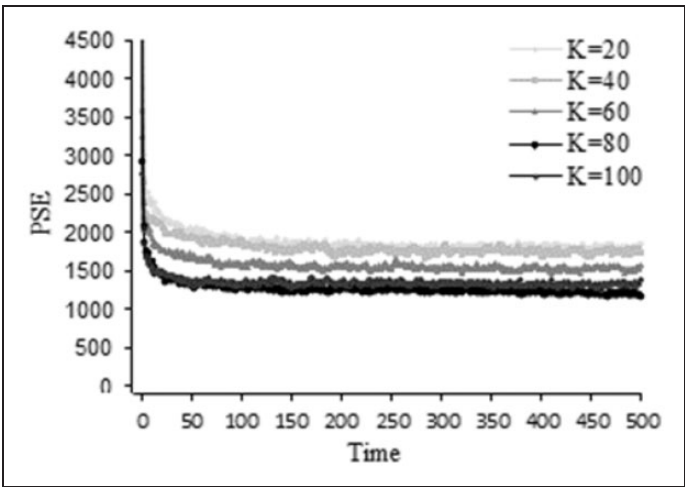


Figure 4. PSE values for varying number of topics (K).

highest probability. Note that the words in this study are regarded as seed words without labeled sentiments.

As shown in Table 4, Topic 76 with the probability ratio of occurrence in all reviews maintains at 7.2%, far higher than the average rate 1.25%. Evidently, Topic 76 is one focused topic among online learners with the highest probability. The topic is mainly about *course content*. Through observing its detailed fine-grained semantic information, we find that most of learners in MOOCs are concerned about the quality and organization of course content. Topic 61 is likely related to a *history course* named “Records of the Grand Historian,” which is one of the China’s famous ancient history. Topic 27 is mainly about

Table 4. Top Five Topics and Their Probability of Occurrence of Typical Words.

Topics	Word	Prob.	Word	Prob.
Topic 27 (0.028)	课程/course	0.027	应用/application	0.018
	统计/statistics	0.016	理论/theory	0.014
	数学/math	0.012	算法/algorithm	0.011
	机器/machine learning	0.011	知识/knowledge	0.011
	方法/method	0.010	概念/concept	0.009
Topic 35 (0.020)	时间/time	0.019	考试/exam	0.019
	证书/certificate	0.016	课程/course	0.016
	结束/end	0.014	习题/quiz	0.014
	完成/complete	0.012	材料/material	0.012
	比较/compare	0.010	阅读/read	0.010
Topic 61 (0.047)	中国/China	0.019	台大/NTU	0.014
	吕老师/teacher	0.014	史记/Shiji	0.012
	学> /study	0.010	付出/pay	0.009
	传统/tradition	0.009	历史/history	0.008
	吸引/attract	0.007	文化/culture	0.007
Topic 76 (0.072)	课程/course	0.042	学习> /study	0.020
	内容/content	0.014	mooc	0.010
	知识/knowledge	0.010	设计/design	0.008
	时间/time	0.007	视频/video	0.006
	基础/foundation	0.006	简单/easy	0.006
Topic 78 (0.013)	字幕/caption	0.046	下载/download	0.037
	视频/video	0.037	开课/class begin	0.016
	在线/online	0.016	听课/listen	0.016
	英文/English	0.016	mooc	0.013
	翻墙/proxy	0.011	coursera	0.011

the courses related to *machine learning* or *mathematical statistics*. It is notable that unlike the above topics, Topic 35 mainly refers to *course logistics*, including time of exam, accomplishment of exercises, and acquirement of certificates; Topic 78 mainly involves *video production*, including language of caption and link of download, and so forth.

Detection of Sentimental Tendencies Toward Learners’ Concerned Topics

The focused topics among learners have been detected and inferred. Learners’ attitudes and opinions toward these topics are equally essential for instructors and administrators. Taking the above five typical topics as an example, their corresponding top eight sentiment words, which have highest probability are listed in Table 5. Note that each topic covers not only the related positive words but also the negative words. Here, we just list the related sentimental words in

Table 5. Top Five Topics and Their Corresponding Sentiment Words.

Typical topics	Sentiment distribution	Sentiment words	Sentiment words
Topic 27 (+)	(0.553, 0.447)	有趣/interesting 收获/gain 丰富/abundant 适合/suitable	易懂/ easy understanding 优点/advantage 深入/deep 扎实/ solid
Topic 61 (+)	(0.896, 0.104)	期待/expect 精彩/splendid 赞/thump up 富有/rich	吸引/attract 推荐/recommend 不错/nice 精品/high quality
Topic 76 (+)	(0.789, 0.211)	用心/dedicated 认真/earnest 棒/good 生动/vivid	有趣/amusing 喜欢/love 有意思/fun 值得/worthy
Topic 35 (–)	(0.215, 0.785)	无聊/boring 放弃/give up 难/difficult 不足/insufficient	吃力/strenuous 拖沓/dilatory 不值得/not worth 烦躁/irritable
Topic 78 (–)	(0.432, 0.568)	严重/serious 糟/bad 延迟/postpone 差/poor	不喜欢/dislike 脱离/break away 不懂/unclear 困难/difficult

terms of each topical sentiment polarity; (+) represents this topic is positive, and (−) represents this topic is negative. Sentiment distribution for each topic in parentheses is positive rate and negative rate, from left to right.

To detect the sentimental tendencies of learners' concerned topics, the experimental results involved in five prominent topics are presented in Table 5. We find that these most frequent topics discussed are assigned with either positive sentiment or negative sentiment, which to some extent could reflect learners' general sentimental states about them. Taking Topic 76 with highest probability as an example, we find that learners make a positive evaluation toward course content which is interesting, vivid, good, and worthy of learning. In particular, sentiment distribution for each topic is obviously different and needs to be further discussed.

The positive probability 0.789 of Topic 76 is much greater than its negative probability 0.211, which is different to the sentimental distribution of Topic 27. While Topic 27 is assigned to be positive as a whole, its positive sentiment probability 0.553 is slightly higher than its negative sentiment probability 0.447. This implies that learners hold a relatively neutral attitude toward Topic 27. In this case, a positive topic is more likely to be transformed to a negative one because of sentiment propagation (Zhao et al., 2014). Hence, instructors need to attach importance to this kind of “unsteady” topic.

Absolutely, understanding the evaluation contents of negative topics is also essential (McGivern & Noret, 2011), which is beneficial for instructors to optimize design of course and improve learners' participation, especially when they encounter learning difficulties or other objective problems in accessing to learning resources. For example, the negative Topic 35 about course logistics is boring, difficult, and insufficient, with a higher negative sentiment probability 0.785; the negative Topic 78 about videos production is serious, postponed, and unclear, with a lower negative sentiment probability 0.568. Thus, we have reasons to believe that the existing problems within Topic 35 should be indeed given precedence.

Learners' Concerned Topics Separated by Dominant Behavioral Tendencies

Like topic-sentiment distribution, behavior is also thought of as a special property within topics in our proposed model. The close examination should be conducted to investigate how learners interact with topics while the focused topics discussed most among learners and their associated sentimental tendencies toward these topics have been identified.

Table 6 shows the number of four dominate behaviors which are coded as *PO*, *LI*, *RE*, and *SC*. Obviously, for learners, posting is the most frequent behavior way to express their opinions and attitudes. Table 7 shows the distribution of five typical topics on the behavior dimensions of *PO*, *LI*, *RE*, and *SC* together with their sentiment labels.

Table 6. Statistics About Dominant Behaviors Among All Learners.

Behavior type	PO	LI	RE	SC
Number of total behaviors	6,532	3,783	2,209	3,588

Table 7. Different Behavioral Tendencies for Typical Topics.

Typical topics	Topic label	Sentiment label	PO	LI	RE	SC
Topic 76 (+)	Course content	Positive	0.324	0.324	0.078	0.274
Topic 61 (+)	History course	Positive	0.493	0.082	0.001	0.424
Topic 27 (+)	Machine learning	Positive	0.323	0.377	0.024	0.277
Topic 35 (−)	Exam, homework	Negative	0.473	0.046	0.016	0.465
Topic 78 (−)	Video production	Negative	0.067	0.110	0.778	0.045

As shown in Table 7, there exists a certain difference in the four behavioral distributions associated with different topics. Topic 76 is mainly dominated by three behaviors of PO, LI, and SC; Topic 61 and Topic 35 is mainly dominated by PO and SC behaviors; Topic 27 is mainly dominated by PO and LI. This suggests that learners tend to post or like reviews to express their personal opinions and attitudes toward most focused topics, whereas they rarely choose to reply to others’ original posts. Note that different from other typical topics, Topic 78 about video production is mainly dominated by just one behavior RE, which indicates that this topic attracts more learners to reply to others’ reviews. Moreover, for each typical topic, most of learners are willing to show earned course certificates when they post reviews. For them, this behavior may be regarded as same as getting the badges of honor which brings incentives for their engagement (Anderson et al., 2014). Specifically, an interesting phenomenon is that Topic 61 and Topic 35 almost share the same distribution of behavior but are related to the different topics and different sentiment polarities. In addition, we performed a correlation analysis between 4 dominant behaviors and all 80 sentiment topics (56 positive topics and 24 negative topics). As a result, only RE presents a significant negative correlation to topics with a higher positive probability ($r = -0.269$, $p < .05$). That is, learners tend to express more negative topics when replying to others’ posts. Intuitively, LI should be related to the topics with a higher positive probability, yet it seems that there does not exist any correlations between them. This indicates that learners may give others’ reviews a thumbs up about both positive contents and negative contents in an irregular way.

Discussion

The results of our study find that several prominent topics which are frequently discussed among learners have been labeled with different sentimental polarities (either positive or negative), and these topics have different sentimental distributions. A particular case is that one topic is assigned to the positive sentiment probability and negative sentiment probability with similar values. For example, the sentiment distribution of Topic 27 about machine learning is denoted as (0.554, 0.445). Although it is labeled as a positive topic, this topic is prone to be a negative one owing to the characteristics of sentiment propagation (Zhao et al., 2014). In addition, it is imperative for instructors to focus on negative topics like 35 about course logistics with higher negative sentiment probability preferentially. These negative topics spread fast (Mcgivern & Noret, 2011) and considerably go against the development of a harmonious and steady online learning community. And they may be responsible for online learners' low participation rate as described by Wen et al. (2014). Thus, in this case, to avoid the negative effects of topics on learners, instructors should take mediate measures in time to regulate and guide the interactive process of discourse context. As Shattuck and Anderson (2013) suggested, lack of clear and practical guidelines may lead to learners' fear and resistance.

A correlation analysis was conducted between 4 dominant behaviors (PO, LI, RE, SC) and all 80 topics with different sentimental polarities. The result shows that there is only one behavior RE closely related to the sentiment probability regarding topics and the RE behavior has a significant negative relevance with the positive sentiment probability. This implies that learners are more inclined to express negative sentiment toward concerned topics by replying to others' review. This case is very interesting and may be related to some specific topics. In addition, there is no direct correlation between the LI behavior and the positive sentiment probability toward topics. That is, learners tend to give others' reviews a thumbs up about both positive and negative aspects. One of the reasons may be that the LI behavior is easily followed among learners and is exclusively concerned with the text contents of the reviews (Sherman, Payton, Hernandez, Greenfield, & Dapretto, 2016).

Although the topic about course content discussed among learners has a higher probability, which are mainly dominated by the PO and LI behavior, only a small proportion of learners tend to reply to others' reviews about this topic. Unlike this topic, while the topic about video production discussed among learners has a lower probability, it actually attracts more learners to take part in responding with a higher behavior probability. With respect to the above two types of topics, the latter type can be thought to be a more enriched interactivity than the former one. Kiemer, Groschner, Pehmer, and Seidel (2015) remarked that courses with more interactive dialogues could significantly improve learners' internal motivation and behavioral participation. Therefore, it is necessary

for instructors to pay more attention to this type of topic with more interactive discussions. Moreover, according to the results of learners' focused topics with diverse behavioral patterns, instructors are able to provide more adaptive and accurate recommendations for different learners, such as peers' reviews.

Previous studies have pointed out that reading through and searching for reviews was extremely time consuming (Jo & Oh, 2011; Wang, 2011; Xu et al., 2015) and thus difficult to operate repeatedly. Hence, compared with the traditional method of postcourse interviews and questionnaires, the method of unsupervised topic modeling BSTM may be a useful and practical alternative for course evaluation, and still needs to be further investigated.

Conclusion, Limitation, and Future Work

This study proposes a unified topic model BSTM by embedding the behavior and sentiment features into the computational analysis of large amounts of learners-generated course reviews in a MOOC. Compared with the benchmark model, BSTM has shown to be dominant. Furthermore, it can be generalizable to other types of learning contexts and can also be utilized in many potential applications such as opinion tracking, course-oriented review summarization, and behavior-driven personalized recommendation. Empirical experiments on real MOOC reviews data show that this approach is effective for automatically exploring the learners' concerned topics, topics-oriented sentimental tendencies, as well as how learners interact with these topics.

Thus, understanding the hidden feedbacks from course reviews can help instructors and administrators conduct timely teaching intervention when learners encounter dissatisfactions (e.g., lecture is boring, homework is difficult) and improve platform construction.

There are, however, still some limitations, which may lead to many directions for future work. First, the specific conditions in a Chinese MOOC by using BSTM may limit the ability to explain the findings of experimental analysis. To further generalize this model in other study contexts, a set of specific rules of phrase segmentation and sentence-oriented sentiment annotation depending on the context should be reconstructed. Second, since our model belongs to an unsupervised method in the field of natural language processing, the computational results should still be examined through logical reasoning, which may cause slight deviations on the obtained results. In such case, human efforts may be required to label textual contents afterwards. Third, because the number of learners' concerned topics is relatively large, it is, to some extent, difficult to elaborate all topics in conjunction with some detailed information like these topics' sentimental and behavioral tendencies. Therefore, how to exhibit these hidden valuable information in a visualized way, in which ought to be in accordance with learners' cognitive preferences (Jyothi, McAvinia, & Keating, 2012; Simsek et al., 2014), is worthy of more attentions. Finally, there seem to be

some correlations among topics within the courses. Thus, it is necessary to determine these special relationships and group them for better understanding the potential semantic structure space of course content.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Research funds from the China Mobile Research Foundation of the Ministry of Education (Grand No. MCM20160401) the Research Funds from National Natural Science Foundation of China (Grant No. 61702207, L1724007).

References

- Adamopoulos, P. (2013). What makes a great MOOC? An interdisciplinary analysis of online course student retention. *Proceedings of the 34th international conference on information systems*. Milan, Italy: ACM. Retrieved from http://idea.btoresarch.com/images/freepaper/63_What%20makes%20a%20great%20MOOC_%20An%20interdisciplinary%20analysis%20of%20student.pdf
- Altrabsheh, N., Cocea, M., & Fallahkhair, S. (2015, June). Predicting students' emotions using machine learning techniques. *International Conference on Artificial Intelligence in Education* (pp. 537–540). Switzerland: Springer International Publishing. doi:10.1007/978-3-319-19773-9_56
- Anderson, A., Huttenlocher, D., Kleinberg, J., & Leskovec, J. (2012). Discovering value from community activity on focused question answering sites: a case study of stack overflow. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 850–858). Beijing, China: ACM. Retrieved from <https://doi.org/10.1145/2339530.2339665>
- Anderson, A., Huttenlocher, D., Kleinberg, J., & Leskovec, J. (2014). Engaging with massive online courses. In *Proceedings of the 23rd international conference on World wide web* (pp. 687–698). Seoul, South Korea: ACM. Retrieved from <https://doi.org/10.1145/2566486.2568042>
- Bahreini, K., Nadolski, R., & Westera, W. (2016). Towards real-time speech emotion recognition for affective e-learning. *Education and Information Technologies*, 21(5), 1367–1386. doi:10.1007/s10639-015-9388-2
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3(1), 993–1022. Retrieved from <http://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf>
- Brinton, C. G., Chiang, M., Jain, S., Lam, H., Liu, Z., & Wong, F. M. F. (2014). Learning about social learning in MOOCs: From statistical analysis to generative model. *IEEE Transactions on Learning Technologies*, 7(4), 346–359. Retrieved from <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6851916>

- Ehrenreich, S. E., Underwood, M. K., & Ackerman, R. A. (2014). Adolescents' text message communication and growth in antisocial behavior across the first year of high school. *Journal of Abnormal Child Psychology*, 42(2), 251–264. doi:10.1007/s10802-013-9783-3
- El-Halees, A. (2011). Mining opinions in user-generated contents to improve course evaluation. *Software Engineering and Computer Systems*, 180, 107–115. doi:10.1007/978-3-642-22191-0_9
- Fu, X., Yang, K., Huang, J. Z., & Cui, L. (2015). Dynamic non-parametric joint sentiment topic mixture model. *Knowledge-Based Systems*, 82, 102–114. doi:10.1016/j.knosys.2015.02.021
- Goffman, E. (1978). *The presentation of self in everyday life* (p. 56). Harmondsworth, England: Penguin Retrieved from http://wps.pearsoncustom.com/wps/media/objects/6714/6875653/readings/MSL_Goffman_Presentation.pdf
- Grafsgaard, J., Wiggins, J. B., Boyer, K. E., Wiebe, E. N., & Lester, J. (2013, July). Automatically recognizing facial expression: Predicting engagement and frustration. *Paper presented at the educational data mining 2013, ACM*. Retrieved from <http://www.educationaldatamining.org/conferences/index.php/EDM/2013/paper>
- Grimmer, J. (2010). A Bayesian hierarchical topic model for political texts: Measuring expressed agendas in Senate press releases. *Political Analysis*, 18, 1–35. <http://doi.org/10.1093/pan/mpp034>
- Guo, P. J., Kim, J., & Rubin, R. (2014, March). How video production affects student engagement: An empirical study of MOOC videos. *Proceedings of the first ACM conference on Learning@ scale conference* (pp. 41–50). Atlanta, Georgia: ACM. doi:10.1145/2556325.2566239
- Halawa, S., Greene, D., & Mitchell, J. (2014). Dropout prediction in MOOCs using learner activity features. *Experiences and Best Practices in and Around MOOCs*, 7, 1–10. Retrieved from https://www.openeducationeuropa.eu/sites/default/files/legacy_files/asset/In_depth_37_1.pdf
- Jo, Y., & Oh, A. H. (2011, February). Aspect and sentiment unification model for online review analysis. *Proceedings of the fourth ACM international conference on Web search and data mining* (pp. 815–824). Hong Kong, China: ACM. doi:10.1145/1935826.1935932
- Jyothi, S., McAvinia, C., & Keating, J. (2012). A visualisation tool to aid exploration of students' interactions in asynchronous online communication. *Computers & Education*, 58(1), 30–42. doi:10.1016/j.compedu.2011.08.026
- Kiemer, K., Gröschner, A., Pehmer, A. K., & Seidel, T. (2015). Effects of a classroom discourse intervention on teachers' practice and students' motivation to learn mathematics and science. *Learning and Instruction*, 35, 94–103. doi:10.1016/j.learninstruc.2014.10.003
- Kizilcec, R. F., Piech, C., & Schneider, E. (2013, April). Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. *Proceedings of the third international conference on learning analytics and knowledge* (pp. 170–179). Leuven, Belgium: ACM. doi:10.1145/2460296.2460330
- Klein, R., Kyrilov, A., & Tokman, M. (2011, June). Automated assessment of short free-text responses in computer science using latent semantic analysis. *Proceedings of the 16th annual joint conference on Innovation and technology in computer science education* (pp. 158–162). Darmstadt, Germany: ACM.

- Klisc, C., McGill, T., & Hobbs, V. (2017). Use of a post-asynchronous online discussion assessment to enhance student critical thinking. *Australasian Journal of Educational Technology*, 33(5), 63–76. doi:10.14742/ajet.3030
- Kontogiannis, S., Valsamidis, S., Kazanidis, I., & Karakos, A. (2014, October). Course opinion mining methodology for knowledge discovery, based on web social media. *Proceedings of the 18th Panhellenic Conference on Informatics* (pp. 1–6). Athens, Greece: ACM. doi:10.1145/2645791.2645827
- Leong, C. K., Lee, Y. H., & Mak, W. K. (2012). Mining sentiments in SMS texts for teaching evaluation. *Expert Systems With Applications*, 39(3), 2584–2589. doi:10.1016/j.eswa.2011.08.113
- McGivern, P., & Noret, N. (2011). Online social networking and e-safety: Analysis of risk-taking behaviours and negative online experiences among adolescents. *British conference of undergraduate research 2011 special issue*. Retrieved from <http://www.warwick.ac.uk/go/reinventionjournal/issues/BCUR2011specialissue/mcgivernnoret>.
- Mei, Q., Ling, X., Wondra, M., Su, H., & Zhai, C. (2007, May). Topic sentiment mixture: Modeling facets and opinions in weblogs. *Proceedings of the 16th international conference on World Wide Web* (pp. 171–180). New York, NY: ACM. doi:10.1145/1242572.1242596
- O'regan, K. (2003). Emotion and e-learning. *Journal of Asynchronous Learning Networks*, 7(3), 78–92. Retrieved from http://www.adesignmedia.com/online/research/emotion_learnv7n3_oregan.pdf
- Qiu, M., Zhu, F., & Jiang, J. (2013, May). It is not just what we say, but how we say them: LDA-based behavior-topic model. *Proceedings of the 2013 SIAM International Conference on Data Mining* (pp. 794–802). Society for Industrial and Applied Mathematics. Texas, USA: SIAM. doi:11.9781/1.9781611972832.88
- Ramesh, A., Goldwasser, D., Huang, B., Daumé, H., III, & Getoor, L. (2013, December). *Modeling learner engagement in MOOCs using probabilistic soft logic*. Paper presented at the NIPS Workshop on Data Driven Education (pp. 62–67). Retrieved from <http://www.academia.edu/download/35876424/daume13engagementmooc.pdf>
- Ramesh, A., Goldwasser, D., Huang, B., Daumé, H. D., III, & Getoor, L. (2014). Understanding MOOC discussion forums using seeded LDA. *Proceedings of the Ninth Workshop on Innovative Use of NLP for Building Educational Applications* (pp. 28–33). Baltimore, MD: Association for Computational Linguistics. Retrieved from <http://www.aclweb.org/anthology/W/W14/W14-1804.pdf>
- Reich, J. (2015). Rebooting MOOC research. *Science*, 347(6217), 34–35. doi:10.1126/science.1261627
- Rodriguez, P., Ortigosa, A., & Carro, R. M. (2012, July). Extracting emotions from texts in e-learning environments. *2012 Sixth International Conference on Complex, Intelligent and Software Intensive Systems (CISIS)* (pp. 887–892). Palermo, Italy: IEEE. doi:10.1109/CISIS.2012.192
- Romero, C., López, M. I., Luna, J. M., & Ventura, S. (2013). Predicting students' final performance from participation in on-line discussion forums. *Computers & Education*, 68, 458–472. doi:10.1016/j.compedu.2013.06.009
- Shattuck, J., & Anderson, T. (2013). Using a design-based research study to identify principles for training instructors to teach online. *International Review of Research*

- in *Open and Distance Learning*, 14(5), 186–210. Retrieved from <http://www.irrodl.org/index.php/irrodl/article/view/1626>
- Sherman, L. E., Payton, A. A., Hernandez, L. M., Greenfield, P. M., & Dapretto, M. (2016). The power of the like in adolescence: Effects of peer influence on neural and behavioral responses to social media. *Psychological Science*, 27(7), 1027–1035. doi:10.1177/0956797616645673
- Siemens, G., & Baker, R. S. (2012, April). Learning analytics and educational data mining: Towards communication and collaboration. *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 252–254). British Columbia, Canada: ACM. doi:10.1145/2330601.2330661
- Simsek, D., Shum, S. B., De Liddo, A., Ferguson, R., & Sándor, Á. (2014, March). Visual analytics of academic writing. *Proceedings of the fourth international conference on learning analytics and knowledge* (pp. 265–266). Indianapolis, IN: ACM. doi:10.1145/2567574.2567577
- Steyvers, M., & Griffiths, T. (2007). Probabilistic topic models. *Handbook of Latent Semantic Analysis*, 427(7), 424–440. doi:10.1145/2133806.2133826
- Sylwester, R. (1994). How emotions affect learning. *Educational Leadership*, 52(2), 60–65.
- Taylor, C., Veeramachaneni, K., & O'Reilly, U. M. (2014). *Likely to stop? Predicting stopout in massive open online courses*, 1–25. Retrieved from <https://arxiv.org/pdf/1408.3382.pdf>
- Tobarra, L., Robles-Gómez, A., Ros, S., Hernández, R., & Caminero, A. C. (2014). Analyzing the students' behavior and relevant topics in virtual learning communities. *Computers in Human Behavior*, 31, 659–669. doi:10.1016/j.chb.2013.10.001
- Wan, X. (2011). Bilingual co-training for sentiment classification of Chinese product reviews. *Computational Linguistics*, 37(3), 587–616. doi:10.1162/COLI_a_00061
- Wen, M., & Rosé, C. P. (2014, November). Identifying latent study habits by mining learner behavior patterns in massive open online courses. *Proceedings of the 23rd ACM international conference on conference on information and knowledge management* (pp. 1983–1986). New York, NY: ACM. doi:10.1145/2661829.2662033
- Wen, M., Yang, D., & Rose, C. (2014, July). Sentiment Analysis in MOOC Discussion Forums: What does it tell us?. In *Proceedings of The 7th International Conference on Educational Data Mining*. (pp. 1–8). London, UK: ACM. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.728.1722&rep=rep1&type=pdf>
- Woolf, B., Bursleson, W., Arroyo, I., Dragon, T., Cooper, D., & Picard, R. (2009). Affect-aware tutors: Recognising and responding to student affect. *International Journal of Learning Technology*, 4(3–4), 129–164. doi:10.1504/IJLT.2009.028804
- Xu, H., Zhang, F., & Wang, W. (2015). Implicit feature identification in Chinese reviews using explicit topic mining model. *Knowledge-Based Systems*, 76, 166–175. doi:10.1016/j.knsys.2014.12.012
- Yu, W. B., & Luna, R. (2013, July). Exploring user feedback of a e-learning system: A text mining approach. *International conference on human interface and the management of information* (pp. 182–191). Berlin, Germany: Springer. doi:10.1007/978-3-642-39226-9_21
- Zhang, H. P., Liu, Q., Cheng, X. Q., Zhang, H., & Yu, H. K. (2003, July). Chinese lexical analysis using hierarchical hidden Markov model. *Proceedings of the second SIGHAN*

- workshop on Chinese language processing* (pp. 63–70). China: Association for Computational Linguistics. doi:10.3115/1119250.1119259
- Zhao, L., Wang, J., Huang, R., Cui, H., Qiu, X., & Wang, X. (2014). Sentiment contagion in complex networks. *Physica A: Statistical Mechanics and its Applications*, 394, 17–23. doi:10.1016/j.physa.2013.09.057
- Zhao, W. X., Jiang, J., Weng, J., He, J., Lim, E. P., Yan, H., & Li, X. (2011, April). Comparing twitter and traditional media using topic models. *European conference on information retrieval* (pp. 338–349). Dublin, Ireland: Springer Berlin Heidelberg. doi:10.1007/978-3-642-20161-5_34
- Zhao, Z., Cheng, Z., Hong, L., & Chi, E. H. (2015, May). Improving user topic interest profiles by behavior factorization. *Proceedings of the 24th international conference on World Wide Web* (pp. 1406–1416). Florence, Italy: ACM. doi:10.1145/2736277.2741656

Author Biographies

Sannyuya Liu is a professor and an associate director in National Engineering Research Center for E-learning (NERCEL), Central China Normal University (CCNU). His research has been funded by National Social Science Foundation. He has led many national researches and technological projects and made contributions in technology development and higher education sectors. His research interests include learning analytics and educational data mining and sentiment recognition in online learning environment.
Email: lsy5918@mail.ccnu.edu.cn

Xian Peng received the master's degree in educational technology from the Central China Normal University in 2012 and 2014 in China. He is currently working toward the PhD degree in education informational technology in National Engineering Research Center for E-learning (NERCEL), Central China Normal University (CCNU), China. His research interests are topic modeling and learning behavior analysis. He is the corresponding author:
Email: px87374006@126.com

Hercy N. H. Cheng was an adjunct assistant professor and a postdoctoral scholar in the Institute of Network Learning Technology at National Central University, Taiwan. He has published 10 SSCI papers on digital learning, some are indexed in Computers and Education Journal. His current research interests include one-to-one learning environments and game-based learning, in particular, challenge design in a one-to-one classroom.
Email: hercycheng.tw@gmail.com

Zhi Liu is a research associate in NERCEL, CCNU, China, since 2015. He is dedicated to conduct the research on learning behavior analysis and educational data mining. His research has been funded by Social Sciences Foundation on the analysis of online learning behavioral patterns and learner modeling.

Email: liuzhi8673@gmail.com

Jianwen Sun received the BS and PhD degrees from CCNU in 2005 and 2011, respectively. He is currently an associate professor in NERCEL, CCNU. Simultaneously, he is a director at the information department in the First Affiliated Middle School of CCNU. His research interests lie in the areas of Educational Data Mining and Intelligent Tutor System.

Email: sunjw@mail.ccnu.edu

Chongyang Yang is currently working toward the master's degree in education informational technology in National Engineering Research Center for E-learning (NERCEL), Central China Normal University (CCNU), China. Her research interests are learning analytics and sentiment detection.

Email: chongyangyang@mails.ccnu.edu.cn