

RESEARCH ARTICLE

Open Access



Topic tracking model for analyzing student-generated posts in SPOC discussion forums

Xian Peng^{1*}, Chengyang Han², Fan Ouyang¹ and Zhi Liu³

* Correspondence: percy0226@zju.edu.cn

¹College of Education, Zhejiang University, Hangzhou 310028, Zhejiang, China
Full list of author information is available at the end of the article

Abstract

Due to an overwhelming amount of student-generated forum posts in small private online courses (SPOCs), students and instructors find it time-consuming and challenging to effectively navigate and track valuable information, such as the evolution of topics, emotional and behavioral changes in relation to topics. For solving this problem, this study analyzed plenty of discussion posts using an improved dynamic topic model, Time Information-Emotion Behavior Model (TI-EBTM). Time, emotion, and behavior characteristics were incorporated into the topic modeling process, which allowed for an overview of automatic tracking and understanding of temporal topic changes in SPOC discussion forums. The experiment on data from 30 SPOC courses showed that TI-EBTM outperformed other dynamic topic models and was effective in extracting prominent topics over time. Furthermore, we conducted an in-depth temporal topic analysis to investigate the utility of TI-EBTM in a case study. The results of the case study demonstrated that our methodology and analysis shed light on students' temporal focuses (i.e., the changes of topic intensity and topic content) and reflected the evolution of topics' emotional and behavioral tendencies. For example, students tended to express more negative emotions toward the topic about the method of data query by initiating the conversation at the end of the semester. The analytical results can provide instructors with valuable insights into the development of course forums and enable them to fine-tune course forums to suit students' requirements, which will subsequently be helpful in enhancing discussion interaction and students' learning experience.

Keywords: Small private online courses, Topic tracking, Topic intensity, Topic content, Discussion forums

Introduction

Small Private Online Courses (SPOCs), considered an extension of Massive Open Online Courses (MOOCs), provide a flexible and hybrid learning mode that combines offline classroom teaching and online distance teaching in higher education (Fox, 2013; Freitas & Paredes, 2018; Wang & Zhu, 2019). By means of SPOC discussion forums, students can deliver their course posts for exchange and feedback of viewpoints (Combéfis, Bibal, & Van Roy, 2014; Filius et al., 2018; Liu et al., 2019, b). With the ever-growing textual contributions from students, instructors find it difficult to manually detect and track students' behavioral patterns and discourse content (e.g., dynamic

topic interests, emotional orientations). A large body of studies have focused on either students' behavioral patterns and social interaction across different discussion posts (Gitinabard, Heckman, Barnes, & Lynch, 2019; Wang, Fang, & Gu, 2020) or the influence of students' forum participation on course performance (Chen et al., 2018; Chiu & Hew, 2018; Moreno-Marcos, Alario-Hoyos, Muñoz-Merino, & Kloos, 2018; Phan, McNeil, & Robin, 2016). However, there is still much to be explored in the matter of discovering valuable semantic information from students' posts to interpret course discussion dynamics in online learning platforms (Almatrafi & Johri, 2018; Ramesh, Goldwasser, Huang, Daume, & Getoor, 2014).

A practical example is that when students participate in course discussion, they can freely express positive or negative opinions toward various course aspects (e.g., specialized knowledge, course examinations, learning resources) by posting or replying. Student-generated discourse content that contains students' focused topics, emotional tendencies and behavioral patterns might be changed over time (Liu et al., 2019, a; Ramesh, Kumar, Foulds, & Getoor, 2015; Wen, Yang, & Rose, 2014). Moreover, the analysis of discussion posts in different time intervals to some extent contributes to the interpretation of students' external behavioral motivation as well as focused topics (Liu et al., 2019, b; Peng & Xu, 2020). Topic intensity about the occurrence probability of students' discussed topics and topic content about the key concepts of students' focused topics typically change in different time periods, resulting in "topic inheritance" and "topic variation" (Griffiths & Steyvers, 2004; Wong, Wong, & Hindle, 2019). Therefore, detecting temporal topic changes derived from students' posts is essential in understanding and assessing the process of online course discussions. Specifically, how students' focused topics evolve can be used as an important clue to predict the priorities of course development for the next course offering. This finding will help instructors obtain a holistic view of the evolution of course forums and conveniently navigate valuable information. Consequently, instructors can provide earlier adjustment of pedagogical methods and learning materials to suit future students' requirements, which might motivate students to participate in course forums. Additionally, manually assigning topic and time labels to discussion posts requires considerable labor and resources; thus it is difficult to adopt this approach on a large scale.

Owing to the overwhelming abundance of documents generated by users, it is challenging for managers to effectively locate and navigate information. Latent Dirichlet Allocation (LDA) considered a standard topic model, has been proposed to solve this task in the business intelligence field (Blei, Ng, & Jordan, 2003; Blei & Lafferty, 2007; Dupuy, Bach, & Diot, 2017; Mo, Kontonatsios, & Ananiadou, 2015). In education, an increasing number of researchers have indicated that using the variants of LDA is an appropriate method to detect and track the semantic information of student-generated discourse content (Ezen-Can, Boyer, Kellogg, & Booth, 2015; Wong et al., 2019).

Concerning the aforementioned aspects, this paper performs an evolutionary analysis of topic detection (topic intensity and topic content) derived from students' discussion posts by using an improved unsupervised dynamic topic model, called Time Information-Emotion Behavior Model (TI-EBTM). It is an extension of the standard LDA model, which characterizes multiple features (i.e., time, emotion, and behavior) associated with course posts to guide the generative process of the language model. TI-EBTM is flexible enough to be applied in other practical application scenarios, such as

peer assignments, chat rooms, and course reviews. It can also be considered a vehicle of information technology (IT) to enhance learning in higher education. The analytical results of topic tracking will help educational practitioners effectively monitor the dynamics of course forums and promote educational practitioners' self-reflection to enhance course intervention and adjustment for meeting students' expectations.

Related work

Topic detection in e-learning

Topic detection and tracking (TDT) was first proposed to discover topically related material in streams of data (Wayne, 1997). In TDT, topic detection is the premise of topic tracking, and a topic is defined as a series of related activities or streams of events. Many studies have detected hidden topics and opinions in user-generated discourse content on commercial websites (Dupuy et al., 2017; Mo et al., 2015; Rossetti, Stella, & Zanker, 2016; Westerlund, Mahmood, Leminen, & Rajahonka, 2019) and social media platforms (Reyes-Menendez, Saura, & Alvarez-Alonso, 2018; Xie, Zhu, Jiang, Lim, & Wang, 2016).

In the online learning context, one of the primary tasks is to capture the latent topics students prefer to talk about, allowing for instructor guidance and interventions to facilitate students' course performance. Ezen-Can et al. (2015) combined a k-medoids clustering algorithm with an LDA topic model to categorize similar posts into different groups and subsequently extract the key concepts of each group. Elgort, Lundqvist, McDonald, and Moskal (2018) provided teachers with a new text mining tool called Quantext to extract key topics of interest expressed by students in MOOC discussion forums. Xu and Lynch (2018) employed deep learning models to automatically classify question posts and detect their question types according to the corresponding meaning. Wong et al. (2019) conducted an empirical study of five Coursera courses to detect topics from the course material and classify discussion posts using both unsupervised and supervised variants of LDA.

Another important task in topic detection is identifying how emotions and behaviors related to different topics are presented, which holds great promise for providing adaptive support to students. Ramesh et al. (2015) proposed a weakly supervised model called hinge-loss Markov random fields in jointly modeling dependencies between aspect and emotion. This model was validated to effectively identify various aspects of MOOC forum posts and infer these aspects' emotional polarities. Liu, Yang, Peng, Sun, and Liu (2017) developed an improved topic model named Emotion Topic Joint Probabilistic Model (ETJM) that was an extension of Sentence-LDA (SLDA). Employing ETJM in online course reviews enabled the automatic identification of various pairs of negative emotions and aspects. Zhao, Cheng, Hong, and Chi (2015) addressed that users' behavioral signals should be regarded as unique characteristics to construct users' topical interests separately. Along the same line, Liu et al. (2019, a) jointly incorporated emotion and behavior into the generation of Behavior-Sentiment Topic Mixture (BSTM) topic model. This model could be applied to large amounts of MOOC review data to unfold learners' focused topics as well as learners' attitudes and behavioral patterns toward these topics. Although these studies can effectively identify hidden topics that learners are concerned about, the neglect of the temporality of topic

evolution limits the ability of these models to mine discourse data to discover the overall development of topic profiles.

Topic tracking using dynamic topic models

To track topic changes over time, dynamic topic models have been shown to be a promising novel technology (Liu et al., 2019, b; Wong et al., 2019). According to the time granularity of documents (e.g., literatures, discussions, reviews, news), some studies have first divided text sets into different sliding windows and then employed topic models for each time stamp to obtain a holistic view of topic evolution (Andrei & Arandjelović, 2016; Blei & Lafferty, 2006). For instance, Blei and Lafferty (2006) proposed a sequential topic model Dynamic Topic Model (DTM) to track the time evolution of topics by discretizing large document sets. DTM employed state space models on the natural parameters of multinomial topic distributions. DTM assumed that the posterior parameters of the model at the current moment were treated as the conditional distribution of the model at the next moment.

Some studies have focused on the qualitative topic evolution over time by examining the posterior discretization results of a topic and the time information of text sets (Garroppo, Ahmed, Niccolini, & Dusi, 2018; Griffiths & Steyvers, 2004; Wang & McCallum, 2006). In this line, Wang and McCallum (2006) developed a non-Markov continuous-time model called Topics over Time (TOT). Based on the hypothesis that each topic followed a multinomial distribution over time, TOT could be utilized to detect topic changes over time by counting the number of words related to each topic in each time window.

Furthermore, some studies have highlighted the importance of time characteristic in guiding the generative process of topic models to capture the evolution of topics involved with emotions (Dermouche, Velcin, Khouas, & Loudcher, 2014; He, Lin, Gao, & Wong, 2014). For example, Dermouche et al. (2014) devised an LDA-based topic model called Time-Aware Topic-Sentiment (TTS) to detect topic-sentiment evolution over time. TTS constructed a direct mapping interdependency that addressed time was associated with topics and sentiments.

Additionally, Ramesh and Getoor (2018) combined the release time of each post and seeded topic model to discover the evolution of fine-grained topics from two MOOC course discussion forums. Liu et al. (2019, b) proposed a temporal emotion-aspect model (TEAM) for identifying students' concerns and capturing the evolution of students' attitudes toward topics in SPOC course forums.

Compared with the aforementioned dynamic topic models, TI-EBTM proposed by this study considers the time information in the generative process of topic modeling without requiring post-processing to infer time evolution. TI-EBTM hypothesizes that each topic is jointly related to corresponding emotional and behavioral features. Therefore, TI-EBTM has great potential of uncovering topic changes from a global view, rather than detecting the topic evolution in each time stamp. Moreover, the evolution of topics' emotional orientations and behavioral patterns can be tracked over time.

Research questions

To our knowledge, some studies have employed static topic models to detect students' topics of interest from textual-based discussion content in online learning

communities. However, these studies have typically ignored topics' temporal granularity as well as their emotional and behavioral characteristics over time. The purpose of this study is to propose and use an optimized dynamic topic model TI-EBTM that embeds time, emotion and behavior attributes of a topic, to track temporality of topics (i.e., the temporal changes of topic intensity and topic content) derived from student-generated posts in SPOC discussion forums.

Specifically, this study addresses the following three questions:

- (1) To ensure the best generalization capability of TI-EBTM, what is the optimal number of topics? Compared with the other dynamic topic modeling approaches, is TI-EBTM better?
- (2) How does topic intensity (i.e., topics' probability distributions, emotional tendencies, and behavioral patterns) that occurs in student-generated posts evolve over time in SPOC discussion forums?
- (3) How does topic content that occurs in student-generated posts evolve over time in SPOC discussion forums?

Time information-emotion behavior topic model (TI-EBTM)

In this section, to answer the research questions above, this study employs a temporal topic model to automatically detect and track student-generated discussion content in SPOC forums. On the one hand, we introduce some prior concepts that make TI-EBTM understandable. On the other hand, we describe the generation principle of TI-EBTM and the process of parameter inference.

Basic concepts

- The seeded words represent a series of key features and domain-oriented concepts that are easily utilized to distinguish different topics, such as “信息化/informatization” and “指针/pointer”. Based on the observation of specialized terms from courses, we construct a situational dictionary including a set of seeded words to compulsively divide and bind words.
- The emotional lexicon contains 22,478 emotion terms, including 9584 positive terms (e.g., “棒/wonderful”) and 12,884 negative terms (e.g., “糟糕/bad”). These emotion terms are selected and summarized from a Chinese sentiment dictionary (Ku, Liang, & Chen, 2006), a Chinese commendatory and derogatory dictionary v1.0 (Li, 2011), and HowNet (Dong, 2013).
- The behavioral categories, including thread posting, replying, quoting, and common posting, are defined according to the students' behavioral interactions with a Chinese SPOC discussion forum. Table 1 presents the coding scheme of the students' four discussion behaviors. Notably, each post is assigned just one behavioral label.
- The course posts are student-generated discourse content involving students' opinions and attitudes toward various aspects of a course in SPOC discussion forums. Each post consists of several words and has a time label corresponding to the division of time stamp, as shown in Fig. 1.

Table 1 The coding scheme of students' interactive behaviors in discussion forums

Label	Behavioral Categories	Basic description of behavioral categories
TP	Thread Posting	The post is written by a student to start a new thread (including hot topics, unsolved problems, etc.)
RE	Replying	The post that involves “回复/reply”, “mention”, or “@” is delivered to response to someone's message.
QU	Quoting	The post that contains “http”, “quote”, “blockquote”, or other similar symbols is written to argue someone's opinions.
CP	Common Posting	Common posting is different from replying and is not a direct response to someone's message.

TI-EBTM description

Considering the time, emotion and behavior attributes of student-generated posts, this study proposes an unsupervised dynamic topic model called Time Information-Emotion Behavior Model (TI-EBTM), as shown in Fig. 2. TI-EBTM is composed of four layers: the document layer, the topic layer, the time-emotion-behavior layer, and the word layer. In this study, each document is equivalent to each post. Each post can be represented as multiple topics; each topic can be constructed as multiple words; each word in a post shares the same time stamp and each word is assigned the corresponding emotion and behavior labels. Like LDA, TI-EBTM assumes that the number of topics is known and fixed; the order of words in the post and the order of posts in the corpus is irrelevant. In addition, TI-EBTM emphasizes that topics follow normalization in the entire time circle without discretizing the corpus in advance. That is, over time, topics draw a global multinomial distribution similar to the topic-word distribution.

In Fig. 2, TI-EBTM is shown as a probability graph model that is a directed acyclic graph based on the Bayesian network. The nodes in the graph represent random variables; the hollow circle refers to unknown hidden variables, such as topics z in the document; and the solid circle represents known observed variables, such as words w in the document. Table 2 describes the symbols used in TI-BETM.

The detailed TI-EBTM generation process is as follows. First, TI-EBTM assumes that all students publish a total number of M course posts in the discussion forum, and each of which can be represented as $R = \{r_1, r_2, \dots, r_m\} (1 \leq m \leq M)$. In TI-EBTM, each post consists of a series of hidden topics, which can be formally denoted as $m = \{z_1, z_2, \dots, z_K\}$. Second, it is assumed that the topic is related to emotion, behavior, and time information. The topic is subject not only to binomial emotion distribution π_{ke} and multinomial behavior distribution ψ_{kb} , but also to multinomial time distribution ζ_{kt} . When a word in a post is randomly sampled, it is given a topic tag and a time attribute. The

Example1: 第一次接触数据结构课程的相关概念, 时间复杂度存在较多的疑点, 怎么计算时间复杂度的问题? 比较难懂, 烦啊! #CP #TW9

It is the first time to contact the related concepts of data structure curriculum, there are some doubts on time complexity. How to compute the problem of time complexity? It's hard to understand, boring! #CP #TW9

Fig. 1 An example of the time label assigned to a post

Fig. 2 Time Information-Emotion Behavior Model (TI-EBTM)

Symbols	Description	Symbols	Description
M	Number of posts	α	Dirichlet prior of θ
N	Number of words	β	Dirichlet prior of φ
K	Number of topics	γ	Dirichlet prior of π
z	Topic index	η	Dirichlet prior of ψ
e	Emotion index	λ	Dirichlet prior of ζ
\mathbf{b}	Behavior index {TP, RE, QU, CP}	θ	Distribution of post-topic
w	Word index	φ	Distribution of topic-word
t	Time index	π	Bernoulli distribution of topic-emotion
ζ	Distribution of topic-time	ψ	Distribution of topic-behavior

$$p(w, b, e, z, t | \alpha, \beta, \gamma, \eta, \zeta) = p(w | z, e, \beta) \cdot p(t | z, e, b, \lambda) \cdot p(e | z, \gamma) \cdot p(b | z, \eta) \cdot p(z | \alpha) \quad (1)$$

where the factors of the formula on the right side are expanded as follows:

$$p(z | \alpha) = \int p(z | \theta) p(\theta | \alpha) d\theta = \left(\frac{\Gamma\left(\sum_{z=1}^K \alpha\right)}{\prod_{z=1}^K \Gamma(\alpha)} \right)^D \cdot \frac{\prod_{z=1}^K \Gamma(n_z^{(k)} + \alpha)}{\prod_{z=1}^K \Gamma\left(n_d + \sum_{z=1}^K \alpha\right)} \quad (2)$$

$$p(b | z, \eta) = \int p(b | \psi) p(\psi | \eta) d\psi = \left(\frac{\Gamma\left(\sum_{b=1}^B \eta\right)}{\prod_{b=1}^B \Gamma(\eta)} \right)^K \cdot \frac{\prod_{b=1}^B \Gamma(n_z^{(b)} + \eta)}{\prod_{b=1}^B \Gamma\left(n_z + \sum_{b=1}^B \eta\right)} \quad (3)$$

$$p(e | z, \gamma) = \int p(e | \pi) p(\pi | \gamma) d\pi = \left(\frac{\Gamma\left(\sum_{e=1}^E \gamma\right)}{\prod_{e=1}^E \Gamma(\gamma)} \right)^K \cdot \frac{\prod_{e=1}^E \Gamma(n_z^{(e)} + \gamma)}{\prod_{e=1}^E \Gamma\left(n_z + \sum_{e=1}^E \gamma\right)} \quad (4)$$

$$p(t | z, e, b, \lambda) = \int p(t | \zeta) p(\zeta | \lambda) d\zeta = \left(\frac{\Gamma\left(\sum_{t=1}^T \lambda\right)}{\prod_{t=1}^T \Gamma(\lambda)} \right)^{K \cdot B \cdot E} \cdot \frac{\prod_{z=1}^K \prod_{b=1}^B \prod_{e=1}^E \Gamma(n_z^{(t)} + \lambda)}{\prod_{z=1}^K \prod_{b=1}^B \prod_{e=1}^E \Gamma\left(n_z + \sum_{t=1}^T \lambda\right)} \quad (5)$$

$$p(w | z, e, \beta) = \int p(w | \phi) p(\phi | \beta) d\phi = \left(\frac{\Gamma\left(\sum_{w=1}^V \beta\right)}{\prod_{w=1}^V \Gamma(\beta)} \right)^{K \cdot E} \cdot \frac{\prod_{z=1}^K \prod_{e=1}^E \prod_{w=1}^V \Gamma(n_z^{(w)} + \beta)}{\prod_{z=1}^K \prod_{e=1}^E \prod_{w=1}^V \Gamma\left(n_z + \sum_{w=1}^V \beta\right)} \quad (6)$$

Then, according to the iterative rule of Gibbs sampling, all words in the post set are randomly sampled. That is, excluding the topic of the current word, the topic label of the current word is assigned through the topic probability distribution of other words in the post. The calculation formula is given by:

$$p(z_i, e_i, b_i, t_i | z_{-i}, e_{-i}, b_{-i}, t_{-i}, w_{-i}, \alpha, \beta, \gamma, \eta, \lambda) = \frac{p(z, e, b, t, w, \alpha, \beta, \gamma, \eta, \lambda)}{p(z_{-i}, e_{-i}, b_{-i}, t_{-i}, w_{-i}, \alpha, \beta, \gamma, \eta, \lambda)} \\ \propto \frac{(n_{d,k})_{-i} + \alpha}{(n_d + K\alpha)} \cdot \frac{(n_{z,e,w})_{-i} + \beta}{(n_{z,e} + V\beta)} \cdot \frac{(n_{z,e})_{-i} + \gamma}{(n_z + E\gamma)} \cdot \frac{(n_{z,b})_{-i} + \eta}{(n_z + B\eta)} \cdot \frac{(n_{z,b,e,t})_{-i} + \lambda}{(n_{z,b,e} + T\lambda)} \quad (7)$$

where z_i denotes the word that belongs to topic z in the i th position of post m ; z_{-i} indicates the words that belong to topic z in positions other than the current word; E represents the emotional category {positive and negative}; B represents the behavior category {TP, RE, QU, CP}; and T represents the granularity of time division $\{t_1, t_2, \dots, t_n\}$.

Finally, we use the following formulas to identify the hidden parameters, which can not only locate the evolutionary change of dynamic topic information as a whole, but also track the evolutionary trend of emotional and behavioral tendencies toward the topics.

$$\theta_{mk} = \frac{N_{mk}^{MK} + \alpha}{\sum_{k=1}^K N_{mk}^{MK} + K \cdot \alpha} \quad (8)$$

$$\phi_{kw} = \frac{N_{kw}^{KW} + \beta}{\sum_{w=1}^V N_{kw}^{KW} + V \cdot \beta} \quad (9)$$

$$\pi_{kew} = \frac{N_{kew}^{KEW} + \gamma}{\sum_{e=1}^E N_{kew}^{KEW} + E \cdot \gamma} \quad (10)$$

$$\psi_{kbw} = \frac{N_{kbw}^{KBW} + \eta}{\sum_{b=1}^B N_{kbw}^{KBW} + B \cdot \eta} \quad (11)$$

$$\zeta_{kebt} = \frac{N_{kebt}^{KEBT} + \eta}{\sum_{t=1}^T N_{kebt}^{KEBT} + T \cdot \eta} \quad (12)$$

Methodology

Research context

The specific research context of this study occurs in the SPOC discussion forum of a Chinese university, named *starC* (<http://spoc.ccnu.edu.cn/starmoonHomepage>). This online hybrid cloud classroom takes advantage of MOOC to conduct extensive learning from conventional classrooms. Based on the cloud computing architecture, this platform mainly provides online learning resources and services for college students. Discussion forum, as an interactive module of this platform, is favored by students for communication and exchange of views.

Data collection

To validate the robustness of TI-EBTM and its effectiveness in practical application research, we stored and archived two types of discussion post data sets, $Data_m$ and $Data_c$, in the next semester (17 weeks) of 2016. To enable further data operation and analysis, these data were uniformly recorded in sheet form. $Data_m$ was derived from a random selection of 30 courses (15 in arts and 15 in sciences), with a total of 15,357 posts. $Data_m$ was used to emphasize the internal validity of TI-EBTM by evaluating the quality of dynamic topics' similarity and segmentation. $Data_c$ was selected from a course called Data Structure based on a comprehensive assessment of both more students and posts, with a total of 100 students and 817 posts. $Data_c$ was employed to analyze the external validity of TI-EBTM by conducting an empirical investigation to obtain an in-depth understanding of the temporal evolution of topics. The basic statistics of $Data_m$ and $Data_c$ are shown in Table 3. According to our previous empirical research experiences, we chose 9 time units (2 week as a time unit) during a period of 17 weeks. The reason came from two aspects: on the one hand, the division of smaller time units tended to result in the lack of data, which made it difficult to exploit the efficiency of

Table 3 Basic statistics of the post datasets

Datasets	Number of courses	Number of students	Number of posts	Number of characters	Avg. number of characters per post	Avg. number of posts across course
Data _m	30	1, 527	15, 357	4, 167, 245	324.36	1706.33
Data _c	1	100	817	105, 700	129.38	90.78

the evolutionary model, whereas the division of larger time units made it difficult to reveal a more detailed evolutionary overview of topic tracking; on the other hand, the time unit of 2 weeks was more consistent with the schedule of teachers' discussion activities on specific teaching topics. Notably, the smaller time granularity or the larger time granularity was a relative selection, which should be combined with the data size of each time unit and the specific activity arrangement of teaching practice.

Evaluation metrics

In this study, we used three evaluation indexes perplexity, similarity, and entropy to measure the performance of TI-EBTM (Liu et al., 2019, b). Perplexity is used to determine the optimal number of topics and to measure the generalization ability of topic models for processing unknown data. The lower the perplexity value is, the better the model's generalization ability is. The formula is as follows:

$$Perplexity(D_{test}|\text{Model}) = \exp \left\{ -\frac{\sum_{d=1}^M \log p(w_d)}{\sum_{d=1}^M N_d} \right\} \quad (13)$$

where Model refers to the currently used dynamic topic model; D_{test} indicates the model input test set; M represents the total number of posts; w_d denotes the joint probability value of words generated in post d ; and N_d denotes the number of words in post M .

Similarity is adopted to assess the overall similarity between topics. The lower the similarity value is, the better the model's topic segmentation ability is. The formula is presented as:

$$Sim(z_i, z_j) = \frac{1}{K(K-1)} \sum_{i=1}^{K-1} \sum_{j=i+1}^K \frac{\sum_{t=1}^T \zeta_{it} \zeta_{jt}}{\sqrt{\left(\sum_{t=1}^T \zeta_{it}^2 \right) \left(\sum_{t=1}^T \zeta_{jt}^2 \right)}} \quad (14)$$

Entropy is used to measure the degree of word aggregation within topics. The lower the entropy is, the higher the information consistency within topics is. The formula is presented as:

$$Entropy = \frac{1}{K} \sum_{k=1}^K \sum_{t=1}^T (-\zeta_{kt} \log \zeta_{kt}) \quad (15)$$

where ζ_{it} and ζ_{jt} represent the probability distribution of topics in the entire time

period; K indicates the total number of topics; and T denotes the total number of time units.

Data analysis

For the first research question, we used the perplexity index to determine the suitable number of topics in TI-EBTM; we then adopted similarity and entropy indicators to compare the evolutionary quality of topic generation among Post-LDA (Griffiths & Steyvers, 2004), ATU-LDA (Author Time Unit-LDA) (Rosen-Zvi, Griffiths, Steyvers, & Smyth, 2004), and TI-EBTM. In these models, the total number of model iterations was 500. The prior parameters α , β , γ , η , and λ were set as 0.1, 0.01, 0.1, 0.01, and 0.1, respectively. Notably, to ensure the validity of model comparison, the former two methods also integrated emotional and behavioral information into the temporal topic models.

For the second research question, we employed the TI-EBTM technique to investigate the evolution of topic intensity across the course. For the last research question, we utilized the KL (Kullback–Leibler divergence) (Andrei & Arandjelović, 2016) index to compute the topics' semantic relevance between different time units, allowing for the presentation of the evolution of topic content across the course.

Results

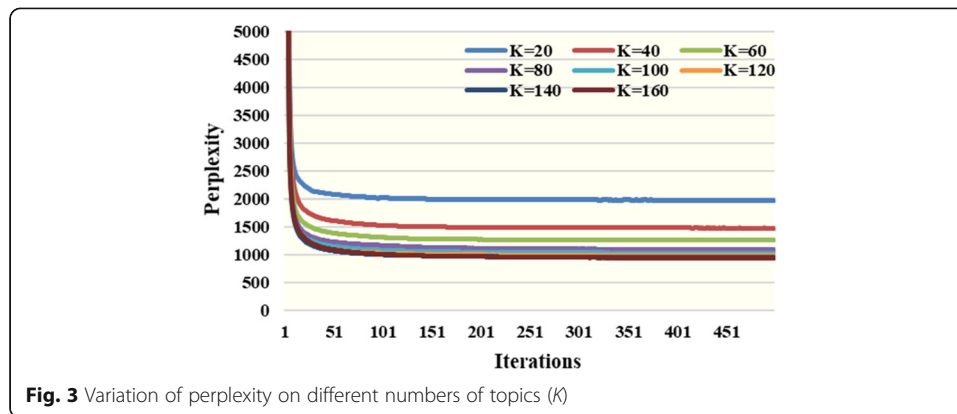
Regarding the three research questions formulated in section 3, we validated the effective performance of TI-EBTM and performed an empirical investigation to obtain an in-depth understanding of the temporal evolution of topics.

Performance of TI-EBTM

The optimal number of topics

To ensure effective model evaluation under the same number of topics and in the same controlled experimental environment, we used the perplexity index to determine the optimal topic number of TI-EBTM by analyzing 30 courses' discussion posts. By gradually increasing the number of topics from 0 to 500 with an interval of 50, we recorded and stored the corresponding model perplexity value. When the model perplexity value no longer decreased, the number of topics was selected as our model's input parameter. The resulting performance curves of model perplexities under different topics are depicted in Fig. 3.

In Fig. 3, all curves on the model perplexity values showed a certain regularity in the entire iteration cycle. Specifically, when the number of model iterations was within the interval of [1, 50], the perplexity value of the model changed greatly and declined suddenly, from 5000 to 1000. When the number of model iterations was within the interval of [50, 500], all curves rapidly reached a stable state. This showed that the model had a strong ability of fast convergence. When the number of topics was increased to 140, the model's average perplexity value reached the lowest level of 1382.04. Moreover, when the number of topics was set as 160, the model's average perplexity value no longer declined and showed a slight upward trend, reaching 1390.32. This indicated that when the number of topics exceeded 140, the generalization ability of the model no

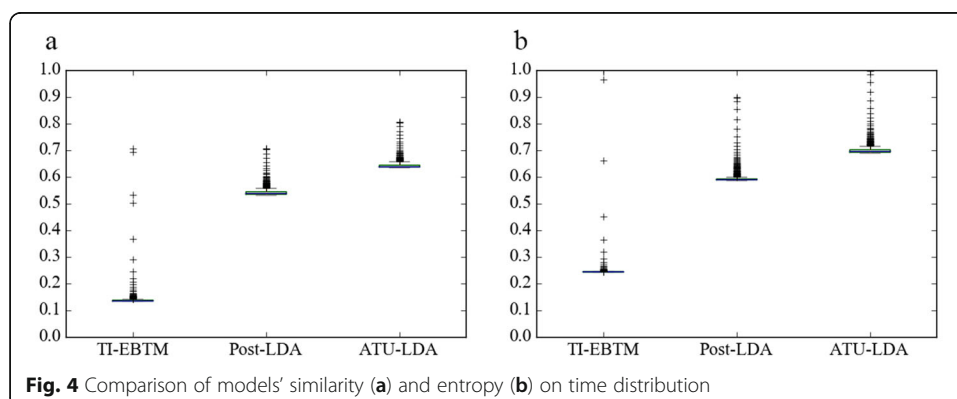


longer increased but rather decreased. Therefore, the optimal number of topics was set as 140.

Comparison of dynamic topics' similarity and segmentation

To verify the quality of topic generation by TI-EBTM, we used a total number of 15,357 posts from the $Data_m$ dataset. We conducted model comparisons in two dimensions of topic-time distribution on similarity and aggregation. Notably, TI-EBTM regards time as an internal unit of global topics, similar to the word granularity of topic correlation. If the similarity between different topics is lower, the quality of the generation of dynamic topics is higher. If the features within topics are more representative, the quality of generating dynamic topics is higher. To ensure the validity of model comparisons, we set the number of topics and the number of iterations as 140 and 500, respectively, for TI-EBTM, Post-LDA, and ATU-LDA.

Fig. 4a and b presented the similarity and aggregation of TI-EBTM, Post-LDA, and ATU-LDA in the quality of the generation of dynamic topics. TI-EBTM obtained lower values for topic-time similarity and topic-time entropy. This indicated that temporal topics generated by TI-EBTM could be effectively distinguished from each other and



characterized by the internal time units. That is, by integrating learners' posting time information, TI-EBTM achieved better performance in tracking dynamic topics.

Evolutionary analysis of topic intensity over time

Evolutionary trends of students' focused topics

Except for the quantitative evaluation of TI-EBTM performance, we conducted a case study of TI-EBTM on the Data_c dataset from the Data Structure course. The total number of topics in Data Structure was set as 20. According to the output topic-time matrix ζ_{kt} , each row vector represented the time probability distribution of a topic; that is, the probability sum of each topic in different time units was 1. Therefore, we could dynamically capture and track the global intensity changes of learners' focused topics in the entire time cycle. The experimental results for four representative topics (the probability of each topic was higher than the average value of 0.05) were displayed in Fig. 5. Notably, the division into smaller time intervals caused a lack of data, while the division into larger time units made it difficult to detect a more detailed overview of the topic evolution. Thus, given that the course duration lasted 17 weeks, 2 weeks were coded into an independent time unit, and 9 time units were obtained.

As shown in Fig. 5, four curves of typical topic evolution presented a certain similarity overall and showed a trend of rapid rising and falling. Moreover, there seemed to be two stages of topic evolution (i.e., 1th week to 10th week, 11th week to 17th week). At the beginning and middle of the semester, the temporal intensity of all topics fluctuated slowly. At the next stage of the semester, the intensity of the topics, except that of topic 14, increased rapidly, peaked, and subsequently descended to near 0. This indicated that students tended to talk about focused topics within the specific time period, especially in the next stage of the semester. Although these topics were mostly discrete, they showed some continuity in similar time units. For example, from the 11th week to the 17th week, students' attention to topic 14 showed continuous evolution, signifying that some topics students preferred to discuss had a certain inheritability.

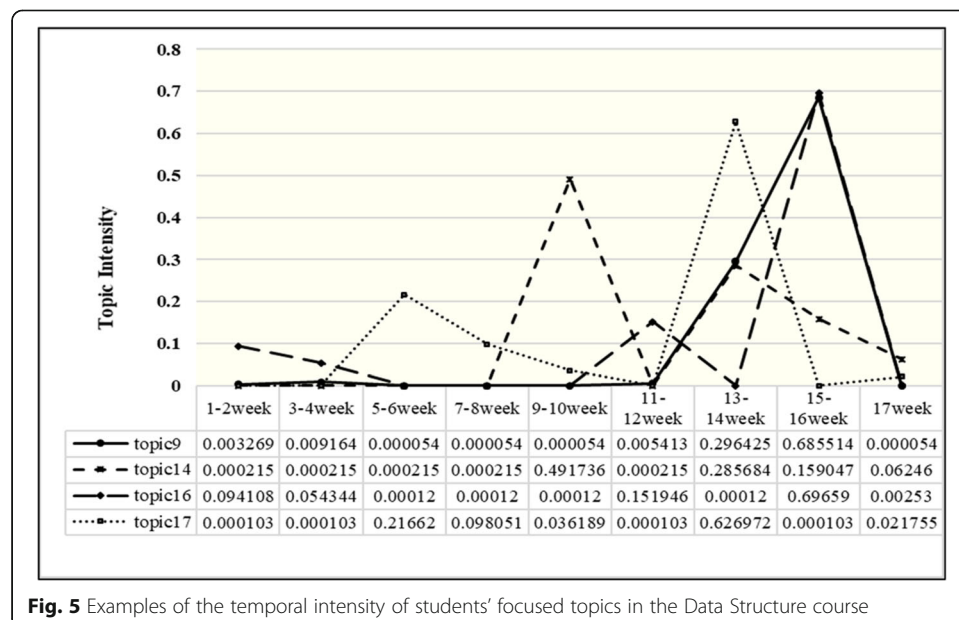


Fig. 5 Examples of the temporal intensity of students' focused topics in the Data Structure course

Additionally, according to the output topic-word matrix ϕ_{kw} for the four typical topics (topic 9, topic 14, topic 16, and topic 17), we listed each topic's representative semantic concepts (topic words, positive emotion words, and negative emotion words). Through empirical logic deduction based on the observation of these concepts, we assigned the corresponding topic label for each topic. A detailed description of each topic's semantic content is displayed in Table 4.

As shown in Table 4, topic 16 had the highest probability among the listed topics that were frequently discussed by students at 0.07, which was higher than the average value of 0.05. Topic 16 was mainly about the foundational concepts and method of data query, including some seeded words, such as tree, list, search, sort, and traverse. Moreover, students expressed emotional opinions of the data query method by using the positive word understand and the negative word conflict. In contrast to topic 16, topic 17 mainly involved prior knowledge and method of data storage using the words list, queue, tree, pointer, etc. Topic 14 mainly referred to programming implementation. Compared with topics 14, 16 and 17, topic 9 seemed to be a global topic involving the development and application of data structure. In addition, the listed topics were all assigned positive labels (+). Hence, the analytic results indicated that as a whole,

Table 4 Examples of four typical topics and their probabilities over words in the Data Structure course. These representative words, including topic words (not underlined) and emotion words (underlined) have higher probabilities. (+) denotes that the topic's overall emotion is positive. The probability value below (+) represents the average probability of each topic appearing in the next semester of the course

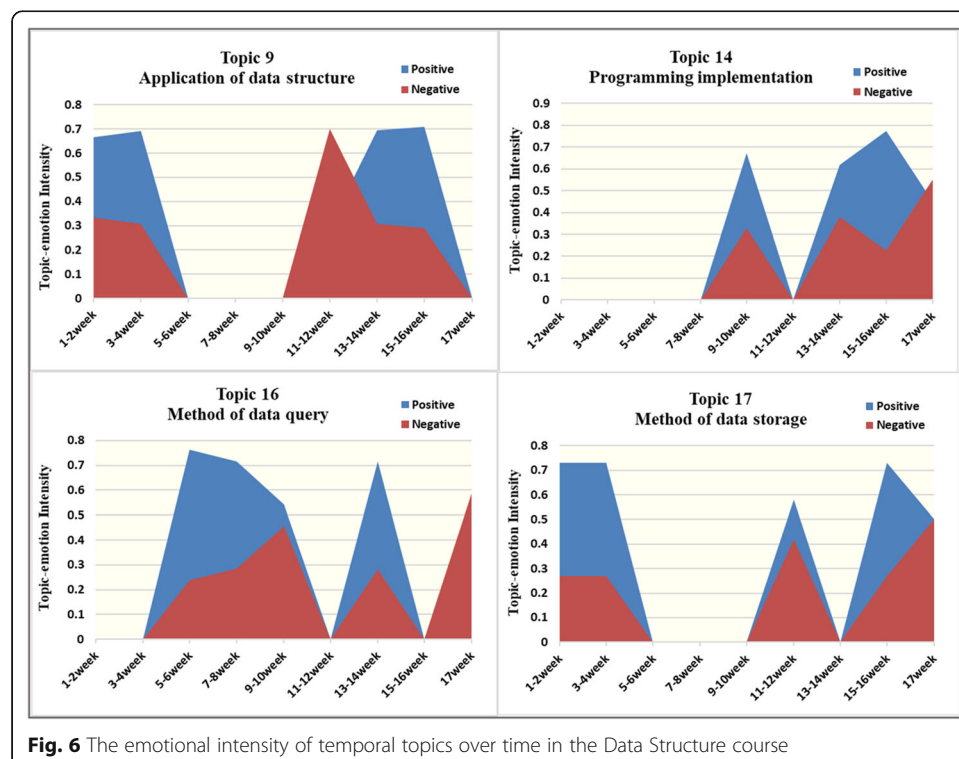
Course	Topics	Topic Label	Words	Prob.	Words	Prob.
Data Structure	Topic 9 (+) (0.054)	Application of data structure	结构/structure	0.0397	数据/data	0.0371
			<u>复杂/complicated</u>	0.0122	设计/design	0.0106
			程序/program	0.0101	人才/talents	0.0090
			信息/information	0.0080	培养/cultivate	0.0080
			<u>正确/right</u>	0.0042	<u>不清楚/unclear</u>	0.0011
	Topic 14 (+) (0.051)	Programming implementation	串/string	0.0943	位置/index	0.0615
			<u>理解/understand</u>	0.0267	语句/statement	0.0143
			循环/loop	0.0123	输出/printf	0.0123
			代码/code	0.0123	<u>错/wrong</u>	0.0102
			掌握/master	0.0080	<u>容易/easy</u>	0.0101
	Topic 16 (+) (0.070)	Method of data query	树/tree	0.0293	表/list	0.0235
			链/chain	0.0141	查找/search	0.0141
			排序/sort	0.0118	遍历/traverse	0.0106
			路径/path	0.0010	插入/insert	0.0094
			<u>理解/understand</u>	0.0059	<u>冲突/conflict</u>	0.0047
	Topic 17 (+) (0.054)	Method of data storage	指针/pointer	0.0212	树/tree	0.0151
			表/list	0.0151	链/chain	0.0141
			结构/structure	0.0131	元素/code	0.0123
			队/queue	0.0081	存储/store	0.0081
			<u>错误/error</u>	0.0051	<u>确定/confirm</u>	0.0020

students tended to express more positive opinions toward the specialized course knowledge over weeks.

Evolutionary trends of topics' emotion and behavior intensity

To better explain the semantic development of dynamic topics, we also captured the temporal changes of topics' emotion and behavior intensity. After the completion of iterative topic sampling by TI-EBTM, each word in the students' discussed posts was given the corresponding topic assignment that was closely correlated with the emotion and behavior characteristics. We first counted the number of emotional and behavioral categories associated with each topic that simultaneously appeared in different intervals of course offering. Then, the normalized method was used to compute the topic-emotion and topic-behavior probability across the Data Structure course. Finally, the evolutionary trends of the topics' emotional and behavioral intensity were visualized, as shown in Figs. 6 and 7.

In Fig. 6, we observed that the four topics all showed a higher positive emotion probability on the whole (the maximum value was close to 0.7 in some time intervals). That is, students who participated in Data Structure were more likely to use positive discourse to express their inner views and attitudes toward these topics. Notably, if students did not discuss a topic in a certain time interval, the emotional intensity of the topic was close to 0 and thus would not be shown in the corresponding area in Fig. 6. In different time segments, the positive emotional value and negative emotional value of the topic could not represent students' overall emotional tendency, but students' different level of emotional attention to the topic in each time unit. In addition, taking



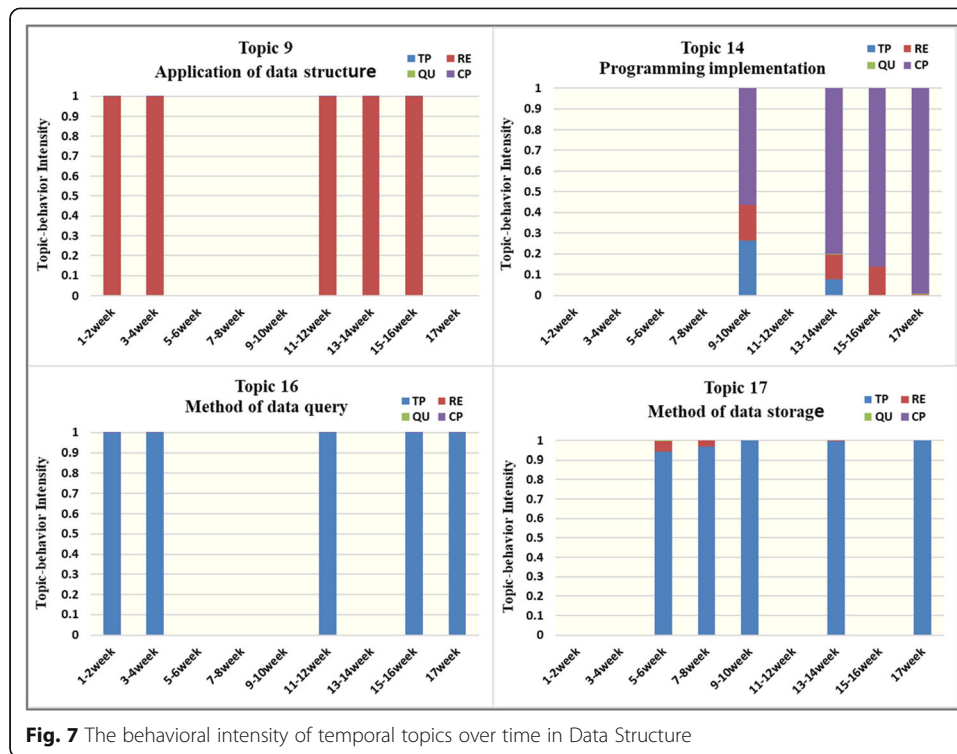


Fig. 7 The behavioral intensity of temporal topics over time in Data Structure

topic 16 as an example, students negatively evaluated the method of data query in the 17th week. Therefore, instructors should attach importance to the topic's negative evaluation in the specific time interval and find out the possible causes of students' adverse emotional responses.

In Fig. 7, the four topics along with four behavioral probability distributions followed similar evolutionary characteristics. These topics were discrete and dominated by one behavior over time. Specifically, topics 16 and 17 were dominated by TP, indicating that students were prone to actively initiate the topic discussion about method of data query and method of data storage. A possible explanation is that students did not have a good grasp of the relevant knowledge, so they needed to take the initiative to express their opinions and seek help from their peers. Compared with topics 16 and 17, topics 9 and 14 were dominated by RE and CP that might be regarded as relatively passive behaviors for students' knowledge construction (Tobarra, Robles-Gómez, Ros, Hernández, & Caminero, 2014). Regarding topic 14, programming implementation, we found that students tended to state their views toward the topic through more richly interactive ways (i.e., CP, RE, and TP) in some time periods and showed a certain continuity at the end of the semester. Thus, the analytic results on the behavioral intensity of dynamic topics are helpful for instructors to establish a macroscopic overview of topic evolution and to find out students' behavioral motivation of expressing different topics.

Evolutionary analysis of topic content over time

With respect to the interactive development rules of online courses, students' focuses usually shift in the temporal changes of topic intensity and topic content. The evolution of topic content presents the differences in the spacing of feature words belonging to

the same topic in different time units, which is specifically represented by the relevant semantic words within the topic. The semantic measurement between topics can be matched by the similarity calculation of topic-word distribution ϕ_{kw} . To conduct the content analysis of dynamic topics, we adopted symmetric Kullback-Leibler divergence to judge the semantic relevance of topic-word ϕ_{kw} . The smaller the distance between KL is, the smaller the difference between topics is. In this case, the stronger topic relevance is more likely to form the continuous evolution of topic content. KL is computed by the following formulas:

$$KL(Q_{k,w} \| P_{k,w}) = \sum_i^{|V|} Q_{k,w}(w_i) \log Q_{k,w}(w_i) / P_{k,w}(w_i) \quad (16)$$

$$KL(Q_{k,w} \| P_{k,w}) = (KL(Q_{k,w} \| P_{k,w}) + KL(P_{k,w} \| Q_{k,w})) / 2 \quad (17)$$

where $Q_{k,w}$ and $P_{k,w}$ are the probability distributions of topic-words within different time units. Taking the Data_c dataset as the input object of TI-EBTM to obtain the topic-word matrix ϕ_{kw} . Each row of the matrix represents the semantic feature concepts of each topic. We employed TI-EBTM to calculate the matrices in different time units. Then we used KL to quantitatively compare the similarity between topics, and the distance threshold was set as 0.6. Combined with the qualitative analysis method of human observation, the dynamic evolution of topic content in different intervals of the semester was depicted.

As shown in Tables 5 and 6, we constructed an example of the semantic evolution of topic content about the application of data structure (topic 9) and programming implementation (topic 14) in different periods of the semester. Through observing the corresponding 10 words closely related to topic 9 and topic 14 within the different time units, students expressed similar discussion content regarding the application of data structure and programming implementation as a whole. Specifically, in terms of topic 9, although some differences were shown in the order of semantic characteristics, these topics' words were particularly associated with each other from the 11th week to the 16th week. This indicates that the topic continuously evolved during this period. With

Table 5 The evolution of topic content about the application of data structure over time

Course	Topic Label	Time Unit 6 11–12 week	Time Unit 7 13–14 week	Time Unit 8 15–16 week
Data Structure	Application of data structure (Topic 9)	信息/ information	信息/information	信息/information
		结构/structure	结构/structure	技术/ technology
		数据/data	数据/data	系统/system
		课程/course	信息化/ informatization	信息化/ informatization
		系统/system	人才/talents	人才/talents
		计算机/ computer	技术/technology	结构/structure
		组织/organize	计算机/computer	管理/manage
		管理/manage	系统/system	课程/course
		服务/service	管理/manage	计算机/computer
		应用/ application	课程/course	数据/data

Table 6 The evolution of topic content about programming implementation over time

Course	Topic Label	Time Unit 5 9–10 week	Time Unit 7 13–14 week	Time Unit 8 15–16 week	Time Unit 9 17 week
Data Structure	Programming implementation (Topic 14)	串/string	串/string	空间/space.	位置/index
		位置/index	语言/language	分配/allocation	出现/appear
		出现/appear	出现/appear	串/string	第一次/first time
		第一次/first time	位置/index	长度/length	字符/character
		理解/understand	字符/character	size	串/string
		字符/character	函数/function	元素/element	理解/understand
		代码/code	java	length	线性/linear
		过程/process	第一次/first time	执行/execute	代码/code
		锻炼/exercise	输入/input	定义/definition	编写/compile
		找到/find	变量/variable	连续/succession	表达式/expression

respect to topic 14, students preferred to use some new terms from weeks 15 to 16. This demonstrates that students' focuses on the topic content have changed to some extent. Moreover, some examples of posts for topics 9 and 14 are presented: for topic 9, "Data structure is the core course of computers and other subjects. Only by mastering data structure can we learn future courses better. Today is an information society, and we must master relevant knowledge about the computer in order to better grasp information", and for topic 14, "How can we understand and apply the space allocation code of strings, and what functions are included in it?" In this case, the analytic results can inform instructors better insights into students' various focuses around the same topic across the course. Subsequently, instructors can guide and adjust students' discussion according to the evolutionary changes of topic content over the duration of the course.

Discussion

This research aimed to detect and track students' post-based dynamic meaning by using an unsupervised topic model called TI-EBTM. Some findings regarding the evolution of topic intensity and topic content over the duration of the course are worth of further discussion.

The evolutionary intensity of the topics that students were concerned about showed a certain similarity over time. That is, the topic intensity fluctuated greatly, as it rose suddenly and then declined rapidly. Moreover, most of the topics' probabilities reached their maximum value in succession and presented some continuity within the specific time period, especially in the next stage of the semester. The possible reasons are that the topics discussed by students depend on the arrangement of knowledge points in the process of course teaching and the presentation of knowledge points follows the small-step teaching method; thus, the discussion of different topics is limited by the specific arrangement of courses. According to the setting mechanism of the topic interaction in SPOC discussion forums, the discussion of a specific topic is usually limited to a specific time period (Fox, 2013). That is, the topics in the corresponding time units are similar, so the topic intensity increases in certain specific time segments. Furthermore, at the end of the semester, the intensity of the topics that students were concerned about declined rapidly. This may be because students need to prepare for the

examination offline and have little time to interact with the course discussions (Ramesh & Getoor, 2018).

As a whole, students tend to positively evaluate specialized knowledge over the duration of the course. This finding is similar to a previous study that addressed learners focused more on course-related content with positive sentiment in MOOC course reviews (Liu et al., 2019, a). Moreover, some studies have indicated that students who discussed topics associated with course content achieved better academic performance (Peng & Xu, 2020; Ramesh et al., 2014). A particular case is that the topic about method of data query in the 17th week has a higher negative probability, close to 0.6. This suggests that students might have difficulty in understanding the method of data query and consequently suffer from a bad learning experience. Thus, instructors should pay attention to the emotional changes of the topic and apply timely topic guidance and emotional feedback. In addition, the temporal topics about the application of data structure and programming implementation are dominated by RE and CP respectively that are generally regarded as passive behavior interactions in online discussions. In contrast, the dynamic topics about the method of query and the method of storage are dominated by TP, which is commonly considered a negative behavior interaction in online discussions (Tobarra et al., 2014). In this case, due to the lack of professional knowledge, students might take the initiative by posting to seek social help in specific time segments.

As the course progressed, topic content that students paid attention to would change to some extent. Perhaps the reason is that when students participate in the same topic discussion, some relevant local topics are generated and extended (Blei & Lafferty, 2006; Ramesh et al., 2015). Thus, to capture various aspects of topics, sentences might be a better choice than documents in modeling dynamic topics.

Conclusions, implications, and future work

A large number of student-generated posts in online learning communities pose an overwhelming workload for instructors to navigate and locate information. Therefore, an automatic semantic analysis for capturing and tracking students' discussions should be exploited. In this study, we leveraged an improved dynamic topic model, TI-EBTM, that characterizes time, emotion, and behavior features to automatically identify the evolution of students' focuses over time. TI-EBTM could be embedded into intelligent program applications, potentially enhancing teaching in practical learning technologies. It could also be generalized to other educational contexts, such as MOOC, LMS, and ITS. The experimental results demonstrated that compared with other dynamic topic models, TI-EBTM could achieve better performance in discovering higher-quality temporal topics. Additionally, TI-EBTM could effectively reflect students' focused topic changes including topic intensity and topic content, as well as the evolution of topics' emotional and behavioral distributions over the duration of the course. Therefore, our methodology and analysis are useful to reliably track the stability of a course and detect students' focused topics or potential issues (e.g., the absence of professional knowledge) over time. Subsequently, instructors can fine-tune the discussion activities and teaching arrangements for the next course offering.

Practically, this study constitutes a step toward employing temporal topic analysis that can be incorporated into an adaptive feedback mechanism for asynchronous

communication, which enables massive-scale automated discourse analysis to enhance the quality of course interaction and to better support students' learning experience. Moreover, tracking student-generated posts in SPOC forums holds the potential to support instructors in gaining insights into an overview of course development and in providing adaptive feedback to students. The analytic results demonstrated that the intensity of students' attention to the topics fluctuated obviously, and the topic content involving some detailed concepts evolved over the duration of the course. Therefore, instructors can reorganize the upcoming topic discussion activities and enable interventions for community management. For example, in the initial stage of the semester, instructors should actively interact with students or organize different partners in the topic discussions to stimulate and strengthen students' enthusiasm for participation, allowing for the promotion of more in-depth interactive activities. Taking topic 16 about the method of data query as an example, the results demonstrated that students tended to express more negative emotions toward topic 16 by TP at the end of the semester. In this way, when negative emotions and active behavioral orientations in relation to this type of topic are displayed during specific time stamps, instructors should judge the discussion content to understand the possible reasons for its occurrence and adopt a positive topic guidance strategy for the maintenance of a healthy community. As Wang and Zhu (2019) noted, negative emotions are directly related to students' course survival, so instructors should provide effective guidance and offer help to students who are in need. Additionally, an adjustable, dynamic dashboard should be built to display and visualize course-related discussions as a network graph of relationships between topics and time periods, emotions, and behaviors (Vytasek, Wise, & Woloshen, 2017). This would enable instructors to navigate information and prompt instructors regarding when and how to intervene in discussion forums.

This study has some limitations. For example, the validation of model performance should not be limited to the educational environment but should be verified on other domain databases. Differences in disciplines might interfere with the effect of empirical experiment application (He, 2013). Therefore, in future work, there is a need to enrich data sources to conduct model generalization and practice. Further research should be conducted to investigate the differences of dynamic topic intensity and topic content between arts students and science students. In addition, integrating dynamic topic model technology into actual teaching tools is helpful for educational practitioners to intuitively use and evaluate the evolution of course forums.

Acknowledgements

Not applicable.

Authors' contributions

PX proposed an improved unsupervised dynamic topic model, called Time Information-Emotion Behavior Model (TI-EBTM), and analyzed the SPOC courses data regarding the dynamic evolution of forum posts. He was also a major contributor in writing the manuscript. HCY, OYF and LZ gave suggestions on adjusting the structure of the article and made some grammar corrections of the article. All authors read and approved the final manuscript.

Funding

This work is funded by National Natural Science Foundation of China (61702207).

Availability of data and materials

The data collected in the current study are not publicly available since they were retrieved under students' authorization and anonymity as well as permission of relevant departments of the university.

Competing interests

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

Author details

¹College of Education, Zhejiang University, Hangzhou 310028, Zhejiang, China. ²School of Psychology, Shenzhen University, Shenzhen 518061, Guangdong, China. ³National Research and Engineering Center for E-learning, Central China Normal University, Wuhan 430079, Hubei, China.

Received: 11 February 2020 Accepted: 4 May 2020

Published online: 02 September 2020

References

- Almatrafi, O., & Johri, A. (2018). Systematic review of discussion forums in massive open online courses (MOOCs). *IEEE Transactions on Learning Technologies*, 12(3), 413–428.
- Andrei, V., & Arandjelović, O. (2016). Complex temporal topic evolution modelling using the Kullback-Leibler divergence and the Bhattacharyya distance. *EURASIP Journal on Bioinformatics and Systems Biology*, 2016(1), 16–32.
- Blei, D. M., & Lafferty, J. D. (2006, June). Dynamic topic models. In *Proceedings of the 23rd international conference on Machine learning* (pp. 113–120).
- Blei, D. M., & Lafferty, J. D. (2007). A correlated topic model of science. *The Annals of Applied Statistics*, 1(1), 17–35.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3(1), 993–1022.
- Chen, W., Brinton, C. G., Cao, D., Mason-Singh, A., Lu, C., & Chiang, M. (2018). Early detection prediction of learning outcomes in online short-courses via learning behaviors. *IEEE Transactions on Learning Technologies*, 12(1), 44–58.
- Chiu, K. F. T., & Hew, K. F. T. (2018). Factors influencing peer learning and performance in MOOC asynchronous online discussion forum. *Australasian Journal of Educational Technology*, 34(4), 16–28.
- Combéfis, S., Bibal, A., & Van Roy, P. (2014). Recasting a traditional course into a MOOC by means of a SPOC. In *Proceedings of the European MOOCs Stakeholders Summit*, (pp. 205–208).
- Dermouche, M., Velcin, J., Khouas, L., & Loudcher, S. (2014, December). A joint model for topic-sentiment evolution over time. In *2014 IEEE International Conference on Data Mining* (pp. 773–778).
- Dong, Z. D. (2013). HowNet's HomePage. Retrieved from <http://www.keenage.com>.
- Dupuy, C., Bach, F., & Diot, C. (2017, July). Qualitative and descriptive topic extraction from movie reviews using lda. In *International Conference on Machine Learning and Data Mining in Pattern Recognition* (pp. 91–106). Springer, Cham.
- Elgort, I., Lundqvist, K., McDonald, J., & Moskal, A. C. M. (2018, March). Analysis of student discussion posts in a MOOC: Proof of concept. In *Companion Proceedings 8th International Conference on Learning Analytics & Knowledge (LAK18)* (pp. 1–7).
- Ezen-Can, A., Boyer, K. E., Kellogg, S., & Booth, S. (2015, March). Unsupervised modeling for understanding MOOC discussion forums: A learning analytics approach. In *Proceedings of the fifth international conference on learning analytics and knowledge* (pp. 146–150).
- Filius, R. M., de Kleijn, R. A., Uijl, S. G., Prins, F. J., van Rijen, H. V., & Grobbee, D. E. (2018). Strengthening dialogic peer feedback aiming for deep learning in SPOCs. *Computers & Education*, 125, 86–100.
- Fox, A. (2013). From MOOCs to SPOCs. *Communications of the ACM*, 56(12), 38–40.
- Freitas, A., & Paredes, J. (2018). Understanding the faculty perspectives influencing their innovative practices in MOOCs/SPOCs: A case study. *International Journal of Educational Technology in Higher Education*, 15(1), 5.
- Garroppo, R. G., Ahmed, M., Niccolini, S., & Dusi, M. (2018). A vocabulary for growth: Topic modeling of content popularity evolution. *IEEE Transactions on Multimedia*, 20(10), 2683–2692.
- Gitinabard, N., Heckman, S., Barnes, T., & Lynch, C. F. (2019). What will you do next? A sequence analysis on the student transitions between online platforms in blended courses. *arXiv preprint arXiv:1905.00928*.
- Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences*, 101(suppl 1), 5228–5235.
- He, W. (2013). Examining students' online interaction in a live video streaming environment using data mining and text mining. *Computers in Human Behavior*, 29(1), 90–102.
- He, Y., Lin, C., Gao, W., & Wong, K. F. (2014). Dynamic joint sentiment-topic model. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 5(1), 1–21.
- Ku, L. W., Liang, Y. T., & Chen, H. H. (2006). Opinion extraction, summarization and tracking in news and blog corpora. In *Proceedings of the 21st National Conference on Artificial Intelligence*, (pp. 100–107).
- Li, J. (2011). Chinese derogatory dictionary v1.0. Retrieved from <http://nlp.csai.tsinghua.edu.cn/site2/index.php/zh/resources/13-v10>.
- Liu, S., Peng, X., Cheng, H. N., Liu, Z., Sun, J., & Yang, C. (2019). Unfolding sentimental and behavioral tendencies of learners' concerned topics from course reviews in a MOOC. *Journal of Educational Computing Research*, 57(3), 670–696.
- Liu, Z., Yang, C., Peng, X., Sun, J., & Liu, S. (2017, December). Joint exploration of negative academic emotion and topics in student-generated online course comments. In *2017 International Conference of Educational Innovation through Technology (EITT)* (pp. 89–93).
- Liu, Z., Yang, C., Rüdian, S., Liu, S., Zhao, L., & Wang, T. (2019). Temporal emotion-aspect modeling for discovering what students are concerned about in online course forums. *Interactive Learning Environments*, 27(6), 598–627.
- Mo, Y., Kontonatsios, G., & Ananiadou, S. (2015). Supporting systematic reviews using LDA-based document representations. *Systematic Reviews*, 4(1), 172–185.
- Moreno-Marcos, P. M., Alario-Hoyos, C., Muñoz-Merino, P. J., & Kloos, C. D. (2018). Prediction in MOOCs: A review and future research directions. *IEEE Transactions on Learning Technologies*, 12(3), 384–401.
- Peng, X., & Xu, Q. (2020). Investigating learners' behaviors and discourse content in MOOC course reviews. *Computers & Education*, 143(1), 1–14.

- Phan, T., McNeil, S. G., & Robin, B. R. (2016). Students' patterns of engagement and course performance in a massive open online course. *Computers & Education*, 95, 36–44.
- Ramesh, A., & Getoor, L. (2018, November). Topic evolution models for long-running MOOCs. In *International Conference on Web Information Systems Engineering* (pp. 410–421). Springer, Cham.
- Ramesh, A., Goldwasser, D., Huang, B., Daume, H., & Getoor, L. (2014, June). Understanding MOOC discussion forums using seeded LDA. In *Proceedings of the ninth workshop on innovative use of NLP for building educational applications* (pp. 28–33).
- Ramesh, A., Kumar, S. H., Foulds, J., & Getoor, L. (2015, July). Weakly supervised models of aspect-sentiment for online course discussion forums. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing* (volume 1: Long papers) (pp. 74–83).
- Reyes-Menendez, A., Saura, J., & Alvarez-Alonso, C. (2018). Understanding world environment day user opinions in twitter: A topic-based sentiment analysis approach. *International Journal of Environmental Research and Public Health*, 15(11), 2537.
- Rosen-Zvi, M., Griffiths, T., Steyvers, M., & Smyth, P. (2004, July). The author-topic model for authors and documents. In *Proceedings of the 20th conference on Uncertainty in artificial intelligence* (pp. 487–494).
- Rossetti, M., Stella, F., & Zanker, M. (2016). Analyzing user reviews in tourism with topic models. *Information Technology & Tourism*, 16(1), 5–21.
- Steyvers, M., & Griffiths, T. (2007). Probabilistic topic models. *Handbook of Latent Semantic Analysis*, 427(7), 424–440.
- 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.
- Vytasek, J. M., Wise, A. F., & Woloshen, S. (2017, March). Topic models to support instructors in MOOC forums. In *Proceedings of the seventh international learning analytics & knowledge conference* (pp. 610–611).
- Wang, C., Fang, T., & Gu, Y. (2020). Learning performance and behavioral patterns of online collaborative learning: Impact of cognitive load and affordances of different multimedia. *Computers & Education*, 143(1), 103683.
- Wang, K., & Zhu, C. (2019). MOOC-based flipped learning in higher education: Students' participation, experience and learning performance. *International Journal of Educational Technology in Higher Education*, 16(1), 33.
- Wang, X., & McCallum, A. (2006, August). Topics over time: A non-Markov continuous-time model of topical trends. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 424–433).
- Wayne, C. L. (1997, October). Topic detection and tracking (TDT). In *Workshop held at the University of Maryland* (pp. 28–30).
- 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 (EDM 2014)* (pp. 1–8).
- Westerlund, M., Mahmood, Z., Leminen, S., & Rajahonka, M. (2019). Topic modelling analysis of online reviews: Indian restaurants at Amazon. Com. In *Proceedings of the International Society for Professional Innovation Management (ISPIM)* (pp. 1–14).
- Wong, A. W., Wong, K., & Hindle, A. (2019). Tracing forum posts to MOOC content using topic analysis. arXiv preprint arXiv:1904.07307.
- Xie, W., Zhu, F., Jiang, J., Lim, E. P., & Wang, K. (2016). Topicsketch: Real-time bursty topic detection from twitter. *IEEE Transactions on Knowledge and Data Engineering*, 28(8), 2216–2229.
- Xu, Y., & Lynch, C. F. (2018). What do you want? Applying deep learning models to detect question topics in MOOC forum posts? In *Wood-stock'18: ACM Symposium on Neural Gaze Detection*, (pp. 1–6).
- Zhao, Z., Cheng, Z., Hong, L., & Chi, E. H. (2015, May). Improving user topic interest profiles by behavior factorization. In *Proceedings of the 24th International Conference on World Wide Web* (pp. 1406–1416).

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)