Yugoslav Journal of Operations Research
34 (2024), Number 4, 707-724
DOI: https://doi.org/10.2298/YJOR2403029L

TEACHER PROFESSIONAL LEARNING COMMUNITY AND INTERDISCIPLINARY COLLABORATIVE TEACHING PATH UNDER THE INFORMATIONIZATION BASIC EDUCATION MODEL

Gen LI

School of Educational Studies, Universiti Sains Malaysia, Penang.11800, Malaysia
hazri@usm.my

Hazri Bin JAMIL*

School of Educational Studies, Universiti Sains Malaysia, Penang.11800, Malaysia
hazribinjamil@163.com

Received: March 2024 / Accepted: July 2024

Abstract: The construction of a learning community cannot be separated from the participation of information technology. The current teacher learning community has problems of low interaction efficiency and insufficient enthusiasm for group cooperative teaching. This study adopts the Latent Dirichlet allocation method to process text data generated by teacher interaction from the evolution of knowledge topics in the learning community network space. At the same time, the interaction data of the network community learning space is used to extract the interaction characteristics between teachers, and a collaborative teaching group is formed using the K-means clustering algorithm. This study verifies the management effect of Latent Dirichlet allocation and K-means algorithm in learning community space through experiments. The experiment showed that the Latent Dirichlet allocation algorithm had the highest F1 value at a K value of 12, which is 0.88. It collaborated with the filtering algorithm on the overall F1 value. At the same time, there were a total of 4 samples with incorrect judgments in Latent Dirichlet allocation, with an accuracy of 86.7%, which is higher than other algorithm models. The results indicate that the proposed Latent Dirichlet allocation combined with K-means algorithm has superior performance in the management of teacher professional learning communities, and can effectively improve the service level of teacher work.

Keywords: Professional learning community**,** interdisciolinary collaborative teaching, knowledge topics, grouping, potential Dirichlet distribution, interactive.

MSC: 97D99.

* Corresponding author

1. INTRODUCTION

With the further updates and iterations of big data and internet technology, more and more social production activities begin to rely on information technology. Education, as the core field of current technological reform, is constantly deepening in applications such as artificial intelligence and information technology. Expanding educational channels and improving teaching quality have become innovative hotspots in the current education field. The improvement of current teacher quality is an important way for educational reform. To enhance the professional competence of teachers in teaching and improve their level, domestic and foreign universities have improved the construction practice of learning communities by building educational resource sharing platforms and teacher exchange and cooperation platforms. However, there are difficulties in data management and vague information exchange among teachers in the construction of the current online learning community space, resulting in unscientific grouping of cooperative teaching among teachers. Therefore, this study will explore the possibility of upgrading and optimizing the management of the teacher learning community service platform system by optimizing the Latent Dirichlet allocation (LDA) model and clustering algorithms. The purpose of the paper is to improve the working efficiency of the Internet+teaching model through information technology, and reduce the time teachers spend in the trivial work of network teaching management. The first chapter of this study is a literature review, which provides a comprehensive analysis of the information technology education theme and LDA model studied. The second chapter discusses the research methods and explores the application of LDA topic model in the learning community space, as well as the teacher collaboration grouping method. The third chapter is experimental analysis, which verifies the proposed model method from two aspects: dataset simulation and practical teaching application. The fourth chapter concludes the research.

2. RELATED WORK

With the innovative application of Internet technology, the integration of education and information technology has become increasingly close. The construction of a professional learning community has become a breakthrough point in current cooperative teaching. Matthew and his team analyzed the impact of learning communities on student performance from the perspective of professional learning communities. The results showed that the attributes of learning communities, collaborative leadership processes, and information-based learning systems are all related to student performance. Therefore, reforming and optimizing the professional learning community in the above aspects can effectively enhance students' learning enthusiasm and efficiency [1]. Do ğ An et al. analyzed the function of professional learning communities from the perspectives of both teachers and students. At the same time, this study also raised concerns about the current situation of learning communities, stating that there are methodological issues in community learning that limit teachers' ability to interpret results [2]. Schaap et al. observed the professional learning communities of different universities through questionnaire surveys and participatory studies. The results indicate that task perception, group composition, tense relationships between roles, beliefs about alliances, reflective dialogue, socialization, and ownership are the seven core elements that affect the construction of professional learning communities [3]. Rilana Prenger and her team believe that a teacher learning community is the key to teacher improvement. Over the course of 10 years, the author observed that some within school learning communities have transformed into inter school communities, where teachers have improved their knowledge, skills, attitudes, and other professional qualities [4]. Admiraa et al. analyzed the optimization path for building a teacher learning community from five perspectives: construction intention, learning opportunities, cooperative learning methods, organizational form, and leadership, providing sustainable cultural support for cooperative learning [5]. Chauraya et al. believe that the way data records teacher conversations can help improve collaborative learning efficiency and help analyze the thinking behind learner errors. Meanwhile, through this approach, teachers have the opportunity to learn how to identify the learning needs of learners, and thus identify their own learning needs [6].

Jelodar H et al. summarized the current application status of LDA models in data mining, data discovery, and document text search, and analyzed the development of LDA in current fields such as software engineering, political science, and linguistics. The author reviewed some highly academic LDA papers from 2003 to 2016 and analyzed their model construction and knowledge structure [7]. Guo C et al. analyzed the semantic relationship distribution in multi word text based on the LDA model and proposed an improved partitioned long text LDA topic model. In the experiment, this model had better performance compared to other models [8]. Maier D et al. analyzed the application of LDA models in the field of communication and summarized the problems of LDA. In the study, the author proposed four perspectives: text set preprocessing, topic quantity selection, reliability, and generation process, and developed a practical user guide for the LDA topic model [9]. Bastani K et al. proposed an intelligent analysis method using the LDA model for complaint texts in the Consumer Financial Protection Agency. This model could effectively improve the consumer complaint experience [10]. Yang S et al. proposed using LDA model for topic analysis of network interaction platforms. Taking Twitter as an example, the author aimed to construct a tweet topic recognition method that can provide emotional analysis. In English Twitter text analysis, the LDA model had effective data mining capabilities [11]. Yuan C et al. proposed a computational method for K-means clustering algorithm (K-means) and analyzed the advantages and disadvantages of four methods: Elbow method, Gap statistic, Silhouette coefficient, and Canopy in their study [12]. Zou H proposed a clustering analysis algorithm based on the role of clustering analysis in data mining and applied it to data mining. After literature comparison and analysis, the author explored the implementation process and instance simulation of K-means, providing a theoretical basis for clustering algorithms to detect and analyze large amounts of data in a timely manner [13].

In summary, the current development of online education is becoming increasingly widespread, but there is a situation where technical practice cannot keep up with theoretical expansion in the management of online teaching. In the current construction of educational community space, the factors of information technology have not been applied in a valuable way. Therefore, this study explores methods for optimizing teacher learning community management based on the LDA model and clustering algorithm.

3. MANAGEMENT OPTIMIZATION METHODS FOR TPLC SPACE

The Teacher Professional Learning Community (TPLC) is a training community space established by teachers through the internet. In the learning community, teachers can independently discuss, interact, learn from each other, and form small groups for cooperative teaching. At the same time, teachers can also explore the knowledge content of relevant subjects through learning communities, construct a public knowledge teaching system, and solve problems in teaching work. Therefore, dynamic monitoring and tracking of knowledge topics within the learning community can help to incorporate information technology into the development of collaborative teaching grouping work, which is also the focus of this study.

3.1. LDA based analysis method for the evolution of spatial knowledge topics in TPLC

The development and construction of teacher professional learning communities cannot be separated from the guidance of relevant teaching theories. Currently, most researchers have designed teacher professional learning community activities from knowledge construction theory, knowledge transfer theory, and activity theory. This study analyzes and evaluates the process and effectiveness of training activities based on previous methods, and proposes corresponding learning assistance strategies for different stages of training activities. As adult learners, teachers are more concerned with solving practical problems related to their profession and acquiring practical knowledge. The practical knowledge of teachers has characteristics such as silence, reflection, and situational nature, and there are certain difficulties in transferring from implicit knowledge to explicit knowledge. Based on this study, the improvement methods of informatization were elaborated from two aspects: knowledge classification and collaborative teaching grouping.

In the interactive dialogue process of TPLC network space, topics can be seen as a manifestation of teacher groups solving teaching problems under specific purpose guidance. The process and inherent rules of the evolution of knowledge topics are the core issues of collaborative knowledge construction research, essentially an analysis of the co creation and generation mechanism of "knowledge". The study of the evolution of knowledge topics within the learning community space can better grasp the operational rules of knowledge in teaching cooperation, and based on this, carry out targeted teaching cooperation grouping work [14-15]. By digitizing topics, teachers can acquire new knowledge on the basis of their existing knowledge. The emergence of knowledge topics brings the process of education closer to the predetermined goals, ultimately achieving educational objectives. However, the dynamic and uncertain characteristics of knowledge topics themselves, as well as the large number of participants in cyberspace interactions, the intersection and mutual influence between content, pose serious challenges to topic mining and evolution analysis. The learning community interaction space between individual knowledge topics and group knowledge topics of teachers is Figure 1.

On the basis of a learning community, this study establishes a research method based on the evolution of knowledge topics. Starting from the complete process of data acquisition, topic mining, decomposition, and application of results, a research method based on topic evolution is proposed, and corresponding research strategies are proposed.

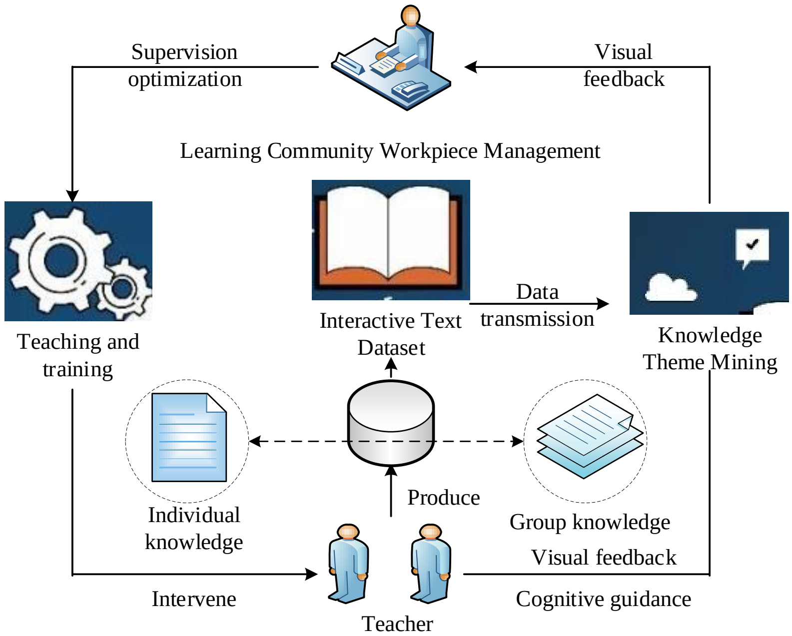

Figure 1: Interactive Space for Teacher Learning Community

Data collection refers to the information collected by participating teachers through collaborative discussions during the online teaching and research process, generating a massive amount of interactive text data. These pieces of information are divided into time windows according to the teacher's work hours, and then segmented into different time periods based on the publication date of the communication post. By preprocessing these documents, a usable interactive text library can be obtained. Thus, LDA models can be used to mine and identify knowledge topics in different time periods. The LDA model is given on Figure 2.

In the LDA model, when identifying knowledge topics for comment posts in the time window, the number of topics to be identified must be set first. If the value is too large, it will lead to identifying too many topics, reducing the dispersion of topics. If the number of topics is too small, it is difficult to reflect the systematic nature of the knowledge framework [16]. Therefore, this study used the confusion index to calculate the number of topics to obtain appropriate K values for each time period, as shown in equation (1).

$$D = \exp\left(-\sum_{x=1}^{X} \log p(w_x) / \sum_{x=1}^{X} Y_x\right) \quad (1)$$

In equation (1), Drepresents the degree of confusion. X represents the total number of posts in the learning community space. x represents the post number. Y represents the number of vocabularies in the document, and $p(w_x)$ represents the probability of vocabulary appearing in the post. After obtaining the degree of confusion under different numbers of topics and comparing them, the optimal number of topics can be determined. After determining the number of topics, calculate the similarity of topics. The research on the evolution of thematic content is an analysis of the temporal evolution of educational curriculum themes. This study will establish the evolutionary relationship between topics by statistically analyzing their similarity. In the process of calculating dispersion, due to the fact that the two topics a, b come from different posts and are not within the same time window. Therefore, according to the sampling estimation method in the LDA model, its formula is equation (2).

$$p(w|T_i^t) = \lambda/(u_T + \lambda W). \quad (2)$$

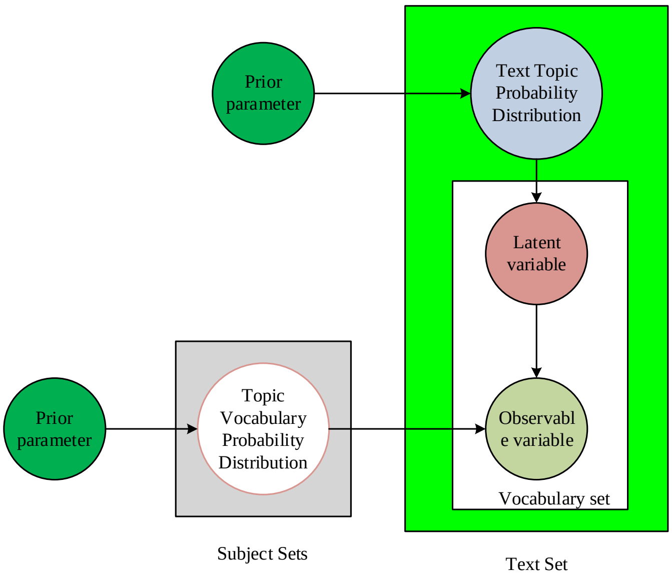

Figure 2: LDA model

In equation (2), w represents vocabulary. t represents the time period. T represents the theme. λ is the prior parameter of the LDA model. u represents the total number of vocabularies included in the topic. W represents the union of word lists between two adjacent time periods. To identify the relevance of the subjects, this study used Relative Entropy (RE) for similarity calculation, and its formula is equation (3).

$$RE(a, b) = \sum_{i=1}^{I} a(m_i)\log\frac{a(m_i)}{b(m_i)} \quad (3)$$

In equation (3), $RE(a, b)$ represents the RE between topic sa, b. m represents a random discrete variable of the topic. I represents the number of random discrete variable sets for the topic. At the same time, introducing JS discretization on the basis of RE for optimization can compensate for the asymmetry of RE discretization, and its formula is equation (4).

$$JS(a, b) = \frac{RE(a, \frac{a+b}{2}) + RE(b, \frac{a+b}{2})}{2} \quad (4)$$

In equation (4), $JS(a,b)$ represents the JS dispersion between topics a,b. By combining equations (2) and (4), it is possible to calculate the similarity of subjects within adjacent time periods and determine the changes in topics related to event continuation. When the topic relevance between two time periods is less than the threshold, it indicates that there is correlation between the subject of the posts in the two time periods. According to correlation analysis, the relationship between subjects in adjacent time periods can be expressed as equation (5).

$$\begin{cases} JS(T_t, T_{t-1}) \leq \varepsilon, \text{forward} \\ JS(T_t, T_{t+1}) \leq \varepsilon, \text{backward} \end{cases} \tag{5}$$

In equation (5), $JS(T_t, T_{t-1})$ represents the correlation between the topics at time t and time $t-1$. ε represents the correlation threshold. When the correlation is less than or equal to W, it indicates that the topics of times t and $t-1$ are forward topic relationships. On the contrary, when $JS(T_t, T_{t+1}) \leq \varepsilon$, it is a backward thematic relationship. According to the relationship between themes, the evolution between themes can be classified. The detailed classification is equation (6).

$$\begin{cases} T_t \notin T_{t-1}, \text{newborn} \\ T_t \notin T_{t+1}, \text{wither away} \\ T_{t-1} \in T_t \in T_{t+1}, \text{inherit} \end{cases} \tag{6}$$

In equation (6), when there is no forward theme, it is represented as a new relationship. When there is no backward theme in the theme, it indicates an extinction relationship. When there is a forward or backward theme, it is represented as an inheritance relationship. In addition to the three relationships in equation (6), there are also splitting and merging relationships. To analyze the level of attention to a topic during a certain period of time during its evolution, this study distinguishes it by calculating the intensity of the topic, as shown in equation (7).

$$\phi_z^t = \sum_{g=1}^{G} \phi_z^g / G^t \tag{7}$$

In equation (7), ϕ_z^g represents the proportion of topic z in document g. G^t represents the text collection of all posts within time period t. ϕ_z^t represents the strength of topic z within the text set of all posts in t. The above is the knowledge topic evolution analysis method based on the LDA model. The focus of studying the changes and transitions in the chronological order of teacher seminar topics is to construct an evolutionary trajectory based on the correlation between topics near the time window, and then further analyze it. Theme intensity evolution analysis is the use of LDA algorithm to analyze the changes in themes of concern to teachers by statistically analyzing the distribution and hotspots of themes in each time period [17-18].

3.2. A grouping method for teacher collaborative teaching based on learning community

This study is based on the communication behavior in the learning community space to group teacher cooperative teaching. In online collaborative learning, teachers build and develop internet-based social relationships through sending or receiving posts. Its main manifestation is the number of posts, comments, or contacts posted or received by teachers. The dialogue and interaction between teachers are purposeful. The higher the importance of a node when it is connected to many other nodes. On this basis, the judgment of grouping is based on the number of responses received from other teachers during the cooperation period. "Importance" refers to the degree to which teachers establish connections with other teachers. The high value of "importance" indicates that teachers have a significant influence on other teachers. The node egress of a member is based on the number of messages received by each participating member in the discussion, which refers to the number of times this member interacts with other students. The larger the value, the more active the participants are [19]. The importance of individual teachers is the sum of their degree of exit and degree of entry, and its calculation formula is equation (8).

$$\begin{cases} IM_{in}(U_i) = \sum_{i=1}^{n} d_i^{in} \\ IM_{out}(U_i) = \sum_{i=1}^{n} d_i^{out} \end{cases} \tag{8}$$

(8) In, IM_{in} represents the degree of importance. IM_{out} represents the degree of importance. U_i represents the teacher with ID number i. d represents the frequency of communication between teachers and other teachers. In addition to the frequency of communication as a basis for teacher grouping, the smoothness of interaction between teachers is an important indicator. Taking the main node as an example, if the teacher is located between the shortest paths of two nodes, it indicates that the teacher has achieved optimal control over the path. The higher the frequency of teachers acting as intermediaries, the smoother the interaction between information [20-22]. The calculation formula for mediation is equation (9).

$$IN(U_i) = \sum_{j,k=1}^{g} s(i,j,k) \tag{9}$$

In equation (9), teacher i acts as an intermediary between teacher j,k. $s(i,j,k)$ indicates that the shortest path between teacher j,k passes through teacher i. The higher the value of mediation, the more important the role of the teacher as an intermediary in communication with other teachers [23-24]. Finally, the interaction intensity between teachers is sampled, and feature centrality is used as a measurement indicator to describe the teacher learner centrality within the learning community space. The calculation formula is equation (10).

$$FC(U_i) = c \sum_{j=1}^{g} v_{ij} IN_j \tag{10}$$

In equation (10), $FC(U_i)$ represents the centrality of the teacher's characteristics. c is a proportional constant. IN_j represents the mediating nature of the teacher node. When teacher i,j interacts with each other, $v_{ij} = 1$, otherwise it is 0. The grouping application of teacher network interaction information is Figure 3.

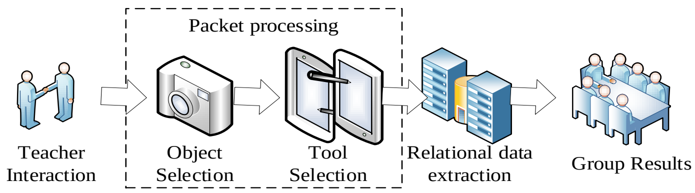

Figure 3: Grouping application of teacher network interactive information

After obtaining network interaction indicators describing teacher participation in the learning community space, teachers were grouped using K-means. Assuming that the teacher in the learning community is defined as $\{t_1, t_2, \ldots, t_n\}$. The importance, intermediary, and feature centrality of teacher interaction is added together as the similarity of teacher interaction, and the similarity is extracted and aggregated into a subset of l. Each subset is a small group, and the formula to determine the clustering effect is equation (11).

$$f_l = \sum_{j=1}^{l} \sum_{k=1}^{M_j} \left\| t_k - m_j \right\|^2 \tag{11}$$

In equation (11), j represents the center of the group, f_n is the function of the data sample and the set center, and m_j is the mean sample in the cluster. The calculation is equation (12).

$$m_j = \frac{1}{n} \sum_{j=1}^{n_j} t_j \tag{12}$$

In equation (12), n represents the quantity of teachers. The smaller the sum of squared errors calculated by the clustering function, the better the clustering effect. The initial mean is set to $k = 2$, assuming that the gross teachers combined is N and the group is divided into $C = \{c_1, \ldots, c_k\}$, the expected value for group division is described as follows (13).

$$P(n_1, n_2, \ldots n_k) = -\sum_{i=1}^{x} n_i/N * \log_2(n_i/N) \tag{13}$$

In equation (13), P represents the expected classification, and n_i represents the number of teachers in category c_i. After obtaining the clustering, the mean points of each group are calculated as the new center, the clustering process is repeated until the cluster center fixed, as displayed in equation (14).

$$E = \frac{\sum_{i=1}^{m}(s_i - Ms)}{m} \tag{14}$$

In equation (14), E is the average error sum of teacher i's group. m is the number of teachers in i's group. s_i is the similarity between i and the group center. Ms is each group's average similarity. If the threshold e is set, the dynamic change of k is expressed as equation (15).

$$k = \begin{cases} k + 1, E > e \\ k, E \leq e \end{cases} \tag{15}$$

Equation (15) represents the condition for the algorithm mean change. When the sum of category average errors (CAE) $>e$, resetting the initial mean, repeating the clustering division and similarity calculation within clusters until the CAE sum $<e$. Based on the LDA knowledge topic mining and K-means group classification method mentioned above, the optimization process of TPLC space is Figure 4.

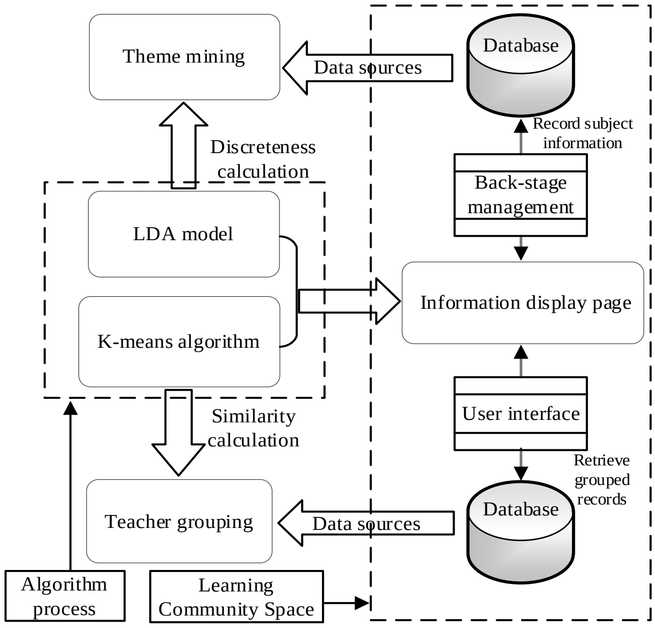

Figure 4: Optimization process of teacher learning community space

In Figure 4, the new TPLC space mainly includes knowledge topic recognition function and collaborative teaching group division function. The topic recognition function relies on the LDA model to calculate the knowledge topic dispersion and extract useful information based on the strength of the topic [25]. The division of cooperative teaching groups relies on K-means, and dynamic k-value optimization is designed to improve the accuracy of teacher group division and provide users with a good retrieval service experience [26]. In the context of a teacher professional learning community, teachers visualize their practical knowledge through dialogue, sharing, and explanation, contributing to the organization. Contribute knowledge to the community, intentionally creating and enhancing valuable knowledge for the organization or community. On the basis of each other's thoughts, teachers can jointly construct a new understanding that may not have existed before they met. At the same time, it is emphasized that teachers can transform practical knowledge through personal reflection, thereby promoting the improvement of personal knowledge and abilities [27]. In the professional learning community of teachers, teachers are both creators and consumers of knowledge. The training activities in teacher workshops have the characteristic of collaborative knowledge construction, and the theory of collaborative knowledge construction has important guiding significance for the design, implementation, analysis, and evaluation of collaborative training activities in teacher professional learning. Therefore, the study adopted an improved LDA model and clustering algorithm to guide the classification of cooperative teaching activities [28].

4. EXPERIMENTAL ANALYSIS OF TPLC SPACE MANAGEMENT OPTIMIZATION

To verify the feasibility of the TPLC spatial optimization management method constructed in this study, dataset simulation and practical application experiments were used to analyze the TPLC spatial optimization management method. The experimental section is mainly divided into two sections. The first section focuses on the classification effect of knowledge topics in the community space, and the second section focuses on the grouping ability of teachers in cooperative teaching.

4.1. Simulation analysis of TPLC spatial knowledge topic mining method

This experiment uses data texts from the professional learning community space of a primary and secondary school teacher as a sample for exploration. After obtaining information through web crawler tools and removing duplicate text content through data cleaning, a network learning community spatial interactive text dataset was obtained, as shown in Table 1.

Table 1: Learning community spatial interactive text dataset

Project	Controller	Teacher	Number of posts	Total Characters	Average number of posts	Maximum Number of text characters	Minimum Text Characters	Average number of characters in text
Data set	2	74	1544	94566	20.9	576	2	61.2
Time window	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8
Number of posts	242	267	147	158	174	198	214	144
Proportion	16%	17%	10%	10%	11%	13%	14%	9%

In Table 1, the text preprocessing tool used in this experiment is the JIEBA word segmentation system. In the recognition of knowledge topics in the learning space, the F1 value of the recognition process was calculated in this experiment. The larger the F1 value, the better the recognition results of topic mining. Table 2 shows the specific experimental results.

Table 2: F1 Comparison of different algorithms

Number of topics	1	2	3	4	5	6	7	8	9	10
LDA	0.054	0.309	0.422	0.585	0.666	0.741	0.821	0.830	0.801	0.840
CMF	0.014	0.232	0.383	0.512	0.610	0.646	0.738	0.791	0.829	0.837
Number of topics	11	12	13	14	15	16	17	18	19	20
LDA	0.871	0.880	0.862	0.857	0.832	0.826	0.817	0.803	0.797	0.789
CMF	0.850	0.837	0.812	0.798	0.795	0.740	0.730	0.725	0.707	0.663

LDA in Table 2 represents the method used in this study, while CMF represents the collaborative filtering method. The CMF algorithm reaches its maximum F1 value of 0.85 when the K value is 11. The LDA algorithm has the highest F1 value of 0.88 when the K value is 12. The data shows that LDA has superiority in overall F1 value. This study extracted three main teaching themes by calculating the intensity of themes within the learning community space. The three themes of student motivation, homework effectiveness, and classroom effectiveness are represented by T1, T2, and T3, respectively. The intensity evolution of the three themes and teacher interaction are shown in Figure 5.

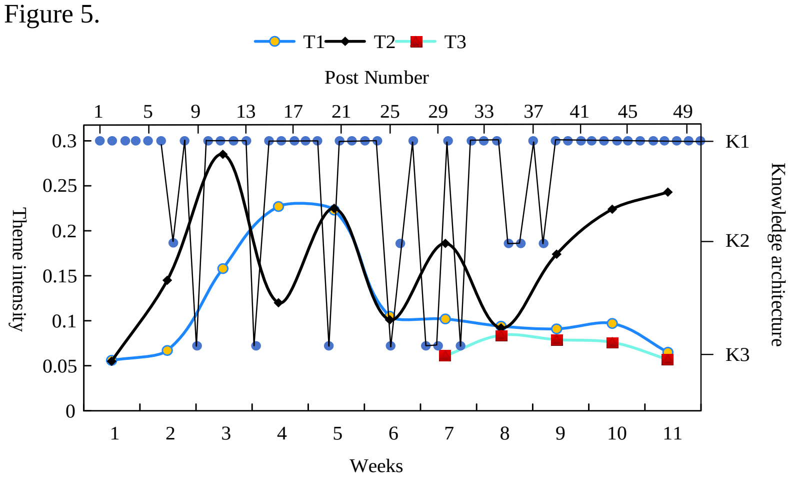

Figure 5: Evolution of different themes

In Figure 5, K1, K2, and K3 respectively represent the teacher's sharing behavior, discussion behavior, and negotiation behavior during the interaction process. The teacher's discussion on student motivation was more frequent 6 weeks ago, with a maximum theme intensity of 0.227. The discussion on the topic of homework effectiveness has been at a high frequency for a long time, with a maximum intensity of 0.285. In the learning community space, there are many discussions on topics in the early stages, indicating that there are many differences in views among teachers. In the later stages of the topic, there is a significant amount of sharing behavior, indicating that most teachers have achieved a coordinated and unified perspective through the learning space.

Finally, 30 teacher comment text data were randomly selected from the test sample dataset and fed into two models, LDA and Support Vector Machine (SVM), to verify the simulation experiment results of the two models. The classification judgments of the two algorithm models are shown in Figure 6, respectively. Figure 6 (a) shows the distribution of LDA model classification judgment and actual classification, and Figure 6 (b) shows the distribution of SVM algorithm topic judgment. Out of a total of 30 test samples, there were 4 samples with incorrect LDA judgments, with an accuracy of 86.7%. The SVM algorithm has 9 error samples and a judgment accuracy of 70%.

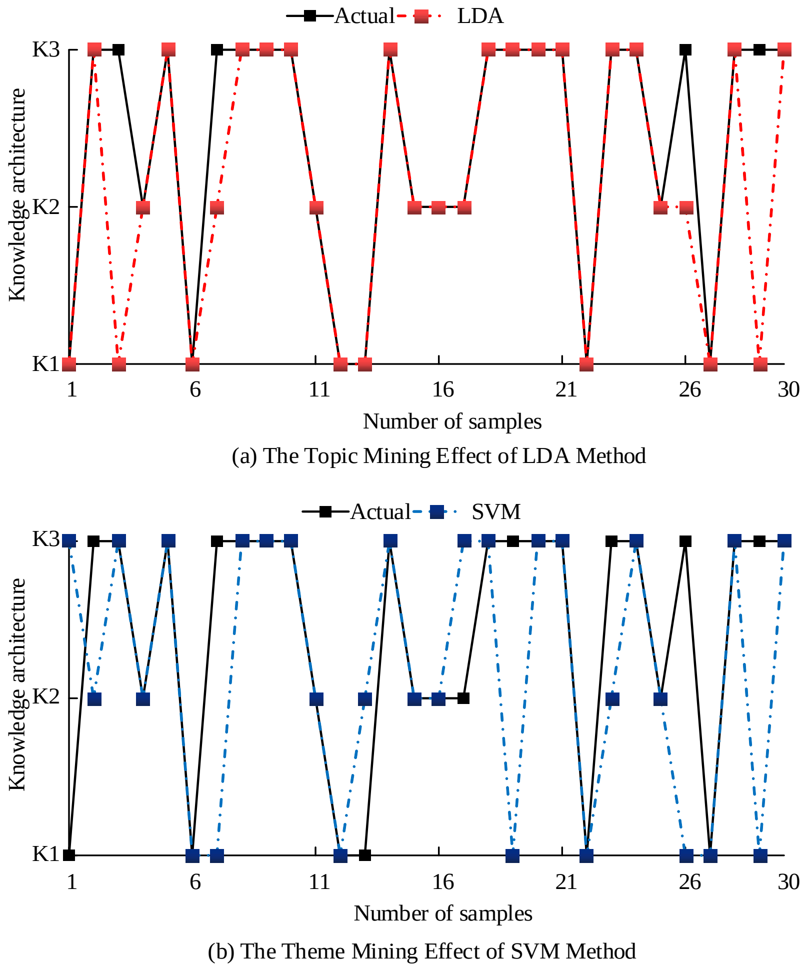

Figure 6: Topic mining distribution of two algorithm models

4.2. Experimental analysis of TPLC collaborative teaching group division method

This study used data texts from the professional learning community space of a certain primary and secondary school teacher as samples for exploration, and differentiated the cooperation grouping of teachers according to the control group and experimental group. The grouping result of the new learning community spatial management is the experimental group, while the traditional method is the control group. Table 3 shows the interactive attributes of learning communities among different groups.

In Table 3, the density and clustering coefficient of the experimental group were generally higher than those of the control group. The average distance of the control group was generally greater than that of the experimental group. The clustering coefficients of each experimental group were all greater than 0.6. Compared with the clustering coefficients of the control group, except for one group which reached 0.6, the clustering coefficients of all other groups were lower. Therefore, the grouping of teachers in the experimental group has stronger cohesion and is more likely to form a collaborative learning space.

Table 3: Interactive attributes of learning communities in different groups

Grouping	Number of interactions	Density	Cluster coefficient	Average distance
Experimental Group 1	60	0.950	0.950	1.000
Experimental Group 2	34	0.841	0.720	1.040
Experimental Group 3	38	0.852	0.734	1.147
Experimental Group 4	66	1.000	1.000	1.240
Control group 1	17	0.345	0.478	1.422
Control group 2	24	0.428	0.661	1.497
Control group 3	47	0.574	0.254	1.614
Control group 4	19	0.374	0.012	1.597

The specific results obtained by visualizing the interaction data between different groups through Netdraw are shown in Figure 7. The experimental group has stronger interactive network relationships. The size of each node of the experimental group teachers is relatively balanced, indicating that the group formed under the new grouping method has higher concentration.

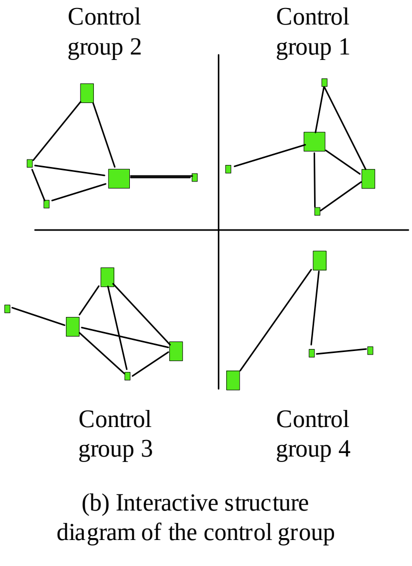

Figure 7: Visualization of interaction networks among various groups

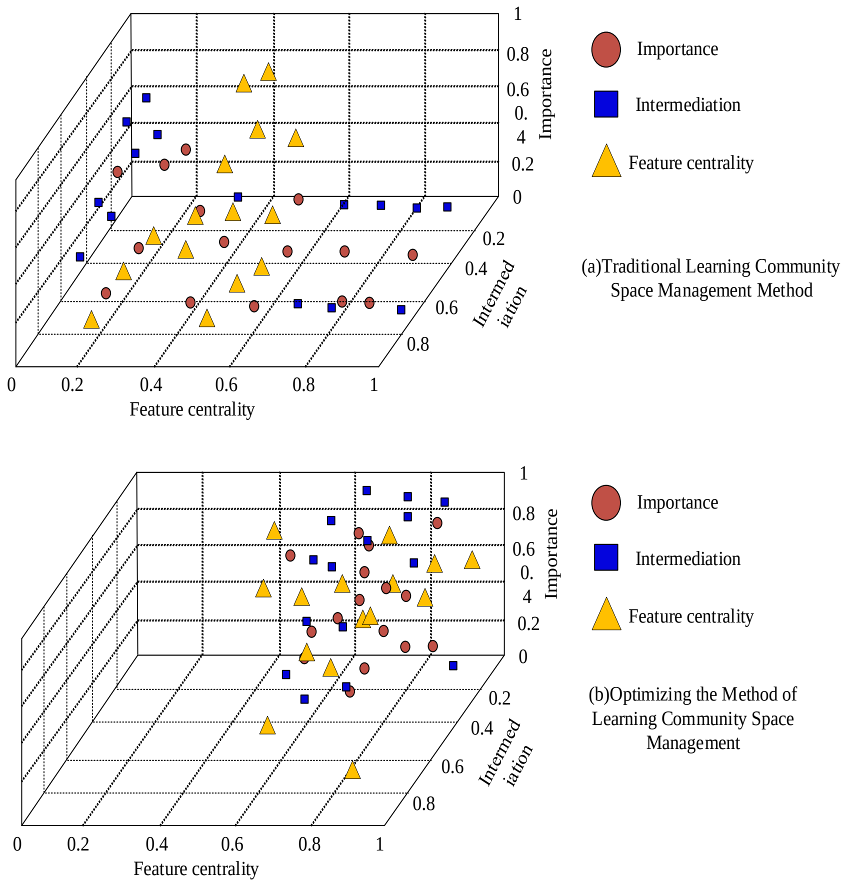

Figure 8: Interaction indicators under different groups

Comparing the interaction indicators between different grouping methods, Figure 8 is obtained. Under the new grouping method, the importance of teacher interaction, mediation, and feature centrality have all been enhanced. The interaction indicators of traditional methods are scattered and generally below the level of 0.6. Under the new grouping method, the interaction indicators of teachers are generally higher than 0.5.

Finally, this study collected evaluations from 40 teachers on the TPLC platform of the primary school on the optimized management system and traditional system. The maximum score is 1. A score below 0.33 indicates failure, a score above 0.66 indicates excellence, and a score range of 0.33-0.66 indicates general evaluation. In the figure, 40 users rated the optimization system higher, with a general rating range of 17 and an excellent rating range of 7. However, traditional dialogue systems only have 4 excellent ratings, with a general evaluation range of 12 people. The experiment shows that the learning community space system optimized by information technology has brought a good user experience to teachers.

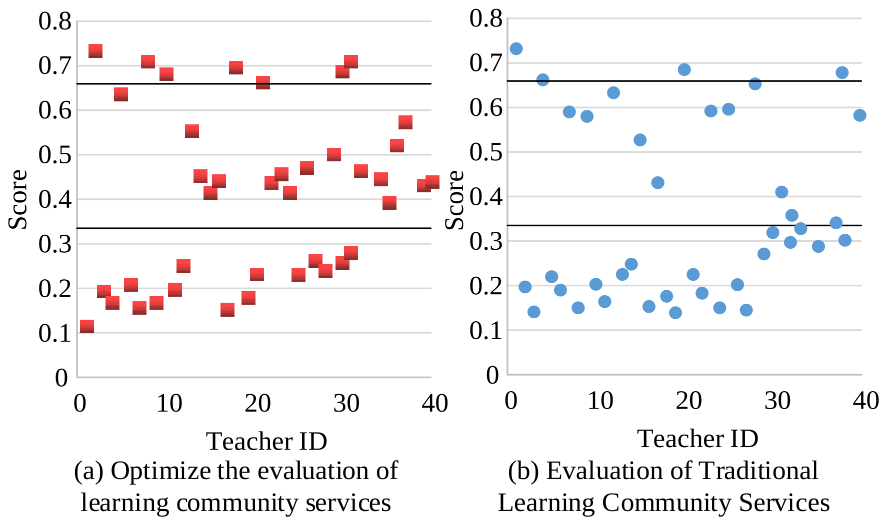

Figure 9: Feedback rating of users on the learning community platform management system

5. CONCLUSION

This study proposed an optimization method based on LDA topic model and K-means to address the management issues in the current TPLC space. The results verified that K-means participation in cooperative teaching groups increased the importance of teacher interaction, mediation, and feature centrality. The interaction indicators of traditional methods were scattered and generally below the level of 0.6. Under the new grouping method, the interaction indicators of teachers were generally higher than 0.5. The LDA topic model could distinguish the topic evolution and interaction status of the teacher learning library community space through the calculation of subject strength. In subject mining, the judgment accuracy of LDA was 86.7%, while the judgment accuracy of SVM algorithm was 70%. The experiment has proven that information technology-based methods can optimize the system management of the learning community space, improve the efficiency of teacher grouping and discussion interaction. The limitation of this study is that the data sources for simulation experiments and actual experiments are relatively single, and more universities are needed as samples for effect discussion. In future research, it is necessary to increase the text recognition function of TPLC space, while improving the effectiveness of current grouping and topic mining functions.

Funding. The research is supported by: Yuncheng University 2022 Applied Research Project; Research on the construction of interdisciplinary teachers' professional learning community in rural middle schools in Central China. (No. YY-202230).

REFERENCES

[1] M. K., Burns, M. R., Naughton, J. L., Preast, Z., Wang, R. L., Gordon, V., Robb, and M. L., Smith, "Factors of professional learning community implementation and effect on student achievement," *Journal of Educational and Psychological Consultation*, vol. 28, no. 4, pp. 394-412, 2018.
[2] S. Doğan, A. Adams, "Effect of professional learning communities on teachers and students: reporting updated results and raising questions about research design," *School effectiveness and school improvement*, vol. 29, no. 4, pp. 634-659, 2018.
[3] H. Schaap, E. de Bruijn, "Elements affecting the development of professional learning communities in schools," *Learning environments research*, vol. 21, pp. 109-134, 2018.
[4] R. Prenger, C. L. Poortman, A. Handelzalts, "The effects of networked professional learning communities," *Journal of teacher education*, vol. 70, no. 5, pp. 441-452, 2019.
[5] W. Admiraal, W. Schenke, L. De Jong, et al. "Schools as professional learning communities: what can schools do to support professional development of their teachers?," *Professional development in education*, vol. 47, no. 4, pp. 684-698, 2021.
[6] M. Chauraya, K. Brodie, "Conversations in a professional learning community: An analysis of teacher learning opportunities in mathematics," *Pythagoras*, vol. 39, no. 1, pp. 1-9, 2018.
[7] H. Eelodar, Y. Wang, C. Yuan, X. Feng, X. Jiang, Y., Li, & L. Zhao, "Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey," *Multimedia Tools and Applications*, vol. 78, pp. 15169-15211, 2019.
[8] C. Guo, M. Lu, W. Wei, "An improved LDA topic modeling method based on partition for medium and long texts," *Annals of Data Science*, vol. 8, pp. 331-344, 2021.
[9] D. Maier, A. Waldherr, Pl. Miltner, "Applying LDA topic modeling in communication research: Toward a valid and reliable methodology," *Communication Methods and Measures*, 2018, vol. 12, no. 2-3, pp. 93-118, 2018.
[10] K. Bastani, H. Namavari, J. Shaffer, "Latent Dirichlet allocation (LDA) for topic modeling of the CFPB consumer complaints," *Expert Systems with Applications*, vol. 127, pp. 256-271, 2019.
[11] S. Yang, H. Zhang, "Text mining of Twitter data using a latent Dirichlet allocation topic model and sentiment analysis," *International Journal of Computer and Information Engineering*, vol. 12, no. 7, pp. 525-529, 2018.
[12] C. Yuan, H. Yang, "Research on K-value selection method of K-means clustering algorithm," *J*, vol. 2, no. 2, pp. 226-235, 2019.
[13] H. Zou, "Clustering algorithm and its application in data mining," *Wireless Personal Communications*, vol. 110, no. 1, pp. 21-30, 2020.
[14] C. Wan, Y. Peng, K. Xiao, X. Liu, T. Jiang, & D. Liu, "An association-constrained LDA model for joint extraction of product aspects and opinions," *Information Sciences*, vol. 519, pp. 243-259, 2020.
[15] M. R. Bhat, M. A. Kundroo, T. A. Tarray, & B. Agarwal, "Deep LDA: A new way to topic model," *Journal of Information and Optimization Sciences*, vol. 41, no. 3, pp. 823-834, 2020.
[16] H. Jelodar, Y. Wang, C. Yuan, X. Feng, X. Jiang, Y. Li, & L. Zhao, "Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey," *Multimedia Tools and Applications*, vol. 78, pp. 15169-15211, 2019.
[17] V. Kukreja, "Recent trends in mathematical expressions recognition: An LDA-based analysis," *Expert Systems with Applications*, vol. 213, pp. 119028, 2023..
[18] M. Hasan, A. Rahman, M. R. Karim, et al. "Normalized approach to find optimal number of topics in Latent Dirichlet Allocation (LDA)," Proceedings of International Conference on Trends in Computational and Cognitive Engineering: Proceedings of TCCE 2020, Springer Singapore, 2021, 341-354.
[19] N. Chinthamu, & M. Karukuri, "Data Science and Applications, "*Journal of Data Science and Intelligent Systems*, vol. 1, no. 2, pp. 83–91, 2023.

[20] M. Gheisari, H. Hamidpour, Y. Liu, P. Saedi, A. Raza, A. Jalili, H. Rokhsati, & R. Amin, "Data Mining Techniques for Web Mining: A Survey", *Artificial Intelligence and Applications*, vol. 1, no. 1, pp. 3–10, 2022.
[21] D. Jackson, J. Temperley, "From professional learning community to networked learning community," Professional learning communities: Divergence, depth and dilemmas, pp. 45-62, 2007.
[22] K. Ismail, R. Ishak, S. H. Kamaruddin, "Development of professional learning communities model using fuzzy delphi approach," [J] *TEM Journal*, vol. 10, no. 2, pp. 872-872, 2021.
[23] M. Hensley, "Assessing the sustainability of social studies virtual professional learning communities on social media: a quantitative study of “sense of community”," *Social Studies Research and Practice*, vol. 16, no. 2, 93-114, 2021.
[24] D. K. Diehl, "Teacher professional learning communities and institutional complexity: Negotiating tensions between institutional logics," *Sociological Spectrum*, vol. 39, no. 1, pp. 5-19, 2019.
[25] C. Swanson, K. Earl Rinehart, J. Mills, "Focusing on teachers as learners in professional learning communities," *Teachers and Curriculum*, vol. 18, no. 1, pp. 1-5, 2018.
[26] V. Moosa, S. Salleh, L. Hamid, "Professional learning communities’ assessment–revised: A measure of schools as learning organisations," *Pertanika Journal of Social Sciences and Humanities*, vol. 28, no. 4, pp. 3047-3068, 2020.
[27] D. Diehl, "The multiplexity of professional learning communities: exploring the co-evolution of teacher social networks," *Research Papers in Education*, vol. 35, no. 5, pp. 623-639, 2020.