Model Essay Recommendation Method Based on Constant Evaluation of Writing Ability and Preference

Li Li(✉)

School of Languages and Cultures, Shanghai University of Political Science and Law, Shanghai, China
lily2211@126.com

Abstract. Model Essay Recommendation (MER) is one of the most important steps in teaching writing. However, most of the MER, especially for English writing teaching in China, still faces two problems to be solved. First, the difficulty in the recommended model essay doesn't match the real level of student, Second, the recommended model essays does not conform to the student's writing preference. To solve the above problems, this paper proposes a new model essay recommendation framework. First, a constant evaluation of English writing models both the writing ability and the writing preference of a student. Then, an interactive essay recommendation method is proposed, which takes teaching objectives, student writing profile, student feedback and other relevant factors into consideration at the same time. Finally, the experiments are conducted. The results show that the proposed method can recommend a more appropriate model essay for a student and improve the effectiveness of writing teaching.

Keywords: English writing teaching · essay recommendation · student profile · learning interaction

1 Introduction

English writing teaching has always been an important part of English education in Chinese schools, and it is also one of the main difficulties faced by Chinese students in English learning [3]. In recent years, a variety of auxiliary teaching software has been applied to English writing teaching, such as automatic scoring system [5], model essay recommendation system [4], which promotes English writing teaching to a large extent.

However, there are still some problems in the model recommendation to be solved. First, the difficulty of the recommended model essay doesn't match the level of students, which hinders the student's learning. Second, the recommended model essays are monotonous, which does not conform to students' expression preferences and is not conducive to students' fast learning. Writing is different from other aspects in the English learning including vocabulary, grammar, in that there are many ways to express each intention. Each student has his own expression preferences, including words, sentence formation, narrative techniques. It is easier to arouse the resonance of students and more

Z. Zhan et al. (Eds.): ICBDIE 2022, AHCS 5, pp. 1059–1069, 2023.
https://doi.org/10.2991/978-94-6463-034-3_108

conducive to students' rapid imitation and learning. Therefore, the essay recommendation that meets the above two points is required to improve the effect of English writing teaching.

Our previous work [2] proposed the method of constant evaluation of L2 students' English writing ability, in which we defined the evaluation dimension of writing ability. It mainly includes 13 dimensions of word, phrase, sentence, paragraph and discourse. The scores of a student in each dimension can be obtained more accurately through continuous evaluation method. Finally, we constructed the knowledge map to express students' overall writing ability. This method can solve the first problem of essay recommendation, but it cannot solve the second one well. The reason is that this method mainly carries out multidimensional writing ability rating and does not pay attention to the preference modelling of authors' writing tendency and preferences in each evaluation dimension. Therefore, this paper tries to build a model essay recommendation method based on constant evaluation of student's writing ability and preference. First, we propose a framework of model essay recommendation, then we propose an improved constant evaluation of writing ability which furthers modelling writing preference of a student. At last, we propose an interactive essay recommendation which can recommend essay that meets the required difficulty and writing preference of the student.

2 Framework of Model Essay Recommendation

BaseD on the above discussion, we propose a new framework to recommend an appropriate model essay for a student, as shown in Fig. 1, which mainly includes Automated Essay Evaluation Module, Model Essay Recommendation Module, Essay Evaluation Index and Model Essay Library, Student Writing Profile and so on.

2.1 Modules

The components and functions of the framework are as follows:

Essay Evaluation Index (EEI) defines the dimensions to evaluate an essay or profile a student's writing ability.

Student Writing Profile (SWP) describes the writing ability and preference of a student in detail from the dimensions in EEI.

Automated Essay Evaluation (AEE) analyzes an essay from the aspect of different EEI and gets the score and preference on different EEI.

Model Essay Library (MEL) consists of model essays which have been scored and preference labelled by AEE. Each essay in MEL has labels of ESI which are the basis of recommendation.

Model Essay Recommendation (MER) selects appropriate model essays from the MEL by matching the student's writing profile and labels on each model essay.

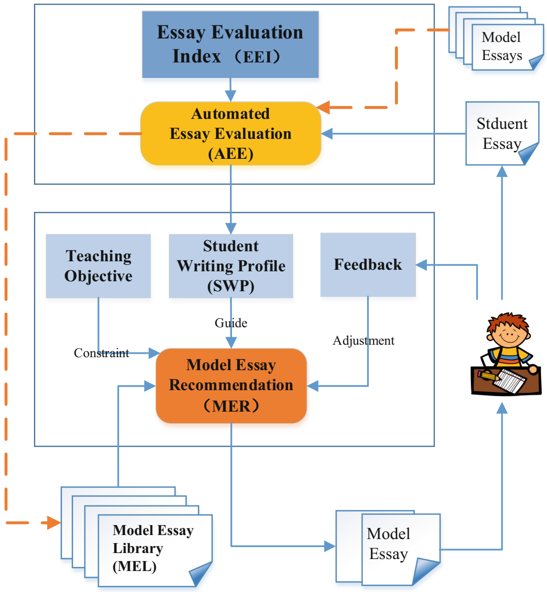

Fig. 1. Framework of model essay recommendation.

2.2 Workflows

Based on the key module AEE, the framework functions are described as follows:

Building MEL. For a special teaching purpose, a certain number of model essays are selected and sent to the AEE. Each one is scored and identified the writing preference from different EEIs. All labelled model essays form the MEL.

Building SWP. The SWP is also organized based on the EEIs. For a student, each of his essay can be scored and writing preference can be identified by AEE. Thus, over a period of time, the student's writing ability and preference can be identified based on a certain number of essays.

Recommending model essay: The recommending is a cycle process. In each cycle, the MER takes into account SWP, teaching objective, and feedback from a student and recommends the most appropriate model essays from MEL to the student. As the cycle goes on, the SWP will become more accurate, and more feedback will be obtained, all of which are helpful to improve the recommendation.

3 Essay Evaluation

The building of MEL and SWP are based on the evaluation of each essay. In our method, the unified essay evaluation indexes are defined firstly, and then an essay are evaluated from two aspects: one is the level on each evaluation index and the overall of the essay, and the other is the writing preference of the essay. In our previous work, the definition of evaluation index and how to get the level of on each index are discussed in detail and turned out to be effective. In this section, we mainly focus on the optimization of evaluation and how to get the writing preference of an essay.

Table 1. Essay evaluation indexes

Primary Index	P-weight	Secondary Index	S-weight
1. word & phrase	0.15	1.1 spelling	0.3
		1.2 grammar	0.3
		1.3 vocabulary	0.4
2. sentence	0.30	2.1 punctuation	0.2
		2.2 sentence structure	0.2
		2.3 sentence grammar	0.3
		2.4 sentence pattern	0.3
3. paragraph	0.25	3.1 sentence coherence	0.4
		3.2 topic relevance	0.6
4. discourse	0.30	4.1 ideas	0.1
		4.2 organisation	0.3
		4.3 paragraph coherence	0.3
		4.4 theme relevancy	0.3

3.1 Essay Evaluation Index

Based on the definition of evaluation indexes in [2], we optimize the weight of each secondary index and set the new weight for each primary index according to teaching practice, and the optimized essay evaluation indexes are shown in Table 1.

3.2 Writing Preference

Considering the operability of modelling and acquisition, we select vocabulary, sentence and discourse from EEI to construct preference model for an essay. It is difficult to adopt a unified expression form and evaluation method because there are great differences among different evaluation indexes.

3.2.1 Vocabulary

The preference of vocabulary dimension reflects the author's vocabulary preferences. Due to the discreteness and large number of words, it is impossible to list all the author's common words, so this paper adopts the hierarchical classification method for modelling vocabulary preference. It is graded according to vocabulary difficulty (G) and classified according to part of speech (P) . The vocabulary classification is the same as the essay evaluation. The vocabulary preference (VP) of the graded subclass (GP) for a single essay is expressed as

$$VP = (\mathrm{GP}_1 \ldots \mathrm{GP}_n). \tag{1}$$

Each vocabulary subclass GP is modelled and dynamically updated separately. GP is represented by the vector model of the author's preference words in this vocabulary subclass, specifically,

$$\mathrm{GP} = \left(\left(\mathrm{wp}_1, \mathrm{tf}_1\right), \ldots, \left(\mathrm{wp}_n, \mathrm{tf}_n\right)\right) \tag{2}$$

where, $\left(\mathrm{wp}_i, \mathrm{tf}_i\right)$ is a two-tuples of wp and tf. Here, the value of tf is not normalized to preserve the differences among different essays.

The process of modelling the vocabulary preference of each essay is as follows: The first step is to determine the GP of each word according to the vocabulary classification and part of speech; The second step is to count the frequency of occurrence of words in each GP and calculate the average the frequency of occurrence. The third step is to retain the words above the average line, and adjust the GP words according the preference benchmark that is 1/3 of the amount of GP words.

3.2.2 Sentence

According to different standards, sentence can be divided into different types. According to their use, sentences are declarative, interrogative, imperative or exclamatory. According to their structure, sentences are simple, compound, complex or compound-complex. From the rhetorical point of view, sentences are loose, periodic and balanced. Sentence modelling is only needed to record the classifications of the sentences used by the author. In order to facilitate the flowing tasks, the sentences are also classified by single layer. The sentence preference is expressed as

$$SP = ((\mathrm{st}_1, \mathrm{tf}_1), \ldots, (st_n, \mathrm{tf}_n)). \tag{3}$$

st_i is the sentence classification, which includes use-declarative, use-interrogative, use-imperative, use-exclamatory, structure-simple, structure-compound, structure-complex, and structure-compound-complex. The tuple $(\mathrm{st}_i, \mathrm{tf}_i)$ means that st_i is used tf_i times in the essay.

The key to modelling the sentence preference is to classify each sentence in the essay. We use part-of-speech tagging and dependency syntax to analyze the grammatical structure of each sentence, and then judge the category of the sentence according to its grammatical structure. After all the sentences are analyzed, the usage frequency of each sentence classification can be gotten.

3.2.3 Discourse

We model the discourse of an essay from the structure division, topic sentence distribution and so on.

The essay structure division (DP_s) is characterized by essay length el, number of paragraphs pn and average length of paragraphs pl,

$$DP_s = (el, pn, pl) \tag{4}$$

The distribution of topic sentences (DP_t) is described as

$$DP_t = \left(\frac{nt}{np}, \frac{nb}{nt}, \frac{nm}{nt}, \frac{ne}{nt}, \frac{nz}{nt}\right) \tag{5}$$

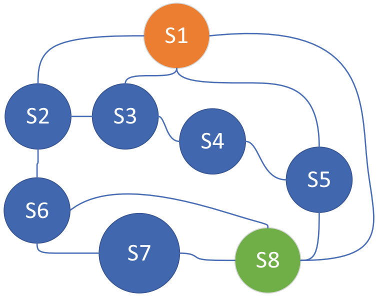

Fig. 2. Example of sentence relation graph

nt is the number of topic sentences. *np* is the number of paragraphs. *nb*, *nm*, *ne*, *nz* are the number of topic sentences at the beginning, middle, end of paragraphs and paragraphs without a topic respectively.

The key of DP_t modeling is to identify topic sentence of each paragraph. In this paper, paragraph topic sentences are identified based on sentence relation graph, listed as follows,

1) Take sentences as nodes and the correlation between sentences as edges to construct a sentence relation graph of each paragraph, shown in Fig. 2.

Among them, the correlation a_{ij} between a pair of sentences $\left(s_i, s_j\right)$ is defined as the sum of the correlation degree of the association rules between this pair of sentences. The association rules between a pair of sentences refer to those association rules whose front keys and back keys happen to be in this pair of sentences respectively. Fig. 2 is an indirected graph, where the direction of the association rule is ignored.

$$a_{ij} = \text{normal}\left(\sum_{s} ar\right) \tag{6}$$

where, s is the set of association rules hit by this pair of sentences, ar is the weight of an association rule, and the final step is normalized.

2) The importance of a sentence (si) is calculated by the centrality and position of the sentence,

$$\text{si} = \alpha d + \beta p \tag{7}$$

d is the degree of the sentence node in sentence relation graph. p is 1, if the sentence is first sentence of the paragraph. p is 0.6, if the sentence is last sentence of the paragraph. Otherwise, p is zero. Generally, $\alpha = 0.8$, $\beta = 0.2$.

3) Select the most important sentence as the topic sentence of the paragraph. If the importance of the sentence has no significant advantage, that is, $\left(\text{t} - \bar{t}\right)/\bar{t} < 0.2$, $\bar{t}$ is the average importance of all sentences in the paragraph, then the paragraph is considered to have no topic sentence.

4 Constant Evaluation of Writing Ability

4.1 Student Writing Profile

In order to continuously track students' writing ability and preferences, we propose a constant student writing profile (SWP_c), which is composed of a student's writing ability and writing preferences in several consecutive cycles and described as

$$\mathrm{SWP}_c = \langle \mathrm{WA}, \mathrm{WP} \rangle \quad (8)$$

WA refers to the writing ability of a student, which is described by *wa* of *n* evaluation cycles,

$$WA = (wa_{t1} \ldots \mathrm{w}a_{tn}) \quad (9)$$

wa is defined using the same EEI as single essay in Sect. 4. The definition *wa* is defined as the same evaluation index as section4 in which single essay and its value is the synthesis of several essays. Since the index is mainly represented by quantitative value, the synthesis can be averaged.

WP refers to a student's writing preferences, which are composed of writing preferences in successive evaluation cycles and are defined as:

$$WP = \left(wp_{t1}, \ldots \mathrm{wp}_{tn}\right) \quad (10)$$

wp is defined in the same way as the modeling method of writing preferences in Sect. 4, and its value is the synthesis of several essays within the evaluation cycle. Different dimensions are expressed in different ways, and the synthesis method is not completely the same. In the vocabulary dimension, the total vector of each category is recalculated based on the preference vector of each category. In the sentence dimension, the average frequency of each sentence classification is calculated directly. In the discourse dimension, the average values of structure division and topic sentence distribution are calculated respectively.

4.2 Constant Evaluation

Accurate modelling of a student's writing ability and preferences must be based on a certain number of essays completed by students in an evaluation cycle. According to the actual writing practice of Chinese students, students write about two or three essays per week and about ten essays per month, so we conduct modelling in a cycle of four weeks. Meanwhile, for the continuity of modelling, we adopted the sliding window method for modelling, as shown in Fig. 3. The size of the window is four weeks, and the sliding time is one week. In one window, WA and WP in *S*WP are constructed by using the method of section 5.1.

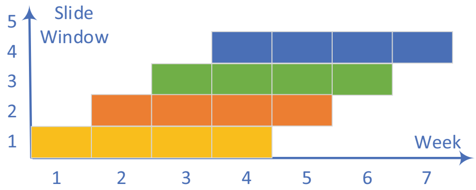

Fig. 3. Constant evaluation based on slide windows

5 Interactive Model Essay Recommendation

As shown in Fig. 1, the proposed model essay recommendation method considers teaching objectives, student writing profile, student feedback and other factors at the same time, and selects the most appropriate model essay from SML library for recommendation. The recommendation function is

$$sim(\mathrm{f} \cdot \mathrm{TO} \mp (1 - \mathrm{f}) \cdot \mathrm{SWP}_c,\ \mathrm{SML}) \tag{11}$$

TO is the teaching objective, including the writing ability target wa and the writing preference target wp. The value of wa can be -1, 0, $+1$ to indicate that the difficulty of the sample is reduced by one level, unchanged, or increased by one level. wp can be 0, 1, where 1 indicates that the student is recommended in accordance with their writing preferences, and 0 indicates that students are not recommended in accordance with their writing preferences, which means students can be recommended to write compositions in different ways if necessary. f is the students' feedback scores which vary from 0.1 to 0.5. When there is no feedback, f takes the median value 0.25. The $\mp$ operator means that SWP_c is adjusted using the value of TO to get the desired model essays. The *sim* function calculates the similarity of adjusted SWP_c and model essays in SML, and recommends model essays to students according to the similarity value. The interaction is mainly reflected in that students influence the recommendation strategy through the feedback value, and this process is repeated until the recommended composition is satisfactory to students. If the students' needs cannot be met after several rounds of recommendation, the recommendation will fail.

6 Evaluations

To verify the effect of the proposed method, we designed two experiments: one is the accuracy evaluation of model essay recommendation, and the other is the comprehensive evaluation of writing ability improvement.

6.1 Participants

We randomly select 45 non-English major students from different departments as the participants. Among these students,15 students are from level one, 15 students are level two, and the rest are from level three. 15 students from each level are randomly divided into 3 groups with 5 students in each group. A total of 9 experimental groups are obtained, named 1A, 1B, 1C, 2A, 2B, 2C, 3A, 3B and 3C respectively.

Table 2. Students' feedback of recommendation

group C	level one	level two	level three	average
week4	0.25	0.26	0.22	0.24
week8	0.31	0.34	0.35	0.33
week12	0.39	0.40	0.41	0.40

6.2 Procedures

1) Construct the model essay library: Teachers select excellent essays and students' essays of different writing levels and preferences, and evaluate the selected essays' writing levels and preferences by using the method in Sect. 4. At last, the evaluation results are optimized by the teachers.

2) Evaluation of initial writing ability: Students from all 9 groups are given a writing test of different types and graded by teachers.

3) Learning and continuous evaluation: The essay recommendation study and evaluation are conducted for 12 weeks in a raw. Among them, group A does not use the recommendation model, directly learn excellent model essays; Group B is recommended based on the writing ability in the student profile. Group C is recommended based on the whole user profile for learning. Tests will be performed every four weeks. At the same time, students are required to give feedback on each essay recommended by the system. The above experiments are carried out independently in the students of three levels.

6.3 Experimental Results and Analysis

The experimental results of accuracy evaluation of model essay recommendation are shown in Table 2. In the experiment, we use the students' feedback as the evaluation. Only the feedback of group C is used for the evaluation. The values range from 0.1 to 0.5, and the best score is 0.5. The values smaller than 0.25 means that the recommendations are not good. The results show that the accuracy of recommendation is being improved with learning going on.

The comprehensive evaluation of writing ability improvement is shown in Fig. 4. In the experiment, the students of 9 groups have similar essay scores in the first week. During the learning process, different groups use different recommendation methods. As a result, the scores of different groups are significantly different. Group C is better than group B, and group B is better than group A, which shows that the proposed recommendation method is really effective.

The learning speed experiments are shown in Fig. 5. The results of periodic tests in the 1st, 4th, 8th, and 12th weeks are used to draw the curve of improvement of writing ability. By comparing the curves of A, B and C, it can be seen that the improvement rate of writing ability in the third group is higher than that in the second group, and the improvement rate of writing ability in the second group is higher than that in the first group, which further verifies the effectiveness of the recommendation model proposed in this paper.

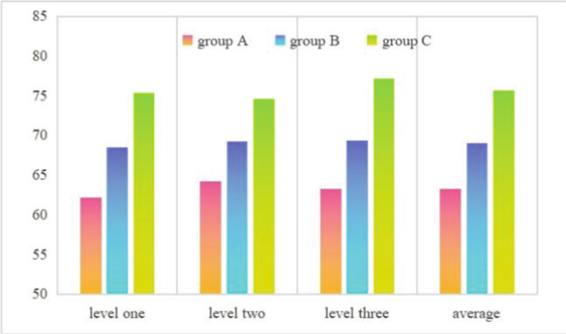

Fig. 4. The evaluation of writing ability improvement.

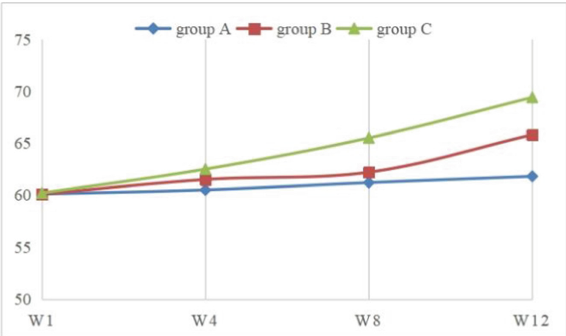

Fig. 5. Speed comparison of writing ability improvement

7 Conclusions

In this paper, we propose a model essay recommendation method based on constant evaluation of student's writing ability and preference.

By recommending the model essays that fit a student's writing ability level and writing preferences, he could learn more efficiently. And the experimental results show that the proposed methods are effective.

In this paper, the modelling of writing preference only uses words, sentences and part of text elements, which cannot fully describe user writing preference. Therefore, the next work of this paper is to improve writing preference indicators and automatic evaluation methods, in order to provide more accurate writing recommendation services.

Acknowledgements. This research was financially supported by the fund from China National Institute for SCO International Exchange and Judicial Cooperation (No.20SHJD021)

References

1. Burstein, J. C., Marcu, D. S., et al., (2001). Towards automatic classification of discourse elements in essays, In ACL2001, the 39th annual meeting of the Association for Computational Linguistics, pp. 90–92, France.
2. Li, L. (2017). Constant evaluation of L2 students English writing ability. In SEKE2017, the 29th International Conference on Software Engineering and Knowledge Engineering, pp.404-409, Pittsburgh, USA.
3. Li, L. & Sugumaran, V. (2019). A cognitive-based AES model towards learning written English. J Ambient Intell Human Comput. 10, 1811–1820.
4. Liang, G.X., On, B.W., Jeong, D., et al. (2021). A text GAN framework for creative essay recommendation. Knowledge-Based Systems, 232, 107501
5. Ramesh, D., & Sanampudi, S. K.. (2021). An automated essay scoring system: a systematic literature review. Artificial Intelligence Review, 55(3), 2495–2527.