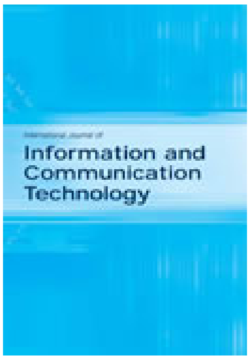


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Behavioural perception-driven evolutionary pathways for vocational English oral communication: fusing graph convolutional networks with multi-objective optimisation

Bowen Yuan, Yating Yang

DOI: 10.1504/IJICT.2025.10075149

Article History:

Received:	09 October 2025
Last revised:	05 November 2025
Accepted:	06 November 2025
Published online:	12 January 2026

Behavioural perception-driven evolutionary pathways for vocational English oral communication: fusing graph convolutional networks with multi-objective optimisation

Bowen Yuan* and Yating Yang

Liao Yuan Vocational Technical College,
Liaoyuan, 136200, China
Email: hebybw2008@163.com
Email: emilyyaya@sina.cn
*Corresponding author

Abstract: Existing English speaking teaching path methods ignore learners' behavioural characteristics, leading to inaccurate generation of personalised teaching paths. This paper first integrates the production-oriented approach to design the teaching process and constructs an intelligent teaching path generation model that integrates learner behavioural preference analysis and the teaching process. The model utilises graph convolutional networks and long short-term memory network to capture semantic associations and learners' dynamic evolutionary characteristics. Then, using non-dominated sorting genetic algorithm II for multidimensional optimisation, multiple paths are optimised for multiple objectives to obtain a set of frontier solutions, and the English speaking teaching path with the highest score is obtained. Experimental results show that the suggested approach improves prediction accuracy by an average of 7.04%–23.96% while significantly reducing the time required to solve for the optimal path, validating the model's efficiency.

Keywords: English speaking teaching path; production-oriented approach; graph convolutional network; long short-term memory network; NSGA-II method.

Reference to this paper should be made as follows: Yuan, B. and Yang, Y. (2025) 'Behavioural perception-driven evolutionary pathways for vocational English oral communication: fusing graph convolutional networks with multi-objective optimisation', *Int. J. Information and Communication Technology*, Vol. 26, No. 49, pp.75–91.

Biographical notes: Bowen Yuan is a Lecturer at the Liao Yuan Vocational Technical College, Liaoyuan, China. She obtained her Bachelor's degree in English Education from Harbin Normal University, Harbin, China in 2011. And she obtained her Master's degree in Subject Teaching (Chinese) from Liaoning Normal University, Dalian, China, in 2013. Currently, she is pursuing her Doctor of Education (EdD). Her research fields mainly include the research on English teaching and student development and education.

Yating Yang is a Lecturer at the Liao Yuan Vocational Technical College, Liaoyuan, China. She obtained a Master's degree in Linguistics and Applied Linguistics from Xi'an International Studies University, Xi'an, China, in 2012. Her research fields mainly include English teaching in higher vocational colleges, linguistics and applied linguistics.

1 Introduction

In the current era of accelerating globalisation, English, as a key language for international communication, has become increasingly important. However, traditional models of English oral instruction have gradually revealed many shortcomings in practical teaching (Sritulanon et al., 2018). On one hand, teaching methods are relatively monotonous and lack real language communication scenarios and interactive elements, resulting in students being passive recipients of knowledge in class (Yang et al., 2024). On the other hand, course content is often disconnected from daily life and fails to meet students' needs for using English to communicate in various situations. The rapid development of artificial intelligence (AI) technology has brought revolutionary changes to the field of education. Its powerful data processing capabilities can enable multi-dimensional analysis and real-time feedback on students' oral performance, effectively compensating for the shortage of teacher resources (Du and Daniel, 2024). The construction of personalised learning paths can better meet different students' learning needs, enabling teaching tailored to individual aptitudes (Romero et al., 2023). It is evident that combining production-oriented approaches with artificial intelligence not only leverages the leadership of production-oriented approach (POA) in English instruction concepts but also utilises AI technology to overcome resource limitations in traditional teaching. Their organic integration provides the potential to build an efficient, precise, and personalised English speaking teaching path.

Early methods organised English learning resources and the semantic relationships between learning content in the form of graphs (Tapalova and Zhiyenbayeva, 2022). This approaches attempted to simulate and assist in constructing the semantic networks within the human brain through external, visual graph structures. This connected isolated linguistic knowledge points and learning resources via rich semantic relationships, ultimately supporting an associative, nonlinear, exploration-driven, and efficient learning model. Anas (2019) considered prerequisites for English courses, degree requirements, and course scheduling factors, proposing three navigation algorithms based on the terminal semester, learners' learning objectives, and learning styles or preferences respectively to generate feasible paths, and using ranking algorithms to filter out top-k teaching paths. Liu (2022) constructed a course graph and a two-layer graph of courses-learning objects, utilising depth-first traversal (DFS) (Du and Su, 2021) to generate all teaching paths that meet time constraints, and assessed the possible grades learners might achieve after completing these paths using indicators such as average scores, ultimately determining the optimal path. The core of the above method lies in searching and pruning; however, with an increase in the length of teaching paths, there is a problem of high algorithm complexity. Liu (2023) constructed an English teaching knowledge graph to organise learning resources according to prerequisites and used graph-based methods to find suitable learning paths for each learner. However, this method does not achieve high accuracy in designing teaching paths. Xu et al. (2024) extracted knowledge units from English courses to construct a thematic map and considered the group members' learning preferences or prior learning backgrounds to build teaching paths based on either learning time or learning preferences.

As the number of courses or the length of generated teaching paths increases, the time and space complexity of the algorithm also increase accordingly. Teaching path design methods based on similarity and the POA (Zhu, 2023) can effectively construct learner profiles through data mining techniques to analyse similarities among learners and thereby provide personalised teaching paths. Compared with graph-based methods, these approaches have lower implementation costs. Maimaiti (2019) determined English teaching objectives according to the standards of English courses, students' actual levels, and teaching needs, and obtained a globally optimal teaching path by improving the genetic algorithm (GA) (Benmesbah et al., 2023). Guo et al. (2022) systematically organised and effectively processed learning materials according to the English teaching process and students' cognitive patterns, and used customised measurement methods to calculate learner similarities for designing teaching paths. Jiang (2022) designed teaching paths based on POA learning resource attributes, learners' learning dynamics, and preferences by using a customised measurement method to compute learner similarity. Although these approaches can construct student profiles through analysing English teaching system data such as course selections and performance records, the actual amount of available data is relatively small, leading to challenges in path design accuracy due to data sparsity.

The core task of teaching path design is to establish a mapping relationship between target learners and teaching paths (Nigenda et al., 2018). Due to the complexity of the English education process, which is difficult to accurately describe with manually defined rules, researchers have turned to deep learning. Deep learning-based English teaching pathways dynamically adjust subsequent learning content, difficulty levels, sequence, and resource types based on learners' real-time performance and long-term progress. When the system detects persistent errors on specific knowledge points, it automatically pushes additional explanatory videos, varied practice exercises, or even reduces question difficulty – rather than forcing learners to advance to the next unit. This approach has strong fitting capabilities by learning potential patterns and hierarchical representations from sample data. Liu (2024) based on long short-term memory network (LSTM) performs teaching effect prediction according to learners' learning processes or materials output by POA, thereby achieving adaptive teaching paths. Iatrellis et al. (2020) analysed the causal relationship between learner study behaviours and study outcomes, with a causal Bayesian network (Kitson et al., 2023) constructing a model to evaluate learners' academic performance, and obtaining optimal learning effect teaching paths by analysing the causal relationships between study behaviours and effects. Su and Shen (2025) constructed learner vectors based on learners' historical submission sequences and capability maps, using graph neural networks to design teaching paths for learners. Zhang et al. (2025) mined students' answer records in learning systems, utilising a multi-scale graph convolutional network (GCN) model (Bhatti et al., 2023) and deep knowledge tracing methods to predict students' mastery levels of knowledge points and generate English teaching paths according to the prediction results.

According to the analysis of existing research on English spoken teaching path design, these methods ignore learner behaviour characteristics, leading to inaccurate personalised teaching paths. To solve this problem, this study focuses on constructing an English spoken teaching path through POA and artificial intelligence technology collaboration. First, a blended online-offline English spoken teaching process based on production-oriented learning is designed, with the production-driven approach as the core engine, building an English spoken teaching path design method that integrates learner behaviour preference modelling and POA theory. GCN obtains initial learning resource embeddings. Furthermore, sequential hypergraphs are used to obtain interaction-based English spoken teaching course resources and learner preference embeddings. Combined with dual-channel gated recurrent unit (GRU) and transformer modules, the long-term and short-term characteristics of learners' historical interaction sequences are extracted. An initial learning path is generated based on the recommendation probability output by multi-layer perceptron (MLP). LSTM dynamically encodes fused information in course resource embeddings, learner answer records, and difficulty information to predict learner performance. Finally, a multi-objective optimisation module is designed using non-dominated sorting genetic algorithm II (NSGA-II) for optimising English spoken teaching paths through multiple objectives, and the final English spoken teaching path is determined from the Pareto frontier by calculating utility functions. Experimental results show that the prediction accuracy of the proposed method reaches 96.22%, with an optimal teaching path solving time of just 1.8 s, which not only improves the efficiency of English spoken instruction but also provides innovative paradigms for personalised language learning path design.

2 Relevant theory

2.1 Production-oriented approach

The purpose of POA is to overcome the drawbacks of the separation between learning and application in foreign language teaching, enabling students to learn for practical use and achieve results, thereby enhancing their language proficiency. POA advocates placing students at the core position, focusing on cultivating student capabilities, promoting autonomous, exploratory, and collaborative learning methods, emphasising production-oriented activities, thus significantly improving English expression skills and linguistic competence (Shu, 2022). The core concept of POA is a closed-loop process. It begins with a challenging output task to stimulate students' learning needs. Teachers then provide targeted input materials as scaffolding to help students overcome difficulties. Ultimately, students successfully complete the output task, achieving knowledge internalisation and skill enhancement. The entire process revolves around effective learning. Compared to traditional methods, POA heightens students' awareness of their own challenges, encourages proactive problem-solving, and clarifies their path to progress. This significantly enhances learning efficiency and precision.

POA holds that the goal of language learning is to effectively use language in practical communication; therefore, teaching activities should revolve around students' production tasks (Ding, 2023). Specifically, the production-oriented approach advocates setting clear communicative goals and language objectives to guide students in experimenting with and experiencing language use in real-world communicative scenarios, thus stimulating their desire to learn and improving their ability to apply language in practical communication. Compared to traditional foreign language teaching methods, POA pays more attention to student agency and learning behaviours, overcoming the drawbacks of 'learning without application' commonly seen in traditional teaching (Sun et al., 2024). In addition, this method also emphasises evaluation and feedback mechanisms; effective evaluation and feedback can help teachers promptly understand students' learning progress, identify problems present in their studies, and adjust teaching strategies and methods accordingly. At the same time, the evaluation and feedback mechanism can also help students recognise their strengths and weaknesses in language production, clarifying directions for improvement and setting clear goals.

2.2 The long short-term memory network model

LSTM is extended from RNN, mainly solving the problem of gradient disappearance and explosion during recurrent neural network (RNN) training (Song et al., 2020). It incorporates a special memory cell structure on the basis of RNN to control information transmission when certain information is input. LSTM has a deep structure along the time dimension; thus, it exhibits superior performance in problems where inputs have sequential correlations (Lindemann et al., 2021). LSTM includes an input gate, a forget gate, and an output gate. Their specific descriptions are as follows. As an improved version of RNN, LSTM effectively addresses the vanishing gradient and exploding gradient issues faced by traditional RNNs when processing long sequence data by introducing gating mechanisms and cell states, thereby significantly enhancing model performance.

1 Input gate: it determines how much new information enters the cell at the current time step and consists of two parts mainly. One is a sigmoid layer that decides which values need to be updated, the other is a tanh layer, updating certain information into the current state. The specific calculation formulas are shown in equations (1) and (2), where x_t is the input variable, W_i, W_c are weights, b_i, b_c are biases, h_{t-1} is the input variable of the implicit level, and σ, tanh are activation functions.

$$i_t = \sigma\left(W_i \cdot \left[h_{t-1}, x_t\right] + b_i\right) \tag{1}$$

$$\hat{C}_t = \tanh\left(W_c \cdot \left[h_{t-1}, x_t\right] + b_c\right) \tag{2}$$

2 Forget gate: it determines how much information from the previous time step is retained. Its inputs include h_{t-1} and x_t. f_t is the value obtained by applying the sigmoid activation function to h_{t-1} and x_t, as shown in equation (3), where W_f is the weight, b_f is the bias, and σ is the activation function.

$$f_t = \sigma\left(W_f \cdot \left[h_{t-1}, x_t\right] + b_f\right) \tag{3}$$

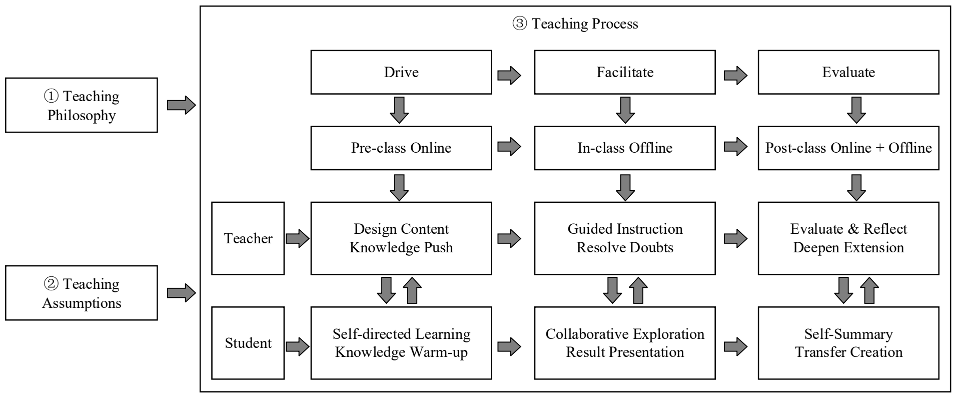
3 Output gate: it determines how much information from the current cell is output. The information within the cell needs to be processed by a tanh function. The calculation equation is as follows, where σ is the activation function, W_o is the weight, and b_o is the bias.

$$o_t = \sigma\left(W_o \cdot \left[h_{t-1}, x_t\right] + b_o\right) \tag{4}$$

3 Designing an integrated online and offline English speaking teaching process based on the output-oriented approach

Under the guidance of POA theory, teachers require students to preview lessons independently online before class and drive them to complete tasks (Li, 2021). In English classrooms, teachers provide feedback in an offline format to promote knowledge acquisition. After class, teachers evaluate student learning through a combination of online and offline methods. The integrated English teaching process based on POA theory transforms into an output-oriented model where students take initiative in their learning, which can be specifically divided into three steps as shown in Figure 1.

Figure 1 The process of oral English teaching based on the POA theory



1 Pre-class motivation: knowledge push and independent study online. In this new mode, language instruction for oral English is moved to an online platform. This article adopts two platforms: iSmart (Jiang et al., 2022) and Desh Real e-Learning (Hsu, 2017). As a mobile learning resource accompanying the textbook, iSmart mainly provides students with vocabulary practice, pronunciation training, and oral recitation scoring as its main content. On the Desh Real e-Learning platform, teachers upload course outlines and related materials according to the college English speaking syllabus, set interactive discussion questions, develop homework banks and exercise banks, etc., as teaching resources. After understanding this unit's learning objectives, content, key points, and difficulties, students fully utilise fragmented time outside class to engage in online self-study (Minghe and Yuan, 2013). Through online autonomy, students warm up and accumulate knowledge, conduct self-assessment, and prepare for subsequent learning. Online independent study not only frees up teaching time and space but also meets the personalised autonomous learning needs of students. The core advantage of online self-directed learning platforms lies in their provision of diverse tools and pathways to meet personalised learning needs. Students can build their own learning systems through a series of fundamental operations.

2 In-class facilitation: offline collaborative study and achievement presentation. The new model implements offline English instruction during class to provide feedback on online learning outcomes, facilitating knowledge transfer and consolidation. In the offline classroom, teachers collaborate with students in joint research. Teachers and students sort out key points and difficulties of the teaching content, consolidate online material; students present their self-study achievements while understanding classmates' progress, analysing and identifying gaps in their own learning. Based on task requirements, students form fixed groups to collaboratively complete tasks during offline classes, presenting the outcomes of their online learning (Ati and Parmawati, 2022). In class, teachers address common questions arising from student pre-study through explanations and evaluations; they design instructional activities

targeting teaching key points and difficulties to eliminate learning barriers for students.

3 Post-class evaluation, online + offline reflective assessment and deepening expansion. Based on the POA theory, the purpose of evaluation is to promote learning through assessment; therefore, evaluation is no longer merely correction and scoring but includes setting evaluation criteria and assessing task completion. To this end, based on network teaching platform statistical analysis, teachers and students jointly evaluate effectiveness. Teachers evaluate output results from students at different levels, offering further study suggestions. Students also reflect on their own learning situations, including their performance in teams, thinking characteristics, and learning habits. Meanwhile, to promote subsequent learning, teachers push practical questions related to the unit teaching content via the online platform as post-class assignments for students, requiring them to apply learned language knowledge to solve real-world problems.

4 Generating English speaking teaching paths through the coordination of output-oriented methods and artificial intelligence

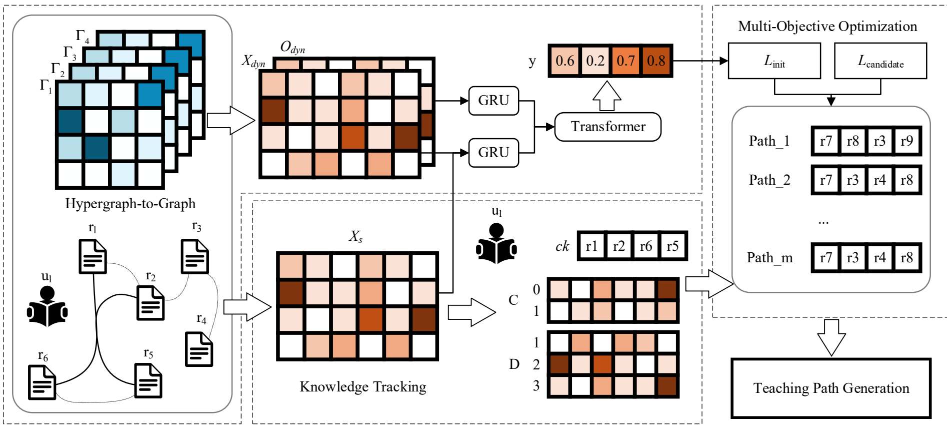
4.1 A model for generating English speaking teaching paths that integrates POA theory

Based on the previously designed English oral teaching process, a generating model for English oral instruction pathways integrating POA theory should be driven by production as its core engine, with input facilitation resources providing support and selective learning serving as a dynamic adjustment mechanism. It constructs a closed-loop teaching system combining learning behaviour preference modelling and knowledge tracking to achieve precise and personalised pathway generation. This model converts theoretical elements into operational teaching segments through modular design, ultimately serving the stepwise improvement of learners' oral production abilities.

Let the set of learners be $U = \{u_1, u_2, \ldots, u_m\}$, English oral instruction courses $R = \{r_1, r_2, \ldots, r_n\}$, historical interaction records $H = \{(\underline{u_i}, r_j, t_k, c_k, d_k)\}$, where t_k is a timestamp, c_k is answering questions, d_k is course difficulty, and $|H| = S$. Treat the learning courses R as graph nodes and the logical relationships between resources as edges ε_s, constructing an undirected graph $g_s = (R, \varepsilon_s)$, where $(r_i, r_j) \in \varepsilon_s$ indicates that there exists a logical association between resource r_i and r_j.

Abstract learners as hyperedges $u \in \varepsilon$, with multiple learning courses R they interact with serving as nodes connected by the hyperedge, constructing a hypergraph structure $g = (R, \varepsilon)$. The length of the learning resource sequence S is divided into T subsets based on timestamps to construct an ordered hypergraph $g = \{g^t = (R^t, \varepsilon^t)|t = 1, 2, \ldots, T\}$, where R^t represents the set of English oral instruction resources. The goal of generating English oral instruction pathways can be stated as: given historical interactions H with English oral instruction courses before a learner's timestamp t, and based on dependency relationships among English oral instruction resources, predict which English oral course resource at $t + 1$ is most likely to be selected by the learner as their next learning task. Traditional methods focus predominantly on single-objective optimisation, whereas this paper further introduces a multi-objective optimisation process, generating diverse teaching path sequences $L = \{r_{t+1}, r_{r+2}, \ldots, r_{t+k}\}$ that cover learner preferences, cognitive gains, English oral instruction resource compatibility, and diversity.

Figure 2 The teaching path generation model integrating POA theory (see online version for colours)



The overall model architecture is shown in Figure 2, which consists of three components: the first part is dynamic learning behaviour preference modelling. GCN obtains the initial learning resource embeddings and acquires English oral instruction course resources and learner preferences based on interactions through a chronological hypergraph. Dual-channel GRU and transformer modules are used to extract short-term and long-term features from learners' historical interaction sequences, and finally generate an initial learning path according to recommendation probabilities output by MLP. The second part is the dynamic knowledge tracking module, where LSTM dynamically encodes integrated information of course resource embeddings, learner answering records, and difficulty levels to predict learners' performance. The third part is a multi-objective optimisation module that uses NSGA-II (Ma et al., 2023) to perform multi-objective optimisation on English oral instruction pathways. The final English oral teaching pathway is obtained by calculating a utility function from the Pareto front.

4.2 Dynamic learning behaviour preference modelling based on graph convolutional neural networks

GCN captures the static logical relationships between English oral instruction course resources and obtains embedded representations of course resources. Through GCN's message passing mechanism, each English course resource embedding vector x_j is iteratively updated based on its neighbouring nodes to generate a course resource representation that includes global dependencies. An adjacency matrix A_s is constructed based on the dependency relationships in the English courses. The GCN encodes the relationship propagation of learning resources as shown in equation (5), where $X_S^{(l)}$ represents the node feature matrix at layer l, N is the total number of course resources, the initial embedding for the course resources is $X_S^{(0)}$, randomly initialised according to a normal distribution, ϕ is the activation function ReLU, $\tilde{D}_s$ is the degree matrix with self-loops, and $W_s^{(l)}$ represents the training parameters.

$$X_S^{(l+1)} = \phi\left(\tilde{D}_s^{\frac{1}{2}}\left(A_S + I\right)\tilde{D}_s^{-\frac{1}{2}} X_s^{(l)} W_s^{(l)}\right) \tag{5}$$

The hypergraph convolutional network (HGCN) learns higher-order relationships. A memory-enhanced sequential hypergraph network dynamically captures the complex dependencies between learners and English oral instruction courses to generate embedded learning resources that capture learner preferences. Since a learner's historical behaviour can accurately reflect their preference, it is possible to utilise the historical interaction between learners and learning resources for graph construction. In this paper, the timeline is divided into T equal-length intervals $\{T_1, T_2, \ldots, T_T\}$, where $T_t = [\tau_{t-1}, \tau_t)$ and $\Delta\tau = \tau_t - \tau_{t-1}$ represent fixed window sizes. For each time window t, a graph $g^t = (R^t, \varepsilon^t)$, $t = \{1, 2, \ldots, T\}$, is constructed, where R^t represents the node set and ε^t represents the hyperedge set. Hyperedge e_i^t connects all English course resources that learner u_i interacts with within T_t, i.e., $e_i^t = \{r_j | (u_i, r_j, t_k) \in H \wedge t_k \in T_t\}$. Each hyperedge e_i corresponds to a u_i's complete learning trajectory within a specific time window.

Node-to-hyperedge aggregation process, the hyperedge e_i feature o_i is weighted aggregated by its nodes r_j embedding representation x_j, as shown in equation (6).

$$o_i^{(l+1)} = \phi\left(\sum_{r_j \in N(e_i)} W_1^{(l)} x_j^{(l)}\right) \tag{6}$$

where ϕ is the activation function ReLU, and $W_1^{(l)}$ is the trainable weight matrix at layer l. $N(e_i) = \{r_j | H(r_j, e_i) = 1\}$ denotes the connection set specified in e_i.

Hyperedge-to-node aggregation process, node r_j feature x_j updates with a weighted sum of co-hyperedges e_i features o_i, as shown in equation (7), where ϕ is the activation function ReLU, and $W_2^{(l)}$ is the trainable weight matrix at layer l. $\varepsilon(r_j) = \{e_i | H(r_j, e_j) = 1$, representing the set containing node r_j.

$$x_j^{(l+1)} = \phi\left(\sum_{e_i \in \varepsilon(r_j)} W_2^{(l)} o_i^{(l+1)}\right) \tag{7}$$

The embeddings learned from the static English oral instruction course relation graph are fused with those learned by hypergraphs, as shown in equation (8), where ϕ is the activation function ReLU. x^s is the feature vector extracted by GCN for static graph convolution, and x^{dyn} represents vectors extracted from dynamic hypergraphs generated by HGCN. W_a, W_b are feature transformation matrices, w_g is the weight vector, and σ is the sigmoid activation function.

$$x^{fused} = gx^s + (1-g)x^{dyn} \tag{8}$$

$$g = \sigma\left(w_g^\top \left[\phi\left(W_a x^s\right); \phi\left(W_b x^{dyn}\right)\right]\right) \tag{9}$$

After modelling historical interaction behaviours, the initial teaching path is generated and the dynamic evolution process during student learning is modelled to generate an initial list. Dual-path GRUs perform short-term temporal modelling on the dynamically generated course resource embeddings X_t from the hypergraph module and the learner behaviour sequence features O_t, outputting hidden state sequences $h_{r,t}$ and $h_{u,t}$, as shown in equations (10) and (11).

$$h_{r,t} = \text{GRU}\left(x_t, h_{r,t-1}\right) \tag{10}$$

$$h_{u,t} = \text{GRU}\left(o_t, h_{u,t-1}\right) \tag{11}$$

The English oral instruction course resource states are merged into the historical learning sequence state to obtain the learner behaviour sequence's implicit state, as shown in equation (12), where $h_{j,t}$ is the state of current resource r_j, and $h_{i,t}$ is the state of u_i.

$$h_t = h_{j,t} + h_{i,t} \tag{12}$$

Based on this implicit state sequence, a transformer-based self-attention encoder is used to capture long-range dependencies within the sequence, enhancing the ability to capture progression evolution. This module consists of multiple identical encoding layers stacked together, each layer containing multi-head self-attention mechanism and position-wise feedforward neural network modules, respectively used for capturing global dependency characteristics among sequence elements and feature transformation. The updated learner embeddings Z are obtained as shown in equation (13).

$$Z = H^{(L)} = \text{TransformerEncoder}\left(H^{(0)}\right) \tag{13}$$

The fully connected layer calculates the unnormalised probability scores for each r_j, as shown in equation (14), where W_p is the weight matrix, and b_p is the bias term.

$$P = ZV_p + b_p \tag{14}$$

Filter out the English oral instruction courses that learners have already studied during generation to create a historical behaviour mask M. The mask marks positions of previously encountered learning resources as $-\infty$. The masked prediction probability P_{masked} is shown in equation (15).

$$P_{masked} = P + M \tag{15}$$

Normalise the recommendation probabilities using the softmax function, sort by probability to generate an initial list L_{init}, as shown in equation (16).

$$\hat{y} = softmax\left(P_{masked}\right) \tag{16}$$

4.3 Graph-enhanced dynamic knowledge tracking mechanism

The knowledge tracing module analyses learners' knowledge mastery states and quantifies their potential for gaining new knowledge. The inputs include the learner's historical data on English oral exercises and problem-solving outcomes, as well as learning resource embeddings X_S generated by GCN to construct a state vector shown in equation (17), where c_k represents answer performance, 0 indicates incorrect answers, 1 indicates correct answers, and d_k denotes the difficulty of practice.

$$s_t = SLTM\left(x \| c_k \| d_k, s_{t-1}\right) \tag{17}$$

Predict the probability that learners will correctly answer questions based on their knowledge states, as shown in equation (18), where σ is a sigmoid function, W_p are learnable parameters, and b_p is a bias term.

$$p_{t+1} = \sigma\left(W_p s_t + b_p\right) \tag{18}$$

Knowledge state prediction loss is calculated using binary cross-entropy loss, as shown in equation (19), where N_{valid} represents the number of effective samples, and $y_{i,j}$ denotes true labels. Weak knowledge points are identified based on prediction results, and knowledge gain $Gain_j$ is quantified as shown in equation (20). In this case, $p(r_j^{after})$ indicates accuracy before the student completes a learning path r_j, while $p(r_j^{before})$ indicates accuracy after completing the path r_j. By adjusting the denominator, knowledge gain for students with different initial levels is made comparable.

$$\mathcal{L}_{BCE} = -\frac{\sum_{i=1}^{|u|}\sum_{r_j \in u_i} L_{i,j}}{N_{valid}} \tag{19}$$

$$L_{i,j} = \left[y_{i,j}\log p_{i,j} + \left(1 - y_{i,j}\right)\log\left(1 - p_{i,\mathrm{j}}\right)\right] \tag{20}$$

$$Gain_j = \frac{p\left(r_j^{after}\right) - p\left(r_j^{before}\right)}{1 - p\left(r_j^{before}\right)} \tag{21}$$

4.4 Generation of oral English teaching paths driven by multi-objective optimisation

Candidate paths are expanded through insertion and swap operations on L_{init} to generate a diverse set of candidate learning paths $C = \{L_1, L_2, \ldots, L_n\}$. NSGA-II multi-objective optimisation is adopted for Pareto optimal path selection, where the multi-objective optimisation problem model is defined as $\max\{f_1(L), f_2(L), f_3(L), \ldots, f_n(L)\}$. Since the initially generated teaching paths are based on predicted learner preferences, unnormalised prediction scores P representing English oral course preference probabilities are derived according to equation (22). The objective function accounting for learner preferences is formulated in equation (23), where $|M|$ indicates the length of a list, with δ_t being the student's current ability value, T denoting the historical window size, ε representing a smoothing term, and Dif_i indicating the difficulty level of the i^{th} English oral teaching course.

$$p_{prefer} = Sigmoid(p) \tag{22}$$

$$f_1(L) = Preference(L) = \frac{\sum_{i=1}^{M}\left(p_{prefer,i}\right)}{|M|} \tag{23}$$

$$\begin{cases} f_2(L) = Adaptiviy(L) = \dfrac{\sum_{i=1}^{M}\left(1-\left|\delta_t - Dif_i\right|\right)}{|M|} \\ \delta_t = \dfrac{\sum_{i=t-T}^{t} Dif_i \cdot r_i}{\sum_{i=t-T}^{t} r_i + \varepsilon} \end{cases} \tag{24}$$

The effectiveness objective function is shown in equation (25), where E_e represents the student's mastery of the target concept after learning path L, E_s denotes the student's mastery of the target concept before learning path L, E_{sup} indicates the upper limit for mastery of the target concept, and $Gain_i$ stands for knowledge gain.

$$\begin{aligned} f_3(L) = E_p(L) &= \frac{E_e - E_s}{E_{sup} - E_s} \\ &= \frac{1}{|M|}\sum_{i=1}^{M} \frac{p\left(r_i^{after}\right) - p\left(r_i^{before}\right)}{1 - p\left(r_i^{before}\right)} == \frac{1}{|M|}\sum_{i=1}^{M} Gain_i \end{aligned} \tag{25}$$

The diversity objective function is shown in equation (26), where r_i, r_j are the i^{th} and j^{th} English oral course resources in path L; x_i, x_j are embedding vectors of r_i, r_j, and sim is cosine similarity.

$$f_4(L) = Diversity(L) = \frac{\sum_{r_i \in L}\sum_{r_j \in L}\left(1 - sim\left(x_i, x_j\right)\right)}{|M||M-1|} \tag{26}$$

Based on crowding distance and non-dominated solutions, the Pareto front $P = \{L_1, L_2, \ldots L_m\}$ is determined, where m represents the number of solutions in the front set. Objective weights are dynamically allocated to determine an overall utility function, and the final recommended path is output as the optimal English oral teaching path L_{final}. The integrated utility function is shown in equation (27).

$$S(L) = \lambda_1 \cdot Preference(L) + \lambda_2 \cdot Adaptivity(L) + \lambda_3 \cdot E_p(L) + \lambda_4 \cdot Diversity(L) \tag{27}$$

where λ_1 represents the weight for the preference objective function, λ_2 denotes the weight for the adaptability objective function, λ_3 indicates the weight for the effectiveness objective function, and λ_4 stands for the weight for the diversity objective function.

5 Analysis of experimental results

This paper uses the English oral teaching process dataset collected from literature (Song et al., 2022), which contains 937 course records, 2,058 student records, and 5,162 teaching process records. To ensure rigor in evaluating the performance of teaching path optimisation design, a stratified sampling strategy is employed to divide samples across datasets into training sets, validation sets, and test sets. The GCN has three network layers, four temporal partitions for the hypergraph, two neural network layers for the hypergraph, an embedding dimension of 64, and two attention heads. The optimiser used is Adam with a learning rate of 0.001 and 60 training epochs. All experiments are conducted using Python as the programming language within the PyTorch framework for building the neural networks. Specifically, Python version 3.7.9, PyTorch version 1.8.0, CUDA version 11.0 are used. All experiments are configured and run on a server equipped with an Ubuntu 18.04 system, AMD Ryzen 7 3800X CPU, and NVIDIA GTX 3090 GPU.

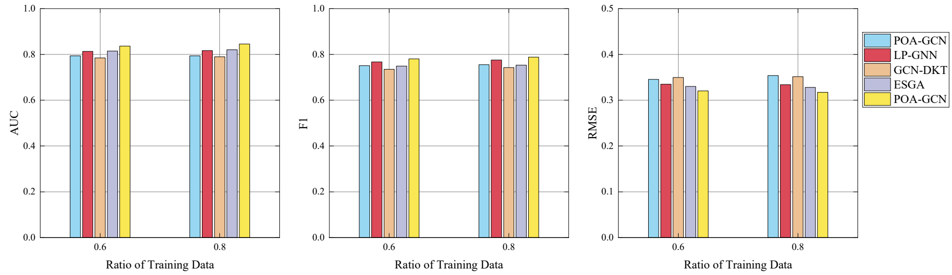
This paper first analyses the performance of the initial path generation prediction for the proposed English oral teaching path design method POA-GCN and baseline methods ESGA (Maimaiti, 2019), SMPOA (Jiang, 2022), LP-GNN (Su and Shen, 2025), and GCN-DKT (Zhang et al., 2025). The length of English oral teaching paths is set to 5, 10, 15, 20, 25, 30. The initial path prediction accuracy rates for different methods are shown in Table 1. With the increase in teaching path length, the prediction accuracy rate improves across all methods. When the path length is 20, the prediction accuracy rates of ESGA, SMPOA, LP-GNN, GCN-DKT, and POA-GCN are 70.05%, 76.18%, 81.42%, 86.97%, and 94.01%, respectively. Compared to ESGA, SMPOA, LP-GNN, and GCN-DKT, the prediction accuracy rate of POA-GCN is improved by 23.96%, 17.83%, 12.59%, and 7.04%, respectively. POA-GCN constructs a hypergraph-graph hybrid architecture to depict semantic associations in English oral teaching processes through static relational graphs, models high-order interaction relationships between learners and courses using the hyperedge mechanism, overcomes the limitation of single edge types in traditional graph models, and designs dual-channel gated recurrent units to capture dynamic learner behaviour evolution, thereby improving the accuracy rate of initial path generation.

Table 1 The prediction accuracy under different lengths of oral English teaching paths

Method	*ESGA*	*SMPOA*	*LP-GNN*	*GCN-DKT*	*POA-GCN*
5	65.22	73.01	77.53	80.02	86.37
10	68.94	74.35	77.96	81.56	88.09
15	69.37	74.97	80.15	83.61	91.52
20	70.05	76.18	81.42	86.97	94.01
25	72.69	77.09	81.97	87.42	94.83
30	74.85	80.48	82.47	89.55	96.22

To further validate the effectiveness of POA-GCN, this paper also selects quantitative metrics for path design, including F1, root mean squared error (RMSE), and AUC, to compare the performance of different methods as shown in Figure 3. POA-GCN maintains optimal performance with training data at different proportions (60%, 70%, 80%), showing excellent results across AUC, F1, and RMSE indicators. With the increase in training data, its prediction performance improves effectively, indicating strong adaptability to datasets of varying scales and the ability to capture associations among teaching paths as well as learners' dynamic behaviours. LP-GNN and GCN-DKT also exhibit good performance, demonstrating that graph neural networks have significant advantages when modelling learning behaviour and POA teaching relationships. The improvements in AUC and F1 for ESGA and SMPOA are relatively small with further increases in training data; even RMSE may rise, which could be due to low model efficiency in processing additional data or mild overfitting.

Figure 3 The predictive performance of different teaching path generation methods, (a) AUC comparison (b) F1 comparison (c) RMSE comparison (see online version for colours)



(a) (b) (c)

Figure 4 The solution times of different methods for optimal English speaking teaching paths

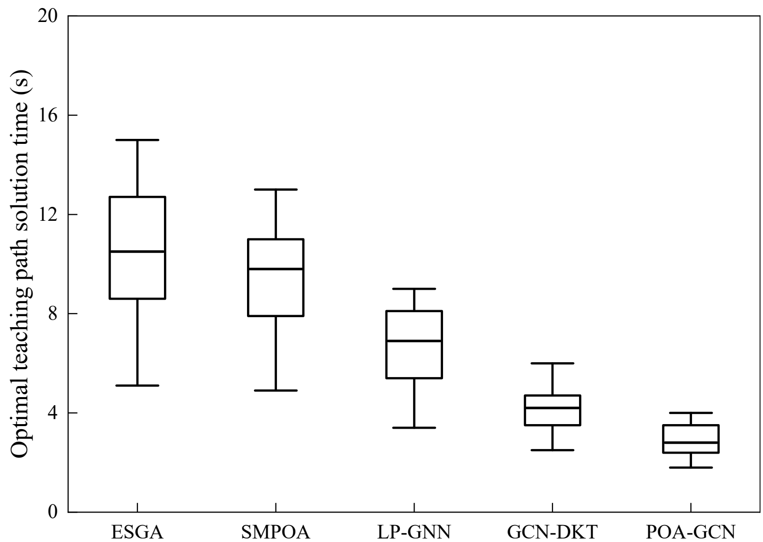


Figure 4 compares the solution times of different methods for optimal English speaking teaching paths. The solution time of POA-GCN for an optimal English speaking teaching path is 1.8 s; ESGA, SMPOA, LP-GNN, and GCN-DKT take 5.1 s, 4.9 s, 3.4 s and 2.5 s respectively to solve the optimal teaching paths. Compared with ESGA, SMPOA, LP-GNN, and GCN-DKT, POA-GCN reduces the solution time by 3.3 s, 3.1 s, 1.6 s and 0.7 s. POA-GCN conducts multi-objective optimisation on multiple paths to obtain a front, and finally acquires the English speaking teaching path with the highest score based on an overall utility function. For learners with different goals, the weights of objective functions in the comprehensive utility function can be adjusted based on the varying degrees of each goal, thereby generating corresponding final teaching paths for learners adopting different strategy objectives according to the Pareto front, significantly improving the solution efficiency of optimal teaching paths.

6 Conclusions

With the acceleration of globalisation and rapid development of artificial intelligence technology, English speaking ability has become a core competency in cross-cultural communication. Traditional teaching models suffer from insufficient extraction of learner behaviour characteristics and inaccurate generation of personalised teaching paths. To this end, this paper proposes an evolutionary path generation method for vocational English speaking based on the synergy of GCN and multi-objective optimisation. First, an online-offline integrated English speaking teaching process based on POA is designed, driven by a production-oriented core engine, and a method for designing English speaking teaching paths that combines learner behaviour preference modelling and POA theory is developed. GCN obtains the initial learning resource embeddings. Then, interactive-based English speaking course resources and learner preference embeddings are obtained through the transposed hypergraph. Combined with dual-channel GRU and transformer modules to extract long-term and short-term features from learners' historical interaction sequences, initial learning paths are generated based on recommendation probabilities output by MLP. LSTM is used for dynamic encoding of fused information in course resource embeddings along with learners' answer records and difficulty levels to predict learner performance. Finally, a multi-objective optimisation module is designed. NSGA-II algorithm is employed for multidimensional optimisation to obtain Pareto front from multiple paths. The final English speaking teaching path with the highest score based on the comprehensive utility function is ultimately determined, effectively enhancing personalisation and accuracy of the teaching paths. Experimental results show that the prediction accuracy reaches 96.22%. The optimal teaching path solving time decreases by 1.6 s. The proposed method not only improves the accuracy rate for generating English speaking teaching paths but also reduces the solving time of the optimal teaching path.

Declarations

All authors declare that they have no conflicts of interest.

References

Anas, I. (2019) 'Educational technology and teacher-student technology competency: a pathway to teaching with technology', *Journal of English Language Teaching and Linguistics*, Vol. 4, No. 2, pp.18–31.

Ati, A. and Parmawati, A. (2022) 'The use of oral presentation in teaching English to improve students speaking skill', *Professional Journal of English Education*, Vol. 5, No. 2, pp.300–305.

Benmesbah, O., Lamia, M. and Hafidi, M. (2023) 'An improved constrained learning path adaptation problem based on genetic algorithm', *Interactive Learning Environments*, Vol. 31, No. 6, pp.3595–3612.

Bhatti, U.A., Tang, H., Wu, G., Marjan, S. and Hussain, A. (2023) 'Deep learning with graph convolutional networks: an overview and latest applications in computational intelligence', *International Journal of Intelligent Systems*, Vol. 20, No. 1, pp.83–104.

Ding, H. (2023) 'Application of production-oriented approach (POA) in college English teaching in normal universities', *Frontiers in Educational Research*, Vol. 6, No. 14, pp.58–66.

Du, J. and Daniel, B.K. (2024) 'Transforming language education: a systematic review of AI-powered chatbots for English as a foreign language speaking practice', *Computers and Education: Artificial Intelligence*, Vol. 6, pp.23–36.

Du, Z. and Su, J. (2021) 'Analysis of the practice path of the flipped classroom model assisted by big data in english teaching', *Scientific Programming*, Vol. 11, No. 1, pp.89–102.

Guo, L., Jia, Z., Li, Q. and Dai, J. (2022) 'Steadiness analysis of means-end conceptual paths and problem-chains based on concept lattices and similarity measuring', *International Journal of Machine Learning and Cybernetics*, Vol. 13, No. 3, pp.691–719.

Hsu, T-C. (2017) 'Learning English with augmented reality: do learning styles matter?', *Computers & Education*, Vol. 106, pp.137–149.

Iatrellis, O., Savvas, I.K., Kameas, A. and Fitsilis, P. (2020) 'Integrated learning pathways in higher education: a framework enhanced with machine learning and semantics', *Education and Information Technologies*, Vol. 25, No. 4, pp.3109–3129.

Jiang, D., Pei, Y., Yang, G. and Wang, X. (2022) 'Research and analysis on the integration of artificial intelligence in college English teaching', *Mathematical Problems in Engineering*, Vol. 10, No. 1, pp.75–89.

Jiang, W. (2022) 'Application of POA theory in innovative teaching model of college English – under the background of curriculum ideological and political education', *Frontiers in Educational Research*, Vol. 5, No. 6, pp.61–74.

Kitson, N.K., Constantinou, A.C., Guo, Z., Liu, Y. and Chobtham, K. (2023) 'A survey of Bayesian network structure learning', *Artificial Intelligence Review*, Vol. 56, No. 8, pp.8721–8814.

Li, Q. (2021) 'Research on the integrated teaching model of college English reading and writing based on the output-oriented approach', *Frontiers in Educational Research*, Vol. 4, No. 9, pp.57–64.

Lindemann, B., Müller, T., Vietz, H., Jazdi, N. and Weyrich, M. (2021) 'A survey on long short-term memory networks for time series prediction', *Procedia CIRP*, Vol. 99, pp.650–655.

Liu, F. (2024) 'Implementing E-learning in English translation teaching using deep learning models and output-oriented methods', *Computer-Aided Design & Applications*, Vol. 21, pp.219–235.

Liu, S. (2023) 'Retracted: Construction and computation of the college English teaching path in the artificial intelligence teaching environment', *Computational Intelligence and Neuroscience*, Vol. 20, pp.98–114.

Liu, X. (2022) 'Construction and implementation path of college English teaching mode in big data environment', *Mathematical Problems in Engineering*, Vol. 6, No. 1, pp.54–76.

Ma, H., Zhang, Y., Sun, S., Liu, T. and Shan, Y. (2023) 'A comprehensive survey on NSGA-II for multi-objective optimization and applications', *Artificial Intelligence Review*, Vol. 56, No. 12, pp.15217–15270.

Maimaiti, K. (2019) 'Modelling and analysis of innovative path of English teaching mode under the background of big data', *International Journal of Continuing Engineering Education and Life Long Learning*, Vol. 29, No. 4, pp.306–320.

Minghe, G. and Yuan, W. (2013) 'Affective factors in oral English teaching and learning', *Higher Education of Social Science*, Vol. 5, No. 3, pp.57–61.

Nigenda, R.S., Padrón, C.M., Martínez-Salazar, I. and Torres-Guerrero, F. (2018) 'Design and evaluation of planning and mathematical models for generating learning paths', *Computational Intelligence*, Vol. 34, No. 3, pp.821–838.

Romero, M.Á.M., Silva, G.J.S., Aguila, O.E.P., Valderas, P., Huallpa, J., Sosa, L. and Arias-Gonzáles, J. (2023) 'Personalized education path for students; a conceptual basis for a digitalized education environment', *Journal of Namibian Studies: History Politics Culture*, Vol. 33, No. 1, pp.394–415.

Shu, J. (2022) 'A POA theory-based network teaching mode for English course in higher vocational college', *International Journal of Emerging Technologies in Learning (iJET)*, Vol. 17, No. 1, pp.224–238.

Song, X., Liu, Y., Xue, L., Wang, J., Zhang, J., Wang, J., Jiang, L. and Cheng, Z. (2020) 'Time-series well performance prediction based on long short-term memory (LSTM) neural network model', *Journal of Petroleum Science and Engineering*, Vol. 186, pp.10–22.

Song, Y., Wei, Y., Shen, Y. and Xu, M. (2022) 'Evaluation of an online oral English teaching model using big data', *Mobile Information Systems*, Vol. 20, No. 1, pp.79–85.

Sritulanon, A., Chaturongakul, P. and Thammetar, T. (2018) 'English speaking teaching model in distance education', *Arab World English Journal (AWEJ) Volume*, Vol. 9, pp.20–31.

Su, X. and Shen, T. (2025) 'Research on path selection and optimization of college English teaching based on deep learning and IoT technology', *Australian Journal of Electrical and Electronics Engineering*, Vol. 3, pp.1–12.

Sun, L., Ismail, H.H. and Aziz, A.A. (2024) 'Current English language teaching using production-oriented approach: a systematic review', *World Journal of English Language*, Vol. 14, No. 4, pp.13–25.

Tapalova, O. and Zhiyenbayeva, N. (2022) 'Artificial intelligence in education: AIEd for personalised learning pathways', *Electronic Journal of E-Learning*, Vol. 20, No. 5, pp.639–653.

Xu, X., Hong, W.C.H., Zhang, Y., Jiang, H. and Liu, J. (2024) 'Learning paths design in personal learning environments: the impact on postgraduates' cognitive achievements and satisfaction', *Innovations in Education and Teaching International*, Vol. 61, No. 4, pp.748–763.

Yang, Y-F., Tseng, C.C. and Lai, S-C. (2024) 'Enhancing teachers' self-efficacy beliefs in AI-based technology integration into English speaking teaching through a professional development program', *Teaching and Teacher Education*, Vol. 144, pp.45–53.

Zhang, G., Gao, X., Ye, H., Zhu, J., Lin, W., Wu, Z., Zhou, H., Ye, Z., Ge, Y. and Baghban, A. (2025) 'Optimizing learning paths: course recommendations based on graph convolutional networks and learning styles', *Applied Soft Computing*, Vol. 175, pp.11–23.

Zhu, A. (2023) 'Path and development strategy of intelligent teaching based on OBE concept', *Frontiers in Educational Research*, Vol. 6, No. 20, pp.14–21.