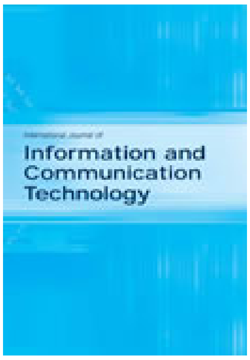


International Journal of Information and Communication Technology

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

Design of college personalised career planning utilising multidisciplinary approaches

Tao Zeng, Jing Zhou, Xiaofang Zeng

DOI: 10.1504/IJICT.2025.10072944

Article History:

Received:	21 June 2025
Last revised:	14 July 2025
Accepted:	15 July 2025
Published online:	08 September 2025

Design of college personalised career planning utilising multidisciplinary approaches

Tao Zeng, Jing Zhou and Xiaofang Zeng*

College of Medicine,
Jingchu University of Technology,
Jingmen 448000, China
Email: 200407018@jcut.edu.cn
Email: 200707023@jcut.edu.cn
Email: 200607011@jcut.edu.cn
*Corresponding author

Abstract: This paper proposes a personalised learning path design method for college students' career planning under an interdisciplinary perspective, aiming to improve the efficiency of learning path generation. Firstly, college students' career planning knowledge graph (CPKG) is constructed, and the set of knowledge points that students have mastered and the target knowledge points are mapped into the CPKG. Then, the neighbourhood expansion rule of graph convolutional neural network is combined to add the association strength of relations to entities to capture learner preferences. The importance of knowledge points is calculated by adjusting the corresponding weights through the features of knowledge points in CPKG, and the optimal personalised learning path is designed for learners by combining learner preferences and the importance of knowledge points. Experimental results show that the proposed method takes only 9.2 s to generate the optimal path, which can provide learners with satisfactory learning paths.

Keywords: learning path design; knowledge graph; learner preference; graph convolutional neural network; knowledge point importance.

Reference to this paper should be made as follows: Zeng, T., Zhou, J. and Zeng, X. (2025) 'Design of college personalised career planning utilising multidisciplinary approaches', *Int. J. Information and Communication Technology*, Vol. 26, No. 32, pp.68–82.

Biographical notes: Tao Zeng received his Master's degree from Wuhan University in 2012. He is currently a Senior Laboratory Technician at the Jingchu University of Technology. His research interests include inflammatory intervention mechanisms, hypertension, diabetes control, and AI applications in emergency rescue.

Jing Zhou received her Master's degree from Chongqing Medical University in 2007. She is currently an Associate Professor at the Jingchu University of Technology. Her research interests include pathology, medical education, and healthcare.

Xiaofang Zeng received her Master's degree from Wuhan University in 2013. She is currently a Lecturer at the Jingchu University of Technology. Her interests include basic oral medicine, oral preventive medicine, and interdisciplinary research.

1 Introduction

As the main force to enter the workplace, college students are faced with many opportunities and challenges. On the one hand, new industries are emerging and the demand for cross-field integration jobs is growing; on the other hand, the connotation and requirements of traditional occupations are also being continuously revolutionised (Gunkel et al., 2010). In this context, single-discipline knowledge reserve has been difficult to meet the demands of college students' career development, and interdisciplinary literacy has gradually become a key element of competition in the workplace. However, most of the current occupation planning education for university students is limited to the conventional path guidance of professional counterparts, and lacks the systematic consideration of the integrated use of knowledge from different disciplines as well as the in-depth exploration of students' individual traits (Serbes and Albay, 2017). This homogenised and patterned approach to education results in many students being confused when making career decisions and having difficulty adapting after entering the workforce. In view of this, it is particularly urgent to explore personalised learning paths for college students' career planning from an interdisciplinary perspective. This not only helps students break through professional barriers, integrate diversified knowledge, and create unique career advantages, but also injects new vitality into vocational education in colleges and universities, accurately empowers students' future career development, and realises the efficient connection between education and employment (Okolie et al., 2020).

Personalised learning paths serve as an alternative path designed to replace a predetermined learning sequence. Current approaches to personalised learning path design fall into three broad categories. The first category is learning path design based on data mining (Hsieh and Wang, 2010). Vanitha et al. (2019) used a collaborative filtering algorithm to select teaching paths used by learners with similar preferences and prior knowledge levels as the learners. Vanitha and Krishnan (2019) used ant colony algorithm (ACO) to optimise teaching paths and enhance teaching adaptability. Shi et al. (2022) gave a studying trajectory planning method relied on swarm intelligence, which can efficiently refer to the experiences of similar populations. Bendahmane et al. (2019) take collaborative recommendation mechanisms into account when performing learning path generation. This approach draws on group experience. Data mining relies on a large amount of data to construct learning path models, and if there are biases in the data sources, such as uneven sample selection that only covers data from a specific group or a specific learning scenario, the mined learning paths may not accurately reflect the needs of all learners. Learners' interests and needs change over time and through personal experience. Data mining models are usually constructed based on historical data. If the dynamic changes in learner interests are not captured in a timely manner, the designed learning paths may not be able to consistently engage the learners, leading to a decline in learning effectiveness.

The second category is learning path design based on resource fitness (Kurilovas et al., 2015). Xie et al. (2017) introduced context-awareness in order to determine the user's style preference, learning status, and so on, to design the optimal path. Researchers such as Benmesbah et al. (2023) used the fit between user style and learning resource type, and the fit between user knowledge and learning resource difficulty as the basis for updating, and combined with genetic algorithms to calculate the learning path. Liu et al.

(2022) constructed an objective function based on the guidelines that the learners' target knowledge points should be maximally included in the learned resources, and the difficulty of the learning resources assigned by the system should be appropriate to the learners' individual abilities, etc., and used an improved particle swarm algorithm (Wang et al., 2018) to generate learning paths, but it may violate the inherent logic between the knowledge points. Pushpa (2012) proposed an ACO based on learning attributes to design the most appropriate learning based on learning styles and learners' knowledge levels.

The third category is knowledge graph (KG)-based learning path design. Based on the learning progress and ability level of college students, suitable practical projects are recommended for them to form personalised practical learning paths. By participating in practical projects, college students can apply the knowledge they have learned to actual work, enhance their practical abilities and problem-solving skills, and at the same time accumulate project experience to prepare for their future careers. Use knowledge graphs to record the learning process of college students, including information such as the knowledge points learned, the learning tasks completed, and the practical projects participated in. By analysing these data, we can understand the learning progress and situation of college students, and promptly identify the problems and deficiencies existing in their studies. Ortiz-Vilchis and Ramirez-Arellano (2023) take the attribute features and relationships of the knowledge points as inputs, and combine them with a topological sorting algorithm to construct a learning path that conforms to the internal logic of the knowledge points, so as to achieve the best learning effect. Davuluri (2021) used convolutional neural network to categorise the test questions and generated learning paths based on Prim's minimum adult tree algorithm. Chahine and Grinshpon (2020) subdivided the main knowledge points and constructed a map of the facets of knowledge to be taught based on the precedence and successor relationships, calculating and comparing the magnitude of the function values on all the alternative paths, and finally selecting the most appropriate path for the scenario. Liu et al. (2023) investigated the learning outcomes of learners using a plain Bayesian classifier to identify learners' weak knowledge points and gave a learning path selection method based on longitudinal combination of KG and antecedent associations.

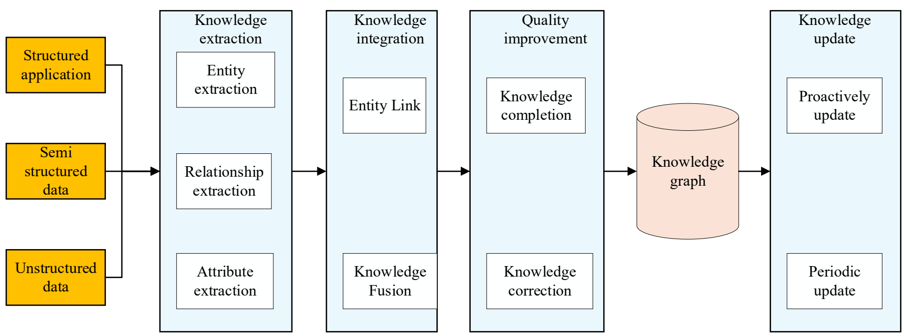
By summarising the current research results, it can be seen that the existing research ignores students' preferences, which leads to the inefficiency of generating learning paths. In order to solve the above problems, this paper first constructs a college students' career planning knowledge graph (CPKG), models the knowledge points of multiple courses involved in career planning and related learning resources as a graph, and establishes the mapping relationships between different entity structures in the resource library. Subsequently, the collection of knowledge points that students have acquired, along with the intended knowledge points, are aligned within the CPKG to produce a personalised directed graph of knowledge points. Then the learner preferences are calculated by changing the neighbourhood expansion rule and by adding the association strength of entity relationships. Meanwhile, on the basis of PageRank algorithm, the KPIRank algorithm is proposed to compute the significance of knowledge points by adding the weight influence of knowledge points and the proximity centrality of nodes in CPKG. Finally, the optimal personalised paths are generated for learners by combining learner preferences and the importance of knowledge points. The experimental outcome indicates that the offered approach generates learning paths with an accuracy of 92.84%, which is improved by 3.19–17.69% compared with other models, and exhibits the optimal learning path generation efficiency.

2 Related works

2.1 Knowledge graph

Knowledge graphs (KGs) are a typical class of structured knowledge representations consisting of entity, relationship and attribute descriptions, which have powerful knowledge modelling and characterisation capabilities and are widely used in the education field (Chen et al., 2018). The entire structure of the KG establishment is implied in Figure 1.

Figure 1 The framework of knowledge graph construction (see online version for colours)



Through preprocessing, the unrefined data is classified into structured, semi-structured, and unstructured categories, and automated or semi-automated approaches are utilised to capture knowledge from the unrefined data to obtain triples containing entities, relationships and attributes. A knowledge integration process is done for triples to link entities, concepts, attributes, etc. by various origins to strengthen knowledge integration. Knowledge is complemented by combining knowledge reasoning to mine the relationships within the graph. By integrating knowledge reasoning, the relationships within the graph are explored and complemented with additional knowledge. To ensure the completeness and accuracy of the knowledge graph at any given moment, a mechanism of active updating, rather than periodic updating, is employed for knowledge updating.

2.2 Graphical convolutional neural network

Graph convolutional neural network (GCN) is a deep learning approach relied on graph data structure. The core structure of a CNN is a convolutional layer in which neurons are connected to only localised regions of the input data. In image processing, for example, a convolutional kernel focuses only on a small square region of the image. GCN works directly on graph data, and its core idea is to update the representation of a node by aggregating the node's own features as well as those of its neighbours. In each layer, each node receives information from its neighbouring nodes and combines it with its own information for fusion and transformation. Compared with the traditional CNN, it can directly perform convolutional operations on the nodes and edges on the graph, so that it can perform various tasks on the graph. The goal of GCN is to compute the embedding vectors of each node by weighting and manipulating the features of the node and its related nodes to perform tasks such as classification, regression, etc. (Zhang et al., 2019). The core idea of GCN is to utilise the local connection information on the graph structure so as to achieve the aggregation and sharing of node features. Suppose there is a batch of graph data, in which there are N nodes, all nodes have its own features, the features of these nodes constitute an $N \times d$-dimensional feature matrix $X = \{x_1, x_2, \ldots, x_n\}$, and then the connection among the nodes constitutes a $N \times N$-dimensional adjacency matrix A, X and A are the inputs to the model.

The convolution operation performed on the graph data can be divided into two parts. First, the initial representation of a node is parametrically linearly transformed using the weight matrix W to obtain the representation of the node as follows:

$$h_i = Wx_i \tag{1}$$

Secondly, the information of the neighbouring nodes around the node is collected by weighting operation to get the output of node i after convolution, and the final representation is shown below.

$$x_i = \sigma\left(\sum_{j \in N_i} \alpha_{ij} h_i\right) \tag{2}$$

where N_i is the set of neighbours of centre node i, α_{ij} is the influence weight coefficient of neighbour node j on centre node i, and $\sigma(\cdot)$is the nonlinear activation function.

3 The construction of knowledge mapping for college students' career planning under interdisciplinary perspective

Since career planning for college students involves multiple disciplines, and factors such as the strong dependence of the discipline KG on manual frameworks and the low frequency of updating the schema layer, the top-down construction model is more suitable. This paper first defines different types of KG entities, including career planning-related courses, knowledge points, teachers, and corporate positions. Among them, to optimise the studying path design for better completion and assist college students organise and memorise knowledge related to career planning logically, the curriculum substance (V) is divided into three dimensions corresponding to KG: conjectural basis (V_B), algorithm/framework (V_A) and project practice (V_T), where $V = V_A \cup V_T \cup V_B$, $V_A \cap V_B \cap V_T = \phi$ is satisfied. To better express the semantic relationship between learning objects, this paper assumes that each relationship is only connected to a specific type of node, and takes these semantic relations as the object property to build the course ontology, and semantic relations are used to link the course entity and the knowledge entity, so as to finally obtain the multidisciplinary career planning ontology including the course level and the knowledge level. Based on the created ontology of college students' career planning, we constructed the college students' career planning KG (CPKG) under the interdisciplinary perspective. In this paper, the knowledge points of multiple courses and related learning resources involved in career planning are modelled as graphs, and the mapping relationships of different entity structures in the resource library are established.

Taking the teaching resource entity knowledge point as a spatial structure and denoting it as (h, r, t), the transformation process of entity information in space can be expressed as follows.

$$g(h, t) = -\|h + r - t\| \tag{3}$$

where g stands for the translation function of entity information in space; h stands for head entity; r stands for entity translation; and t stands for tail entity. In order to realise the entity-based representation of KGs and establish the semantic relationship between different knowledge points, the CPKG shown in equation (4) is designed using the transformed entity information.

$$R = g(h, t)h^t Kt \tag{4}$$

where R denotes the CPKG in the interdisciplinary perspective; K denotes the entity dimension corresponding to the knowledge point.

4 Designing personalised learning pathways for university students' career planning from an interdisciplinary perspective

4.1 Personalised knowledge points directional map

For the goal of fully considering the preferences of college students' career planning, the group of assimilated knowledge points and the key studying targets are firstly mapped into CPKG to obtain a personalised knowledge point directed graph containing the beginning point and the final point. Then the neighbourhood expansion rules in CPKG and GCN neighbourhood aggregation operations are combined and the strength of association of relations to entities is added to capture learner preferences more accurately. Based on the characteristics of the knowledge points in CPKG, the corresponding weights are adjusted to compute the significance of the knowledge points. Combined with learner preferences and the significance of knowledge points, the optimal personalised learning path is generated for learners, and the entire flow of the offered approach is shown in Figure 2.

In this paper, the constructed CPKG mapping generates vertices in a directed graph of knowledge points. Dependency relationships between knowledge entities are represented as unilateral connection relationships in the directed graph of knowledge points, and dependency relationships between knowledge entities are represented as bilateral connection relationships in the directed graph of knowledge points; congruent, irrelevant and other relationships between knowledge entities are represented as unbounded connection relationships in the directed graph of knowledge points.

Subsequently, according to the learning information of college students through the online learning platform, the corresponding knowledge points are included in the cluster of knowledge items C that students have owned, while the intended learning objectives p of students can be obtained by querying the relevant information of students. By labelling all the mastered knowledge points in set C as the beginning point and p as the end point in the CPKG structured as a directed graph, this paper can get a personalised knowledge point directed graph belonging to the university students. For example, according to the practice response behaviour of student S in the learning platform, it can be judged that the knowledge point set C_S that student S has mastered includes knowledge points v_1 and v_2, and the target knowledge point is p_s, and then the personalised knowledge point digraph G_s of the student is generated. In order to facilitate the description, only some knowledge points in G_s are selected for display, as shown in Figure 3.

Figure 2 The entire flow of the offered approach (see online version for colours)

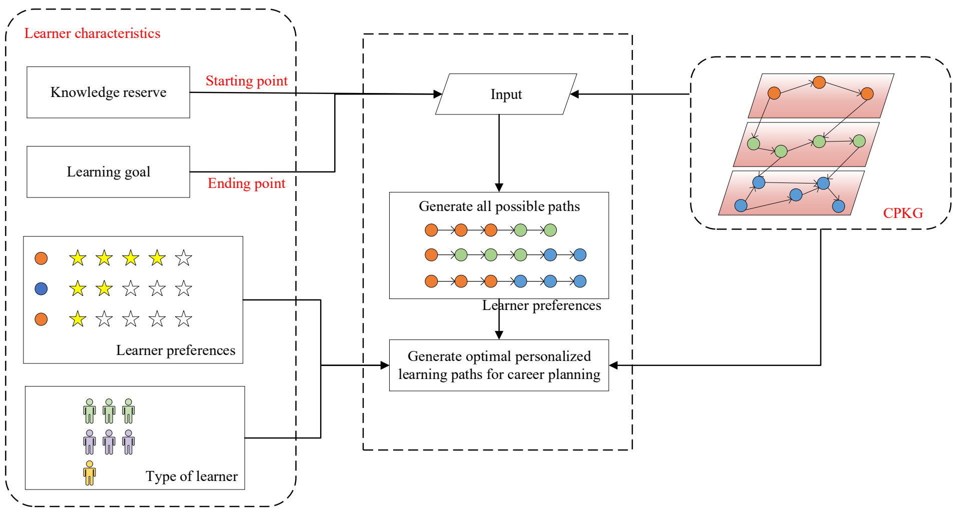
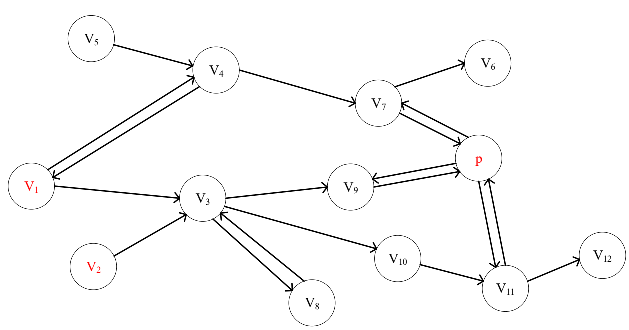


Figure 3 Personalised knowledge points directional map (see online version for colours)



4.2 *Preference mining for college students based on KG and GCN*

After mapping CPKG into directed graph, this paper fully exploits the preferences of college students. By combining the CPKG and GCN algorithms to obtain college students' preference for course knowledge points, and taking the importance of relationship r to college students u (u-r) as the condition for choosing neighbourhoods, we propose a specific neighbourhood selection algorithm based on CPKG and GCN (CPKG-CNH), as shown in Figure 4. The strength of association between user relationships is shown in equation (5), where the operation g is the scalar product between the vectors, and π_r^u is *u-r* association's strength. In this article, the first *k* neighbours exhibiting higher levels of π_r^u are selected for expansion, as implied in equation (6), where $\mathcal{N}(v)$ is the set of neighbours linked with entity *v*, $\mathcal{N}_k(v)$ is the first *k* neighbours chosen, and $s\pi_r^u$ is the assembly of connection forces of *u* to all relations in *v*'s neighbours, as implied in equation (7).

$$\pi_r^u = g(u, r) \tag{5}$$

$$\mathcal{N}_k(v) = top_k\left(s\pi_r^u, k\right) \tag{6}$$

$$s\pi_r^u = \bigcup_{e_i \in \mathcal{N}(v)} \pi_{r_v, e_i}^u \tag{7}$$

To establish the strength of the *u-r* relationship within the range of (0, 1), this paper standardises all association scores pertaining to v and computes the corresponding *u-r* relationship strength $\tilde{\pi}$ for each link as follows:

$$\tilde{\pi}_{r_{v,e}}^u = \frac{\exp\left(\pi_{r_{v,e}}^u\right)}{\sum_{e \in \mathcal{N}_k(v)} \exp\left(\pi_{r_{v,e}}^u\right)} \tag{8}$$

Since only the association strength scores of *u-r* are weighted in the CPKG-CNH neighbourhood expansion, the association strength between entities-relationships (*e-r*) in the triad is ignored. Existing methods for calculating correlation strength include TransD (Zhang et al., 2020), TransE (Song et al., 2021) and TransH (Wang et al., 2017). However, TransD and TransE models is capable of addressing only the singular triplet correspondence, and can not deal with the one-to-many/many-to-one/many-to-many relationships in the knowledge graph well. Therefore, the CPKG-CNH algorithm (KGCN + TransH) is proposed to compute the *e-r* relation strength using the TransH model idea. As implied in equation (9), the related score *g* of connection *r* to entity *e* in the triad is computed, and as *g* decreases, the strength of the *e-r* association increases.

$$g_r(h, t) = \left\| h_\perp + d_r - t_\perp \right\|_2^2 \tag{9}$$

where $h_\perp$ and $t_\perp$ are the representations of the entity on the hyperplane after mapping, d_r is the distance from the entity to the hyperplane, W_r is the common vector to the hyperplane and $||W_r||_2 = 1$.

$$\begin{cases} h_\perp = h - w_r^{\mathrm{T}} h w_r \\ t_\perp = t - w_r^{\mathrm{T}} t w_r \end{cases} \tag{10}$$

As π_r^u increases, the strength of the *u-r* correlation is demonstrated to rise. Thus, for consistency in trend, g is modified by applying a sigmoid function, and the correlation score *s* of *e-r* is proposed as in equation (11) and normalised based on equation (8) as in equation (12).

$$s_{r_{h,t}} = sigmoid\left(\frac{1}{1 + g_r(h, t)}\right) \tag{11}$$

$$\tilde{s}_{r_v,e} = \frac{\exp\left(s_{r_v,e}\right)}{\sum_{e \in \mathcal{N}_k(v)} \exp\left(s_{r_v,e}\right)} \tag{12}$$

To capture the local structure surrounding v, the paper integrates the intensity of the *e-r* relationship with the intensity of the *u-r* relationship for computation, as outlined in equation (13), in which *e* denotes the embedded representation of adjacent entities, w_1 is the ratio of *u-r* to *e-r* association strength, and w_2 is likewise for another corresponding relationship and $w_1 + w_2 = 1$.

$$v_{\mathcal{N}_k(v)}^u = \sum_{e \in \mathcal{N}_k(v)} \left(w_1 \tilde{\pi}_{r_v,e}^u + w_2 \tilde{s}_{r_v,e} \right) e \tag{13}$$

Finally, the entity v and its neighbourhood representation $v_{\mathcal{N}_k(v)}^u$ are computed by SUM aggregation to revise the depiction of v as illustrated hereinafter, in which w_0 and b_0 are weights and biases, individually, and σ is a nonlinear function.

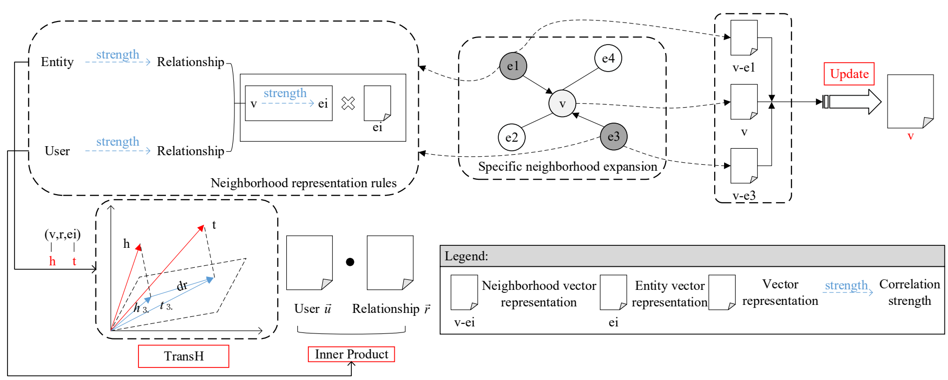
$$V = \sigma\left(w_0 \left(v + v_{\mathcal{N}_k(v)}^u \right) + b_0 \right) \tag{14}$$

This paper considers the learner-knowledge point interaction matrix Y, alongside the knowledge graph G, both of which are derived from learner behaviours, the CPKG-CNH approach is adopted to extract knowledge point features and user interests. Equation (15) is used to predict u's interest in knowledge point LO_i, which is uninteracted thus far, $pref_{LO_i}$. To map the students' favourite choices between (0, 1), a sigmoid operation is used to normalise the final students' favourite choices.

$$pref_{LO_i} = sigmoid\left(f\left(u, LO_i\right)\right) \tag{15}$$

where LO_i is the updated entity illustration.

Figure 4 The structure of CPKG-CNH (see online version for colours)



4.3 Modelling the importance of multidisciplinary knowledge points

In addition to mining learners' preferences, the calculation of the importance of multidisciplinary knowledge points is also an important factor influencing the design of learning paths. In this paper, the PageRank algorithm (Kamvar et al., 2004) is used to rank the importance of knowledge points. If knowledge points *A* and *C* contain relationships that point to knowledge point *B*, then the importance of knowledge point *B* is decided by the significance of knowledge points *A* and *C* as below:

$$PR(K) = \sum_{i=1}^{n} \frac{PR(Y_i)}{C_{out}(Y_i)} \tag{16}$$

where $PR(K)$ is the PageRank value of knowledge point *K*, $PR(Y_i)$ is the PageRank value of Y_i that links to knowledge point *K*, and $C_{out}(Y_i)$ is the number of outgoing links of *n*.

Since the traditional PageRank algorithm ignores the semantic relationship between knowledge points, the same weight setting also ignores the actual need for jumping to knowledge points and the level of knowledge points themselves in the calculation process. To cope with the above issues, this article offers a knowledge point importance ranking algorithm, KPIRank, which adds knowledge point importance level weights and difficulty weights to the PR value calculation, as shown in equation (17).

$$PR(k)' = \sum_{m \in Parents(k)} \frac{\left(\dfrac{W(k)}{\sum_{n \in Children(m)} W(n)} + \dfrac{D(k)}{\sum_{n \in Children(m)} D(n)} \right)}{2} PR(K) \tag{17}$$

To prevent the PR value of 0 caused by the out-degree of a knowledge point in CPKG and the absolute averaging of semantic jump relations caused by the same access probability, this paper introduces the proximity centrality of nodes. The proximity centrality of a node can be used as a measure of the importance of the node. According to equation (18), CC_k is the proximity centrality of *k* nodes, d_i is the average distance from node *i* to the rest of the points, and the reciprocal of the average distance is the proximity centrality, which ultimately yields the importance of knowledge point *k*, as shown in equation (19).

$$\begin{cases} CC_k = \dfrac{1}{d_i} \\ d_i = \dfrac{1}{N-1} \displaystyle\sum_{j=1}^{N} d_{ij} \end{cases} \tag{18}$$

$$KG\_PR(k) = (1-\sigma) \times CC_k + \sigma PR(k)' \tag{19}$$

4.4 Optimal personalised learning path design

The final knowledge point match is obtained by combining the college students' preference for knowledge points $pref_{LO_i}$ and the importance of knowledge points $KG\_PR(k)$ as follows:

$$Rec_{LO_i} = w_{pref} \times pref_{LO_i} + w_{import} \times Import_{LO_i} \tag{20}$$

where w_{pref} and w_{import} represent the weights of learner preference as well as course importance, respectively.

Based on equation (20), all paths' score is gained, and the highest-scoring path is chosen as the designed optimal learning path to be recommended to college students.

$$score_{P_i} = \frac{\sum_{j=1}^{N} Rec_{LO_j}}{N} \tag{21}$$

5 Experimental results and analyses

The dataset used in this paper comes from the career planning data of college students in a large recruitment website and the course learning data, which contains 150 courses, 3,271 knowledge points, 342 vocational skill positions, and 4,193 records of hierarchical relationship between knowledge points. 70% of this dataset is used as a training set, 20% as a test set and 10% as a validation set. The initial learning rate in the experiment is 0.01, the batch size is 64, and the experimental environment is based on the hardware device Intel Core 4.2 GHz i7 7700K CPU and NVIDIA GeForce GTX 2080 Ti GPU, the software device Win10 operating system, the Python3.6 runtime environment, and the use of the open source framework Pytorch to build the neural network model was built using the open-source framework Pytorch.

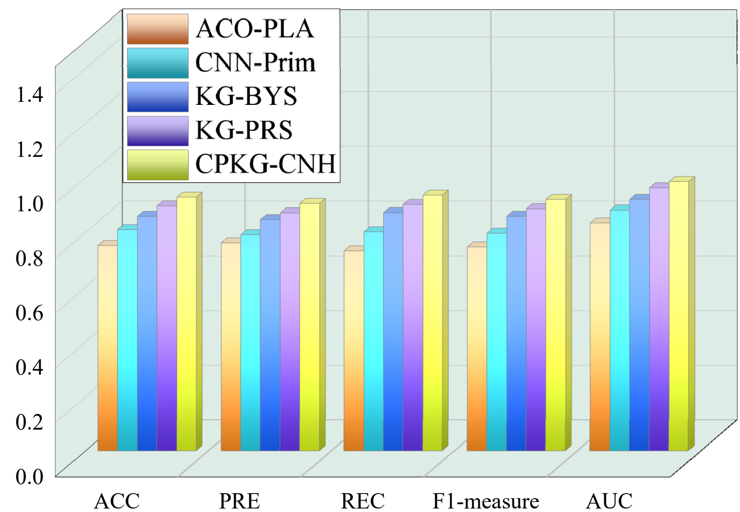
In order to validate the effectiveness of the personalised learning path CPKG-CNH designed in this paper, the commonly used evaluation metrics of path generation efficiency, knowledge level improvement IKL (Li et al., 2020), accuracy (ACC), precision (PRE), recall (REC), F1-measure, and AUC values are selected to evaluate the CPKG-CNH as well as the comparison methods ACO-PLA (Pushpa, 2012), CNN-Prim (Davuluri, 2021), KG-BYS (Chahine and Grinshpon, 2020), and KG-PRS (Liu et al., 2023) for comparison experiments. The IKL for path lengths N of 5, 10, 15, 20, 25, 30, 35 and 40 are taken in all models and the experimental results are shown in Table 1. With the increasing path length, the IKL of all five models showed increasing until the optimum, and then stabilised with a slightly decreasing trend. The difference is that for ACO-PLA, CNN-Prim, and KG-BYS, IKL performs best when the path length *N* is taken around 35, 30 and 25, respectively. For both KG-PRS and CPKG-CNH models, IKL performs best when the path length *N* is taken around 20. Combining all the models to obtain the best performing IKL value, the IKL value of CPKG-CNH is 0.4247, which is at least 5.70% better compared to the other four models, indicating that CPKG-CNH has better path generation.

Table 1 IKL comparison results for different path lengths

Method	*5*	*10*	*15*	*20*	*25*	*30*	*35*	*40*
ACO-PLA	0.1625	0.1842	0.2193	0.2264	0.2519	0.2545	0.2719	0.2635
CNN-Prim	0.1815	0.2076	0.2358	0.2516	0.2791	0.2843	0.2611	0.2635
KG-BYS	0.2396	0.2571	0.2814	0.3018	0.3547	0.3325	0.3295	0.3337
KG-PRS	0.2601	0.2954	0.3283	0.4018	0.3914	0.3901	0.3891	0.3917
CPKG-CNH	0.2703	0.3158	0.3476	0.4247	0.4102	0.3984	0.3916	0.3942

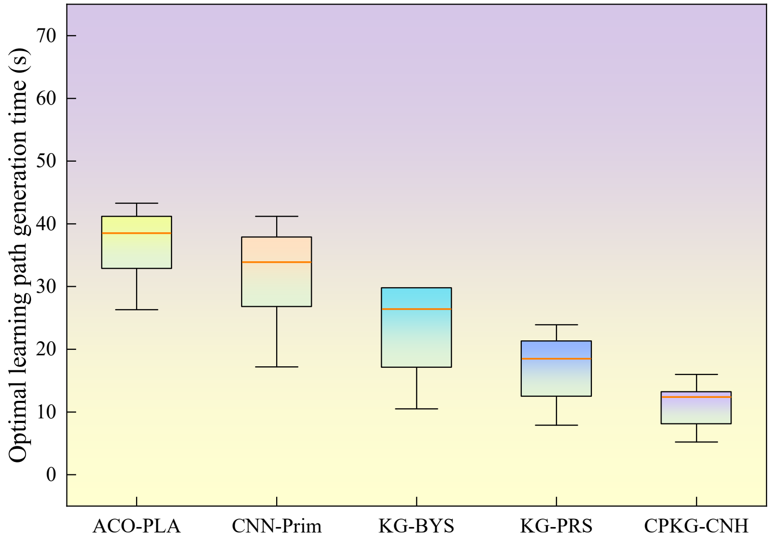
The ACC, PRE, REC, F1-measure and AUC pairs for different learning path design methods are shown in Figure 5. The learning path generation accuracy of CPKG-CNH is 0.9284, which improves 17.69%, 11.8%, 6.9%, and 3.19% compared to ACO-PLA, CNN-Prim, KG-BYS, and KG-PRS, respectively. The F1-measure is the reconciled average of PRE and REC, which best reflects the accuracy of path generation, and the F1-measure of CPKG-CNH improves at least 3.82–23.28% compared to the other four models. AUC is the area of the offline surface of the ROC, and the closer its value is to 1, the higher the accuracy of learning path generation, and the AUC of CPKG-CNH is 0.9861, which is closest to 1, indicating a higher accuracy rate. In summary, CPKG-CNH can capture college students' interests more accurately and provide a basis for personalised learning path design.

Figure 5 Comparison of performance metrics of different learning path design methods (see online version for colours)



In addition to analysing the quantitative performance metrics mentioned above, this paper also compares the learning path generation time of different methods, as shown in Figure 6. The total elapsed time of CPKG-CNH fluctuates up and down no more than 5.1 s, with a maximum of only 16 s, while the elapsed time of ACO-PLA fluctuates the most, with a maximum elapsed time of 43.3 s. KG-PRS has the next best time performance, with the longest time being only 23.9 s. CNN-Prim and KG-BYS have the longest time of 41.2 s and 29.7 s, respectively. CPKG-CNH has the lowest time-consuming learning path generation, which is due to the fact that ACO-PLA searches for the optimal learning path through ACO without considering the preferences of college students and the importance of the knowledge points, which results in a long time-consuming method. CNN-Prim Although CNN is utilised to classify various types of knowledge points, the algorithm of Prim minimum adult tree is time consuming. KG-BYS and KG-PRS are both based on KG in order to find the optimal learning path, and both consider the importance of knowledge point modelling, neither of them consider the interest of college students, and the generation of learning paths is not as efficient as that of CPKG-CNH. Therefore, through observation, it can be seen that the numerical distribution of CPKG-CNH is more centralised than that of the other methods, which suggests that the stability of CPKG-CNH is better.

Figure 6 Comparison of learning path generation time for different methods (see online version for colours)



6 Conclusions

Focusing on the issue that the existing personalised learning path design methods for college students' career planning ignore learner preferences and lead to low efficiency of learning path generation, CPKG is first constructed and CPKG entities from an interdisciplinary perspective are defined. The course ontology is constructed with these semantic relations as object properties, and the semantic relations are used to link the course entity with the knowledge point entity. Finally, the CPKG ontology including the course level and knowledge level is obtained. On this basis, the set of knowledge points that students have mastered and the target knowledge points are mapped to the CPKG to obtain a personalised knowledge point directed graph containing the starting point and the end point. Next, learner preferences are computed by varying the neighbourhood expansion rule and by adding the association strength of entity relationships. Meanwhile, on the basis of PageRank algorithm, the KPIRank algorithm is proposed to compute the significance of knowledge points by adding the weight influence of knowledge points and the proximity of nodes to the centre in the graph. Finally, the optimal personalised paths are generated for learners by combining students' preferences and the significance of knowledge points. The experimental results show that the proposed method not only improves the accuracy of learning path generation, but also reduces the generation time, and can quickly and accurately offer students with the most matching personalised studying paths.

To achieve highly personalised learning path design, it is necessary to comprehensively and accurately collect a variety of information about college students, including learning interests, ability levels, personality traits, and career goals. However, in practice, it is difficult to obtain complete and accurate information through existing data collection means. For example, students may provide untrue information in questionnaires for various reasons, or some potential traits are difficult to measure through conventional methods, which will affect the accurate design of personalised learning paths. In the future, a variety of data collection methods, such as learning behaviour analysis, psychological assessment, and social network analysis, will be used comprehensively to obtain more comprehensive and accurate information about college students. At the same time, advanced data mining and machine learning technologies will be used to deeply mine and analyse multi-source data to uncover students' potential learning needs and career development tendencies, providing a more reliable basis for personalised learning path design.

Declarations

This work is supported by the Provincial Teaching Reform Research Project for Undergraduate Universities in Hubei Province named: Exploration of the 'Four Integration' Talent Training Model in the Medical and Health Care Direction of Local Colleges and Universities from the Perspective of Great Health (No. 2023-504), and the Research on the Training Strategy of Medical and Health Care Talents in Local Colleges and Universities from the Perspective of Great Health (No. JX2023-005).

All authors declare that they have no conflicts of interest.

References

Bendahmane, M., El Falaki, B. and Benattou, M. (2019) 'Toward a personalized learning path through a services-oriented approach', *International Journal of Emerging Technologies in Learning*, Vol. 14, No. 15, p.52.

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.

Chahine, I.C. and Grinshpon, M. (2020) 'Using knowledge space theory to delineate critical learning paths in calculus', *International Journal of Learning, Teaching and Educational Research*, Vol. 19, No. 3, pp.123–148.

Chen, P., Lu, Y., Zheng, V.W., Chen, X. and Yang, B. (2018) 'Knowedu: a system to construct knowledge graph for education', *IEEE Access*, Vol. 6, pp.31553–31563.

Davuluri, M. (2021) 'AI in education: personalized learning pathways using machine learning algorithms', *International Meridian Journal*, Vol. 3, No. 3, pp.1–14.

Gunkel, M., Schlaegel, C., Langella, I.M. and Peluchette, J.V. (2010) 'Personality and career decisiveness: an international empirical comparison of business students' career planning', *Personnel Review*, Vol. 39, No. 4, pp.503–524.

Hsieh, T-C. and Wang, T-I. (2010) 'A mining-based approach on discovering courses pattern for constructing suitable learning path', *Expert Systems with Applications*, Vol. 37, No. 6, pp.4156–4167.

Kamvar, S., Haveliwala, T. and Golub, G. (2004) 'Adaptive methods for the computation of PageRank', *Linear Algebra and its Applications*, Vol. 386, pp.51–65.

Kurilovas, E., Zilinskiene, I. and Dagiene, V. (2015) 'Recommending suitable learning paths according to learners' preferences: experimental research results', *Computers in Human Behavior*, Vol. 51, pp.945–951.

Li, J., Li, Y-x., Tian, S-s. and Xia, J-l. (2020) 'An improved cuckoo search algorithm with self-adaptive knowledge learning', *Neural Computing and Applications*, Vol. 32, pp.11967–11997.

Liu, C., Zhang, H., Zhang, J., Zhang, Z. and Yuan, P. (2023) 'Design of a learning path recommendation system based on a knowledge graph', *International Journal of Information and Communication Technology Education (IJICTE)*, Vol. 19, No. 1, pp.1–18.

Liu, X., Zhang, D., Zhang, T., Zhang, J. and Wang, J. (2022) 'A new path plan method based on hybrid algorithm of reinforcement learning and particle swarm optimization', *Engineering Computations*, Vol. 39, No. 3, pp.993–1019.

Okolie, U.C., Nwajiuba, C.A., Binuomote, M.O., Ehiobuche, C., Igu, N.C.N. and Ajoke, O.S. (2020) 'Career training with mentoring programs in higher education: facilitating career development and employability of graduates', *Education + Training*, Vol. 62, No. 3, pp.214–234.

Ortiz-Vilchis, P. and Ramirez-Arellano, A. (2023) 'Learning pathways and students performance: a dynamic complex system', *Entropy*, Vol. 25, No. 2, p.291.

Pushpa, M. (2012) 'ACO in e-Learning: towards an adaptive learning path', *International Journal on Computer Science and Engineering*, Vol. 4, No. 3, p.458.

Serbes, M. and Albay, M. (2017) 'Importance of career planning and development in education', *International Journal of Social Sciences & Educational Studies*, Vol. 4, No. 2, pp.149–154.

Shi, Z., Wu, Z., Zhang, Z., Chen, Y. and Liu, X. (2022) 'Learning path planning algorithm based on career goals and artificial intelligence', *International Journal of Emerging Technologies in Learning (IJET)*, Vol. 17, No. 10, pp.256–272.

Song, D., Zhang, F., Lu, M., Yang, S. and Huang, H. (2021) 'DTransE: distributed translating embedding for knowledge graph', *IEEE Transactions on Parallel and Distributed Systems*, Vol. 32, No. 10, pp.2509–2523.

Vanitha, V. and Krishnan, P. (2019) 'A modified ant colony algorithm for personalized learning path construction', *Journal of Intelligent & Fuzzy Systems*, Vol. 37, No. 5, pp.6785–6800.

Vanitha, V., Krishnan, P. and Elakkiya, R. (2019) 'Collaborative optimization algorithm for learning path construction in E-learning', *Computers & electrical Engineering*, Vol. 77, pp.325–338.

Wang, D., Tan, D. and Liu, L. (2018) 'Particle swarm optimization algorithm: an overview', *Soft Computing*, Vol. 22, No. 2, pp.387–408.

Wang, Q., Mao, Z., Wang, B. and Guo, L. (2017) 'Knowledge graph embedding: a survey of approaches and applications', *IEEE Transactions on Knowledge and Data Engineering*, Vol. 29, No. 12, pp.2724–2743.

Xie, H., Zou, D., Wang, F.L., Wong, T-L., Rao, Y. and Wang, S.H. (2017) 'Discover learning path for group users: a profile-based approach', *Neurocomputing*, Vol. 254, pp.59–70.

Zhang, S., Sun, Z. and Zhang, W. (2020) 'Improve the translational distance models for knowledge graph embedding', *Journal of Intelligent Information Systems*, Vol. 55, pp.445–467.

Zhang, S., Tong, H., Xu, J. and Maciejewski, R. (2019) 'Graph convolutional networks: a comprehensive review', *Computational Social Networks*, Vol. 6, No. 1, pp.1–23.