Analysis of the dilemma of higher vocational thinking education in China under the background of “Internet+”

Yanhua Hu¹, Banghua Xu¹†

1. School of Marxism, Shenzhen Polytechnic, Shenzhen, 518055, China

Abstract

Network culture is a culture based on computer network information technology. With the popularity of “Internet+”, the influence of network culture on college students’ ideological views, behavioral patterns, value orientations and psychological development is also increasing. While Internet culture brings opportunities to higher education ideological and political education, it also makes education work face serious challenges. This paper proposes a collaborative knowledge graph construction method for Civic Education based on Internet + group intelligence. The core of the method is a continuously running loop that contains 3 activities: free exploration, automatic integration, and active feedback. The experimental results show that the knowledge graph fusion algorithm can effectively utilize the structural information of the knowledge graph as well as the semantic information of the nodes to form a high-quality knowledge graph fusion scheme. The collaborative approach based on the “explore-fusion-feedback” loop can improve the scale of group construction of knowledge graphs and the efficiency of individual construction of knowledge graphs, and exhibit better scalability of group scale.

Keywords: network culture; ideological and political education; human group intelligence; Internet+; knowledge graph

AMS 2020 codes: 68T05
1 Introduction

Higher vocational education focuses on cultivating the ability to work in front-line positions and places great importance on the intensive training of vocational skills, emphasizing both theory and practice, and combining engineering, with a large proportion of practical training teaching in students’ training plans. Moreover, higher vocational education places more emphasis on the openness of schooling, and the school system is relatively flexible in that teaching programs can be jointly formulated by schools and industries and enterprises to determine the cultivation requirements of senior skilled talents. Teachers and students go to industries or enterprises for internship according to the teaching plan, and experienced engineers and technicians from industries and enterprises teach in schools regularly in order to implement two-way communication. In the teaching process, try to make use of the conditions and facilities already available in the industry or enterprises to establish professional top-up internship and practical training bases for students. In the skills assessment, jointly organized with industry and business professionals, the implementation of the vocational qualification certificate system recognized by industry and business. At the same time, since a considerable number of higher vocational colleges and universities are upgraded from the original secondary colleges and universities, the higher education stage is still relatively short, which makes many higher vocational colleges and universities lack the experience and tradition of systematic thinking and political education for college students. Facing the new situation, new tasks and new challenges, there are still many weak links in the ideological and political education work of higher education institutions, and many aspects such as working methods, approaches and mechanisms need to be further improved and perfected, and the effectiveness needs to be further enhanced. How to make full use of “Internet+” and actively face the opportunities and challenges brought by the network culture environment to the ideological and political work of higher education institutions is a topic to be studied and solved in front of the ideological and political education workers in higher education institutions.

With the development of “Internet+”, an important task of knowledge representation is to provide a unified description framework for the huge amount of data and information existing in the Internet, so as to facilitate the structured representation, interconnection and sharing of large-scale knowledge [1-2]. Compared with earlier knowledge representations, modern knowledge graphs (e.g., Freebase, Yago, Wikidata, etc.) all weaken the requirement for logical semantic representation and emphasize large-scale fact-based knowledge. Among them, resource description framework (RDF) is a mainstream representation of factual knowledge, i.e., it represents the entities and their relationships in the knowledge graph through the form of <subject, predicate, object> triad. Also, the semantic information of RDF is described in a lightweight manner by means of RDF paradigm (RDF schema), metadata (metadata), and so on [3-4].

Knowledge graphs formed by manual methods of construction have high accuracy, usability and credibility [5]. However, limited by the individual ability of the constructor, this approach suffers from narrow knowledge coverage and slow update [6-7]. Although “Internet+” crowdsourcing has greatly increased the scale of knowledge graph construction, there is still a strong reliance on a small core group of experts in this approach. For example, inconsistencies between data submitted by different users still need to be adjudicated by core community members [8].

Automated construction algorithms for knowledge graphs can be broadly classified into two categories, rule-based and statistical-based [9-12]. In rule-based construction algorithms, knowledge extraction, fusion, and complementation rules applicable to a specific dataset need to be given in advance by domain experts, and then the algorithm applies these rules to the specific dataset to form a knowledge graph [13-16]. Statistical-based construction algorithms, on the other hand, automatically identify the statistical features of domain-specific data sources and automatically
complete the construction of knowledge graphs [17-18]. At present, the mainstream statistical-based automated construction algorithms generally adopt supervised learning approach, which relies on large scale training datasets manually labeled in advance, and different training datasets need to be built for different problem domains. For the problem of sparse sample data in open domains, some scholars have also explored the use of weakly supervised learning for automated construction of knowledge graphs [19-22].

Automated algorithms have improved the efficiency and reduced the cost of constructing knowledge graphs to some extent, but there are still two fundamental problems. (1) Automated algorithms, especially knowledge graph construction algorithms using supervised learning, heavily depend on the size and quality of the training dataset [23-24]. (2) In the foreseeable future, the ability to understand general unstructured knowledge possessed by automated algorithms is still far from the ability of individual humans, which largely limits the scope of application of automated algorithms. The knowledge graph used in the Google search engine contains a large amount of information about the knowledge graph constructed by manual means in the Freebase project [25-26]. Several research works have also shown that the inclusion of human feedback in the automated construction of knowledge graphs can significantly improve the quality of the construction of knowledge graphs [27-30].

In the free exploration activity, each participant independently performs the knowledge graph construction activity. In the automatic fusion activity, the individual knowledge graphs of all participants are fused together in real time to form a group knowledge graph. In the active feedback activity, the supporting environment recommends specific knowledge map fragment information to each participant based on that participant’s individual knowledge map and the group knowledge map at the current moment to improve its efficiency in constructing the knowledge map. For these three activities, a hierarchical individual knowledge map representation mechanism is established, an individual knowledge map fusion algorithm with the goal of minimizing generalized entropy is proposed, and two types of context-independent and context-relevant information feedback methods are designed. In order to verify the feasibility of the proposed method and key technologies, three types of experiments are designed and implemented: simulation graph fusion experiments containing only structural information, fusion experiments of large-scale real knowledge graphs, and collaborative construction experiments of real knowledge graphs.

2 Automatic fusion activities

2.1 Entropy of the population knowledge graph

Given the set of individual knowledge graphs \( G = \{G_i = (V_i, T_i, E_i) \mid i \in [1, N]\} \), the fusion scheme \( M \in M(G) \), and the group knowledge graph \( G(M) = (V, E, T, edges) \) of \( G \) under \( M \), the entropy \( H(G) \) of \( G \) is calculated as follows.

\[
H(G) = \sum_{m \in M} H(m) \cdot (1 + \theta(m))
\]  

\[
\theta(m) = \Delta - \Delta^{[m]/N}
\]
The entropy $H(G)$ of the population knowledge graph is equal to the sum of the regularized entropy $H(m) \cdot (1 + \theta(m))$ of all nodes $m \in V$ in the population knowledge graph (Equation (1)). where $H(m)$ is the entropy of node $m$; $(1 + \theta(m))$ is the corresponding regularization coefficient. The definition of $\theta(m)$ is given in Equation (2). Where $\Delta$ is a constant; $|m|_N$ denotes the number of $N$ components of $m$ for which $N$ is not equal to $\perp$. $H(m)$ is defined in Equation (3), which measures the mean value of the entropy generated by the participation of node $m$ in the relationship with different roles.

$$H_4(m | Rol) = \frac{\sum_{\text{Role} \in V(\nLeftarrow)} H_4(m | Rol)}{\sum_{\text{Role} \in V(\nLeftarrow)} 1(\text{ref}(\text{Role} \nLeftarrow \nrightarrow \nrightarrow \nrightarrow m) > 0)}, \quad \ell(m) = 4 \text{ and } ref(\nLeftarrow \nrightarrow \nrightarrow \nrightarrow m) > 0$$
$$H_3(m) = \begin{cases} H_3(m | " ), & \ell(m) = 3 \text{ and } ref(\nLeftarrow \nrightarrow m) > 0 \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

For instance-level node $m \ell(m) = 4$, $H_4(m | Rol)$ denotes the entropy generated by the participation of node $m$ in the relation as an instance bearer of $Rol$; $H(m)$ denotes the average of the entropy generated by $m$ under different $Rol$. For the model layer node $m \ell(m) = 3$, $H(m)$ denotes the entropy generated by the participation of node $m$ in different relations as an instance bearer of ", i.e., $H_4(m | ")$. For the metamodel layer and the metamodel layer, customization of their contents is not currently supported, so the entropies of the nodes in them are all 0.

For the entropy $H_4(m | Rol)$ and $H_4(m | ")$ generated by the nodes in the instance and model layers in a specific role, a generalized entropy for discrete random variables and considering the similarity between different values of the random variables is used to calculate and normalize the values to the interval $[0,1]$. For any discrete random variable $X$, its conventional information entropy $H(X)$ is defined as $-\sum_{x \in X} p_X(x) \cdot \log(p_X(x))$. Given a similarity calculation function $S: X \times X \rightarrow [0,1]$
between all values taken about \( X \), then another form of entropy
\[
H(X, S) = -\sum_{x \in X} p_X(x) \cdot \log \left( \sum_{y \in X} p_X(x) \cdot S(x, y) \right)
\]
can be defined called generalized entropy. The
generalized entropy \( H(X, S) \) has the following three basic properties [31-33]. (1) Degradability. The
generalized entropy degenerates to the conventional entropy, \( H(X, S) = H(X) \), if the similarity
between all two different values \( x, y \) of the random variable \( X \) is \( S(x, y) = 0 \). (2) Mergeability. If
the similarity \( S(x, y) = 1 \) between two different values \( x, y \) of the random variable \( X \), the two values
can be combined into one value (the probability of the new value is the sum of the probabilities of \( x, y \)), and the value of the generalized entropy remains unchanged. (3) The greater the similarity, the
smaller the generalized entropy. Given a discrete random variable \( X \) and two similarity functions
\( S_0, S_1 \), for any two values \( x, y \) of \( X \) if \( S_0(x, y) \leq S_1(x, y) \) is satisfied, then there must be
\( H(X, S_0) \geq H(X, S_1) \). To facilitate understanding, different values of the random variable can be
interpreted as different points of view, and the similarity between different values means the
similarity between different points of view, then the generalized entropy of the random variable portrays the inconsistency in a particular group due to different individuals having different
viewpoints and the inconsistency arising from different individuals in a given group.

\[
p_4(X, m, \text{Rol}) = \sum_{x \in X} \left[ i \in [1, N], \left( \text{Rol}_i \triangleright \rho \triangleright m \right) \in G \right]
\]

\[
S_4(X, Y \mid m, \text{Rol}) = \frac{\sum_{(x, y) \in \text{MATCH}(X, Y) \mid (x, y) \neq M} s_4(x, y \mid m, \text{Rol})}{|X| + |Y| - \max M(X, Y \mid m, \text{Rol})}
\]

\[
\text{MATCH}(X, Y \mid m, \text{Rol}) = \left\{ M \in X \times Y, \forall (x, y) \in M \implies (x, y) \neq (u, v) \implies x \neq u \land y \neq v \right\}
\]

\[
s_4(x, y \mid m, \text{Rol}) = \sum_{\text{Rol}(x, y \mid m, \text{Rol})} \frac{\text{sim}(v^\text{Rol}_{x}, v^\text{Rol}_{y})}{|\text{Rol}(x, y \mid m, \text{Rol})|}
\]

\[
\text{Rol}(x, y \mid m, \text{Rol}) = \left\{ \text{Rol} \mid \begin{array}{l}
\text{Rol} \in V(\triangleright), \text{Rol} \neq \text{Rol}', \\
\text{Rol}'(x \leq \text{rol}'_{x}) \in G, \text{rol}'_{x} \in V(\text{Rol}'), \\
(y \leq \text{rol}'_{y}) \in G, \text{rol}'_{y} \in V(\text{Rol}') \end{array} \right\}
\]

\[
v^\text{Rol}_{x} \equiv \left\{ \psi \left( x \leq \text{rol}'(\rho') \triangleright v \right) \in G, \text{rol}' \in V(\text{Rol}') \right\}
\]
For the instance layer node $m$, the probability distribution function $p_3(\chi^l | m, Rol)$ of the discrete random variable corresponding to its entropy $H_4(m | Rol)$ generated under $Rol$ is defined in Equation (16), and its physical meaning is the proportion of occurrences of the viewpoint $X$ in all individual knowledge graphs containing the structure “$m$ nodes connected to a particular relationship through $Rol$ instances”; the denominator represents the number of individual knowledge graphs containing the structure “$m$ nodes connected to $Rol$ instances”; the numerator represents the number of occurrences of view $X$ in these individual knowledge graphs. The similarity between two different viewpoints $X$, $Y$ is calculated in Equation (9). Where the numerator denotes the sum of the similarity of all matched relational pairs $(x, y)$ in the maximal similarity pie between the elements in the two views; the denominator denotes the sum of the elements contained in the two views minus the number of elements in the maximal similarity pie. This formula is a generalized form of the Jaccard coefficient between two sets. The similarity of a pair of relations $x$, $y$ is calculated in the form of equation (12), whose physical meaning is the average of the similarity of the bearers $v_x^{Rol}$, $v_y^{Rol}$ of each of the other roles in these two relations excluding the role $Rol$. When the elements contained in both sets of bearers are instances of symbolic concept instances, there exists a corresponding literal for each element and the words contained in each literal have a corresponding embedded vector representation. At this point, we calculate the similarity between the literal quantities corresponding to the elements of these two sets and the similarity between the corresponding embedded vectors, and take the greater of these two similarities as the similarity between the two bearer sets. Otherwise, if the elements contained in two bearer sets are not instances of symbolic concept instances, the Jaccard coefficient is used as the similarity between the two bearer sets.

$$\text{sim}(v, v) \doteq \begin{cases} \max \left( \text{sim}(\alpha(v), \alpha(v)) \right), & \exists u \in V(\#): u \neq 1 \ (v \cup v) \\ \text{Jaccard}(v, v), & \text{otherwise} \end{cases}$$  

(15)

For a modal instance layer node $m$, the probability distribution function $p_3(\chi^l | m, "e")$ of the discrete random variable corresponding to its entropy $H_5(m | "e")$ generated under $"e$ is defined in Equation (16), and its physical meaning is the proportion of occurrences of the viewpoint $X$ in all individual knowledge graphs containing $m$ nodes connected to the structure of a particular relationship through "$e$" instances. The denominator represents the number of individual knowledge graphs containing the structure of $m$ nodes connected to "$e$" instances; the numerator represents the number of occurrences of view $X$ in these individual knowledge graphs. Equation (17) gives the calculation of the similarity between two viewpoints. Since the nodes in the model layer do not have literal quantities, the Jaccard coefficient of the two sets is used directly as the similarity between the two views.
2.2 Fusion algorithm for individual knowledge graphs

Given a set \( G = \{G = (V_i, T_i, E_i) \mid i \in [1, N]\} \) of individual knowledge graphs, its fusion problem is a combinatorial optimization problem. The optimization goal is to find a fusion solution with the lowest possible entropy value in the fusion solution space \( M(G) \), and then form the corresponding group knowledge graph. In addition, it should be noted that in the process of collaborative group knowledge graph modeling, the individual knowledge graph of each participant may evolve continuously; accordingly, the individual knowledge graph fusion problem itself will evolve continuously.

Figure 1 gives the basic process of fusion of individual knowledge graphs, which consists of four main activities: initial fusion activity, which forms an initial group knowledge graph based on the results of fusion of individual knowledge graphs in the previous moment. Entropy calculation activity, which calculates the entropy of the current knowledge graph according to the formula; termination condition determination activity, which determines whether the fusion needs to continue to the next round according to the number of iterative rounds and the change of entropy value. Incremental fusion activity, based on the entropy value of each fusion node in the current population knowledge graph, screens out fusion nodes with possible fusion errors and re-fuses all individuals contained in them at the knowledge graph nodes. Where two activities, initial fusion and incremental fusion, make corresponding fusion decisions based on the participants’ responses to feedback information during the construction of individual knowledge graphs. The entropy calculation activity calculates the similarity between different words based on the vector embedding representation of the words and the tagging distance of the words. Incremental fusion activities will also improve the efficiency of re-fusion by narrowing the search for node alignment relationships using the associated word list built through word vectors.

![Figure 1. Automatic fusion of individual knowledge graphs](image-url)
2.3 Proactive feedback

In the active feedback activity, the supporting environment feeds specific knowledge map fragments to each participant based on the current group knowledge map to improve its efficiency in constructing the knowledge map. Each participant autonomously decides whether to accept, reject, or ignore the feedback information provided by the supporting environment. The accepted feedback information is entered into the participant’s individual knowledge map. Participants’ responses to the feedback are also recorded and used to assess individual knowledge preferences and the acceptance of specific information by the group.

In this paper, two types of information feedback are designed: context-independent feedback, and context-relevant feedback. In context-independent feedback, the support environment feeds a certain number of high-quality knowledge graph nodes directly to the participants without considering their current individual knowledge graphs and current operations. To achieve context-independent feedback, the support environment determines the feedback intensity of each fused node in the group knowledge graph by the following formula (the higher the feedback intensity of a node, the higher its probability of being fed back).

\[
ri(m) = \frac{1}{1 + H(m)}^{\text{ref}(m)}
\]

Where \(H(m)\) denotes the entropy value of node \(m\) and \(1/1 + H(m) \in (0,1]\) denotes the degree of consistency of nodes. \(\text{ref}(m) = \{m_i | m_i \in m, m_i \neq \perp\}\) denotes the number of occurrences of the node in the individual knowledge graph. It is known that the feedback strength of nodes is proportional to the consistency and the number of occurrences of the nodes themselves.

3 Experiments and results analysis

The performance of user task execution under the adaptive scheduling scheme is analyzed and modeled above, and the next step is to verify the analysis results through a series of simulation experiments and to investigate the effects of different parameters on the utility function. The parameters in the simulations are shown in Table 1.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of local nodes</td>
<td>2</td>
<td>Number of cloud nodes</td>
<td>3</td>
</tr>
<tr>
<td>Number of tasks that can be served simultaneously by the node</td>
<td>Local side:[4,3] Cloud:[5,5,4]</td>
<td>Arithmetic power of nodes</td>
<td>Local side:[0.4,0.3] Cloud:[2,2,1.5]</td>
</tr>
<tr>
<td>Task data volume upper and lower limits/Mbit</td>
<td>[0.5,3]</td>
<td>Upper and lower limits of computing power required per unit of data volume</td>
<td>[0.1,3.2]</td>
</tr>
<tr>
<td>Task arrival rate / task per time slot</td>
<td>1</td>
<td>Task tolerable latency limit/time slot</td>
<td>10</td>
</tr>
<tr>
<td>Queue length limit</td>
<td>20</td>
<td>Local to cloud data rate (Mbit/ms)</td>
<td>0.02 0.03 0.01</td>
</tr>
<tr>
<td>Reliability factor</td>
<td>0.8</td>
<td>Time delay factor</td>
<td>0.2</td>
</tr>
</tbody>
</table>
This section verifies the effectiveness and feasibility of the above knowledge graph construction method based on group intelligence through three types of experiments. Among them, (1) simulation graph fusion experiments verify the ability of the knowledge graph fusion algorithm proposed in this paper to effectively utilize the graph structure information by fusing the band type directed graphs generated by simulation that contain only structure information. (2) Knowledge graph fusion experiments to verify the effectiveness of the knowledge graph fusion algorithm proposed in this paper through automatic fusion of real-existing large-scale knowledge graphs. (3) Knowledge graph construction experiment verifies the feasibility of collaborative knowledge graph construction based on the “exploration-fusion-feedback” loop through online knowledge graph construction activities of multiple people under controlled conditions.

3.1 Simulation of graph fusion experiments

The knowledge graph fusion algorithm proposed in this paper makes comprehensive use of the semantic information of the nodes in the graph itself as well as the structural information of the graph. To examine the ability of the algorithm to effectively utilize the structural information of the graph, a set of directed graph pairs with different disparities of band types (in which, the nodes in the graph do not have any semantic information) are randomly generated in this paper, and the variation of the accuracy of the fusion scheme formed by the algorithm when fusing on the pairs with different disparities is analyzed.

For the experimental data generation, in this paper, the number of types of nodes in the graph is set to 6 and the type of edges is set to 3. The three values of 1000, 5000, and 10000 are chosen for the number of nodes, and the four values of 5, 10, 15, and 20 are chosen for the average degree of nodes. Then, six scale values of 5%, 10%, 15%, 20%, 25%, and 30% were used to randomly remove the corresponding percentage of nodes or edges from a graph.

In the experiments, the maximum number of iteration rounds maxRound of the algorithm is set to 5, the entropy reduction sensitivity $\epsilon$ is set to 0, and the re-fusion ratio per is set to 10% in this paper. The type of node and the type count of node-associated edges and points are used as the local structure vector of the node, and the initial fusion scheme is generated based on the node local structure vector, and the final fusion scheme is obtained by using the principle of minimum entropy for re-fusion. As shown in Figure 2, when the average degree of nodes is high and the variation rate is low, the accuracy rate is still close to 100%; as the size of nodes increases, the accuracy rate decreases slightly; as the average degree of nodes decreases and the variation rate increases, the accuracy rate decreases significantly.

![Figure 2. Accuracy of simulation graph fusion](image-url)
3.2 Knowledge graph fusion experiment

In this paper, two sets of real knowledge graph data are selected for knowledge graph fusion experiments. The data are derived from two pairs of knowledge graphs formed by sampling three real large-scale knowledge graphs, DBpedia, Wikidata and YAGO3: DBP-WD and DBP-YG. In which, each knowledge map contains 100,000 pairs of entities and millions of attributes and relationships. On these two pairs of knowledge maps, we compare and analyze the knowledge map fusion algorithm proposed in this paper with the traditional knowledge map fusion algorithm (LogMap) and the embedding-based knowledge map fusion algorithm (MTransE, BootEA, GCN-Align, TransD, JAPE, and MultiKE).

The embedding-based fusion algorithm generally uses one or more classes of information in the names, attributes and relationships of entities for training and constructing entity vectors; then a set of most similar entity nodes is obtained by the similarity between entity vectors as the candidate nodes for alignment relationships. These algorithms usually use the Hit@N metric (i.e., the probability that the first N candidate nodes contain correctly aligned nodes) as a criterion for alignment quality. In contrast, traditional algorithms and the algorithms in this paper directly generate one alignment result between two graphs (i.e., each node is given the unique one with which it is aligned). To ensure fairness, we use the Hit@1 metric to compare the effectiveness of these three types of algorithms.

The maximum number of iteration rounds maxRound of the algorithm is set to 5, the entropy reduction sensitivity ε is set to 0, and the re-fusion ratio per is set to 5% and 10%, respectively, in the experiments. The initial fusion scheme is generated using the similarity of entity names in the knowledge graph. And the final fusion scheme is obtained by using the principle of minimum entropy for re-fusion. Table 2 gives the fusion results of the 3 types of fusion algorithms on the two pairs of knowledge maps DBP-WD and DBP-YG.

Table 2. Comparison of different knowledge graph fusion methods

<table>
<thead>
<tr>
<th></th>
<th>Traditional Method</th>
<th>Based on Embedding method</th>
<th>Methodology of this paper</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LogMap</td>
<td>MTransE, BootEA, GCN-Align</td>
<td>EIR&amp;5%</td>
</tr>
<tr>
<td>DBP-WD@Hit1</td>
<td>80.68</td>
<td>74.79, 47.70</td>
<td>93.20, 93.69</td>
</tr>
<tr>
<td>DBP-YG@Hit1</td>
<td>83.73</td>
<td>76.10, 60.05</td>
<td>99.33, 99.33</td>
</tr>
</tbody>
</table>

The algorithm in this paper needs to reconverge a certain percentage of nodes in the iterative process; we choose two reconvergence ratios (5% and 10%) to run the algorithm in this paper in practice respectively. Experimental results show that these two re-fusion ratios do not have a significant effect on the Hit@1 index. However, compared with other algorithms, the algorithm in this paper achieved the highest Hit@1 metric on both experimental data. Compared with the previous algorithm with the best effect, on DBP-WD data, the Hit@1 index of this algorithm has increased by 1.75%–2.24%; on DBP-YG data, the Hit@1 index of this algorithm has increased by 11.3%.

Table 3 gives the change process of Hit@1 metrics and fusion graph entropy values on these two experimental data when the algorithm of this paper uses 5% and 10% re-fusion rate conditions, respectively. It can be seen that in the iterative fusion process, the Hit@1 metric and the fusion graph entropy value are converged rapidly on both data sets. Among them, on DBP-WD data, the algorithm in this paper enters the convergence state after three iterations (both the fusion graph entropy value and Hit@1 metric no longer change). And during the iteration, the entropy value continues to decrease and the Hit@1 index continues to improve. On DBP-YG data, the algorithm in this paper
enters the convergence state after one iteration. The experimental results show that with the minimization of entropy value as the optimization goal, a fusion result with high quality can be formed.

Table 3. Hit@1 Change in entropy value of indicator and fusion graph in iteration

<table>
<thead>
<tr>
<th></th>
<th>DBP-WD@5%</th>
<th></th>
<th>DBP-WD@10%</th>
<th></th>
<th>DBP-YG@5%</th>
<th></th>
<th>DBP-YG@10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Round0</td>
<td>Hit@1 (%)</td>
<td>Entropy</td>
<td>Hit@1 (%)</td>
<td>Entropy</td>
<td>Hit@1 (%)</td>
<td>Entropy</td>
<td>Hit@1 (%)</td>
</tr>
<tr>
<td>Round1</td>
<td>92.92</td>
<td>8991.86</td>
<td>92.92</td>
<td>8881.72</td>
<td>99.33</td>
<td>518.15</td>
<td>99.33</td>
</tr>
<tr>
<td>Round2</td>
<td>93.20</td>
<td>8939.41</td>
<td>93.69</td>
<td>8876.72</td>
<td>99.33</td>
<td>518.15</td>
<td>99.33</td>
</tr>
<tr>
<td>Round3</td>
<td>93.20</td>
<td>8940.76</td>
<td>93.68</td>
<td>8878.53</td>
<td>99.33</td>
<td>518.15</td>
<td>99.33</td>
</tr>
<tr>
<td>Round4</td>
<td>93.19</td>
<td>8940.24</td>
<td>93.67</td>
<td>8878.07</td>
<td>99.33</td>
<td>518.15</td>
<td>99.33</td>
</tr>
<tr>
<td>Round5</td>
<td>93.20</td>
<td>8940.24</td>
<td>93.68</td>
<td>8878.13</td>
<td>99.33</td>
<td>518.15</td>
<td>99.33</td>
</tr>
</tbody>
</table>

3.3 Impact of task arrival rate

Figure 3 shows the performance variation of the three task scheduling schemes with the service arrival rate. As mentioned before, task arrivals for each user obey an arbitrary distribution with rate $\lambda$ (assumed to obey Poisson distribution in the simulation), so the larger the value of $\lambda$, the larger the interval between two adjacent task arrivals, which means that the probability of a task being eliminated due to no free position in the queue decreases gradually, i.e., the probability of being successfully computed increases. In view of this, an obvious conclusion of Figure 2 is that the utility function of all schemes is an increasing function with respect to $\lambda$.

![Figure 3](image)

Figure 3. Utility function versus task arrival rate for different task scheduling schemes

As can be seen from this figure, when the value of $\lambda$ is very small, the value of the utility function tends to zero for all schemes, which is because almost all tasks are discarded at this time. As $\lambda$ increases, the performance of the 3 schemes appears to differ. Considering the features such as hierarchical intelligence and scalability of visual data, the proposed adaptive scheduling scheme outperforms the other two. While the cloud task scheduling algorithm and the non-adaptive task scheduling scheme still perform very close to each other when $\lambda$ is small, the former performs worse when $\lambda$ is larger, while the adaptive scheduling algorithm still performs best. As $\lambda$ increases, the performance of each scheme tends to be flat. This is because when $\lambda$ is large, tasks arrive more slowly and almost all requests are serviced without suffering from lack of resources. An increase in $\lambda$ does not continue to cause much change in task execution performance. The gain from the adaptive
scheduling algorithm can be seen in Figure 2, especially when the task arrives slowly and the network load is not high.

3.4 Effect of reliability factor

The impact of the reliability coefficient on the performance of the compared solutions is shown in Figure 4. The reliability coefficient characterizes how much importance the network places on reliability, i.e., the larger the value, the more the system cares about the reliability of task execution; and conversely, if the value is smaller, it indicates a higher requirement for delay performance. From Figure 3, the following conclusions can be drawn: first, it can be seen that the performance of all scenarios increases with the increase of the reliability coefficient. The increase of reliability coefficient can intuitively increase the value of the utility function. But the improvement of reliability performance inevitably causes deterioration of delay performance. Therefore, the overall utility function is not a linear function with respect to the reliability coefficient. Moreover, as the network pays more attention to reliability performance, each scheduling prefers to offload all the data to improve reliability, so the gap between adaptive and non-adaptive scheduling schemes shrinks as the reliability coefficient increases. Compared with the non-adaptive scheduling scheme and the cloud computing task scheduling scheme, the former performs less well than the latter when the reliability factor value is small, while its performance gradually outperforms the latter as the reliability factor increases. This indicates that when the reliability coefficient is small, the gain from considering the content data characteristics is higher than the selection of computation nodes, while with the increase of the reliability coefficient, the selection of the best computation node can bring higher gain. Finally, comparing the performance of the three schemes, it can be found that the adaptive scheduling scheme outperforms the other two significantly, especially when the reliability coefficient is small and the network cares more about the delay performance this gain reaches the highest.

![Figure 4. Relationship between utility function and reliability parameters for different task scheduling schemes](image_url)

4 Conclusion

This paper proposes a knowledge mapping method based on Internet+ group intelligence, which aims at constructing high-quality knowledge maps through the collaboration of human groups based on Internet+, in the context of “Internet+”, on how to do a good job of Civic Education in higher education institutions in the network culture environment. The goal is to construct a high-quality knowledge map through the collaboration of human groups based on “Internet+”. At the heart of the method is a loop containing 3 activities: exploration, integration, and feedback. The loop fuses the individual knowledge graphs constructed by participants in real time to form a group knowledge
graph, and feeds the knowledge fragments in the group knowledge graph to different participants in a targeted manner, providing an indirect interaction mechanism based on the environment (group knowledge graph) for effective collaboration of the group. In this paper, a multi-person online knowledge graph construction environment supporting the method is developed, and three different types of experiments are designed and implemented to provide a preliminary verification of the feasibility of the proposed method and key techniques. Based on the knowledge graph construction method proposed in this paper, a structured knowledge modeling ecosystem based on group intelligence is enabled. At the micro level, the ecosystem allows any autonomous human individual or intelligent algorithm to build and maintain an individual knowledge graph for a specific problem domain as it sees fit, and to collaborate effectively with other participants supported by an “exploration-fusion-feedback” loop. At the macro level, the ecosystem will emerge to form a large-scale, diverse, structured, and continuously evolving complex information artifact.

Acknowledgements

1) Higher vocational colleges systematically develop and implement ideological and political education through curriculum relying on professional curriculum development tools—A case study of “Fundamentals and applications of Economics” (educational science planned project of Guangdong Province of China, 2021JKDY061).

2) Ideology and politics education though curriculum of higher vocational colleges design and implement based on online courses in post epidemic era—taking “Fundamentals and applications of Economics” for example (Network ideological and political work center of Guangdong’s university 2021 project, GDWL21YB55).

3) Research and practice on business and trade major of higher vocational colleges teaching contents merging with ideology and politics education though curriculum relying on modern tools such as occupational ability analysis table, Mind map, fishbone diagram—taking “Fundamentals and applications of Economics” for example (2021 major project of Guangdong higher vocational college teaching steering committee, SM202101).

4) Research and practice on courses teaching contents merging with ideology and politics education though curriculum relying on modern tools of course development and project management tools —A case study of “Fundamentals and applications of Economics (Shenzhen Polytechnic 2021 ideology and politics education planning project, 10018021310018Z).

References

[4] Fang