Online-Journals.org

iJET International Journal of Emerging Technologies in Learning

iJET | eISSN: 1863-0383 | Vol. 18 No. 17 (2023) | OPEN ACCESS

https://doi.org/10.3991/ijet.v18i17.43487

PAPER

Construction and Application of a Complex Network-Based Case Knowledge Base in an Assisted Instruction System

Shanshan Wang(✉),
Xiangpeng Chang,
Dong Ding

Shijiazhuang University of Applied Technology, Shijiazhuang, China

2017010823@sjzpt.edu.cn

ABSTRACT

With the development of the business administration major, case teaching has become an effective teaching method. However, a lot of case resources exist in the business administration field, and the correlation between these cases is complex, which poses challenges to teachers and students in terms of information overload and knowledge organization. Existing case knowledge bases in assisted instruction systems often have shortcomings in terms of knowledge relationship strength calculation, text semantic similarity calculation, hierarchical knowledge clustering, and propagation evolution. To solve these problems, this study proposed a method of constructing and applying a case knowledge base in an assisted instruction system based on a complex network. This method mainly includes three aspects: calculation and reasoning of ontology relationship strength, calculation of ontology text semantic similarity, and hierarchical knowledge clustering and propagation evolution in complex networks. Through a comprehensive study of these three aspects, a more efficient and intelligent case knowledge base in an assisted instruction system was constructed, which not only improved the teaching efficiency and quality of the business administration major but also had the potential to promote teaching in other disciplines and fields. In addition, this study also provided new perspectives and methods in related fields, which is of great significance for efficiently organizing and utilizing knowledge.

KEYWORDS

complex network, assisted instruction system, teaching cases, knowledge base construction

1 INTRODUCTION

With the vigorous development of the business administration major, the teaching demand in related fields has been increasing day by day. Business administration covers multiple sub-disciplines, such as marketing, human resource management,

Wang, S., Chang, X., Ding, D. (2023). Construction and Application of a Complex Network-Based Case Knowledge Base in an Assisted Instruction System. *International Journal of Emerging Technologies in Learning (iJET)*, 18(17), pp. 153–167. https://doi.org/10.3991/ijet.v18i17.43487

Article submitted 2023-06-24. Revision uploaded 2023-07-26. Final acceptance 2023-07-27.

financial management, strategic management, and so on, with rich theoretical and practical cases in each sub-discipline. In the teaching process, case teaching helps students gain a deeper understanding of theoretical knowledge and improves their ability to solve practical problems [1–5]. However, there are lots of case resources in the business administration field, which often causes the problem of information overload for teachers and students when they search and use relevant cases. In addition, the correlation between different cases is often complex, which is of great significance for students' learning, but it is difficult to demonstrate it through traditional teaching methods. Therefore, building a case knowledge base in an assisted instruction system based on a complex network can greatly promote the teaching effect of a business administration major [6–10].

In the business administration field, building such a system not only provides a rich case resource library, but also further taps into potential knowledge and insights using complex network technology to analyze the correlation between cases [11–14]. This system not only enables students to learn and understand the core concepts and methods of business administration more systematically, but also enables teachers to prepare courses and guide students more efficiently [15–18]. Therefore, the construction and application of a case knowledge base in an assisted instruction system based on a complex network is of great significance for improving the teaching quality and efficiency of business administration.

However, the existing case knowledge bases in assisted instruction systems often lack effective knowledge relationship strength calculation and reasoning mechanism, which cannot accurately reflect the internal relationships between cases [19] [20]. Meanwhile, when dealing with large-scale and multi-domain case texts, traditional text semantic similarity calculation methods often fail to capture deep semantic information well [21] [22]. In addition, the existing knowledge clustering and propagation evolution methods have some limitations when dealing with hierarchical knowledge in complex networks [23] [24].

In view of the above problems, this study proposed a construction and application method of a case knowledge base in an assisted instruction system based on a complex network. First, the calculation and reasoning of ontology relationship strength were studied to reveal the deep connection between cases. Second, the ontology text semantic similarity was calculated to capture and analyze the deep semantic information of case texts. Furthermore, the hierarchical knowledge clustering and propagation evolution in complex networks were studied to reveal and optimize knowledge organization and propagation at different levels.

This study not only improves the teaching efficiency and quality of the business administration major but also has broad application value. The research method of this study provides references for teaching in other disciplines and fields and promotes the development of case knowledge bases in assisted instruction systems based on complex networks. The innovations in ontology relationship strength calculation, text semantic similarity calculation, hierarchical knowledge clustering, and propagation evolution may provide new perspectives and research methods in related fields. Efficient knowledge organization and utilization have become important challenges in the era of information explosion, and this study provides a powerful tool to address this challenge.

2 CALCULATION OF ONTOLOGY RELATIONSHIP STRENGTH

To reveal the internal connection between cases and help students better understand concepts and knowledge, this study first calculated the ontology relationship strength of the case knowledge base in an assisted instruction system. The calculation enhanced the efficiency of knowledge retrieval. When students or teachers searched specific topics in the knowledge base, the system quickly located relevant cases based on relationship strength, thereby reducing the search time.

2.1 Construction of PageRank diagram

To calculate the ontology relationship strength of the case knowledge base, this study extracted corresponding context, problem, and knowledge point ontologies from cases of the assisted instruction system and constructed a PageRank diagram. The diagram quantified the importance of each element (context, problem, and knowledge point) in the network. The link between elements represented their relationships, which helped reveal the internal connection between contexts, problems, and knowledge points, thereby better understanding the relationships between various cases and the knowledge system structure. The PageRank diagram enabled teachers to clearly see the importance of each knowledge point and problem and optimize the allocation of teaching resources accordingly, such as time, attention, and materials. Figure 1 shows an example of the PageRank diagram architecture.

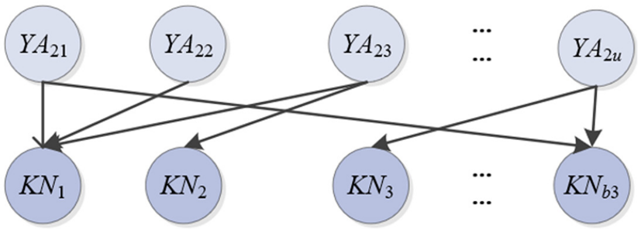

Fig. 1. Example of PageRank diagram architecture

The following three sets were defined in accordance with the corresponding context, problem, and knowledge point ontologies, which were extracted from cases of the assisted instruction system. Let $ZZJ_f = \{(YA_{fu}, \{ZZJ_1, ZZJ_2, ..., ZZJ_{b1}\})\}$ be the context set, ZZJ_f be the context set involving assisted instruction cases of course f, YA_{fu} be the u-th assisted instruction case among all cases of course f, and $\{ZZJ_1, ZZJ_2, ..., ZZJ_{b1}\}$ be the context set corresponding to the u-th assisted instruction case. Similarly, according to the above rules, let $ZHJ_f = \{(YA_{fu}, \{ZHJ_1, ZZJ_2, ..., ZHJ_{b2}\})\}$ be the problem set, $ZYJ_f = \{(YA_{fu}, \{ZYJ_1, ZYJ_2, ..., ZYJ_{b3}\})\}$ be the knowledge point set, ZHJ_f be the problem set involving assisted instruction cases of course f, $\{ZHJ_1, ZZJ_2, ..., ZHJ_{b2}\}$ be the problem set included in assisted instruction case m of course f, ZYJ_f be the knowledge point set involving assisted instruction cases of course f, and $\{ZYJ_1, ZYJ_2, ..., ZYJ_{b3}\}$ be the knowledge point set included in assisted instruction case u of course f.

The PageRank diagrams of contexts, problems, and knowledge points were constructed, respectively, with assisted instruction case nodes as the outlink nodes, and the nodes contained in the context set, problem set, and knowledge point set of existing course assisted instruction cases as the inlink nodes. The PR values of the codes of corresponding contexts, problems, and knowledge points in the three PageRank diagrams were calculated, respectively. Let *s* be the damping factor, *B* be the number of nodes in the diagram, *c* be a certain assisted instruction case containing *i*, B_u be the assisted instruction case set containing *i*, and *M*(*c*) be the number of other contexts, problems, or knowledge points included in the assisted instruction case *c*. The equation for calculating the PR value was given as follows:

$$PR(i) = \frac{1-a}{B} + s * \sum_{c \in N_i} \frac{PR(c)}{M(c)} \tag{1}$$

2.2 Construction of logic rules

Logic rules ensure that information and relationships in the ontology are logically consistent and accurate. By defining clear rules and constraints, contradictory or erroneous information in the knowledge base can be avoided, thereby improving the reliability of the knowledge base. At the same time, the construction of logic rules provides the basis for automated reasoning. Application of logical rules enables the system to automatically derive new knowledge and relationships, thereby enriching the contents of the knowledge base and helping users have a deeper understanding of relevant concepts and cases. Therefore, this study further constructed logic rules for the case knowledge base ontology in the assisted instruction system. The following two rules, e_1 and e_2 were made:

$$e_1 : NX(N,X) \wedge NG(N,H) \rightarrow XX(X,G) \tag{2}$$

$$e_2 : NG(N,G) \wedge NT(N,T) \rightarrow XT(G,T) \tag{3}$$

In e_1, *NX*(*N*, *X*) represents that case *N* has context *X*, and *NG*(*N*, G) represents that case *N* is identified as problem *G*, so context *X* is correlated with problem *G* in *N*. In e_2, *NG*(*N*, *G*) represents that case *N* is identified as problem *G*, and *NT*(*N*, *T*) represents that case *N* uses knowledge point *T*, so problem *G* is correlated with knowledge point *T* in *N*.

In teaching cases, the relationships between knowledge points may be uncertain or fuzzy. Probabilistic soft logic better handles this uncertainty by introducing probability distributions, thereby providing quantitative metrics for fuzzy or uncertain relationships. Probabilistic soft logic provided more fine-grained evaluations using continuous numerical values as soft truth values, which enabled the system to distinguish subtle differences and represent the relationship strength between knowledge points more accurately. Figure 2 shows a schematic diagram of probabilistic soft logical reasoning.

Let *E* be the rule set, η_e be the rule weight, *X* be the normalization factor, *f*(*e*) be the distance satisfaction degree of the rules, and *o* be the order of predicate logic. The equations used to solve the probability distribution of probabilistic soft logic were as follows:

$$O(U) = \frac{1}{X} \exp\left[-\sum_{e \in E} \eta_e (f(e))^o\right] \tag{4}$$

$$X = \int \exp\left[-\sum_{e \in E} \eta_e (f(e))^o\right] \tag{5}$$

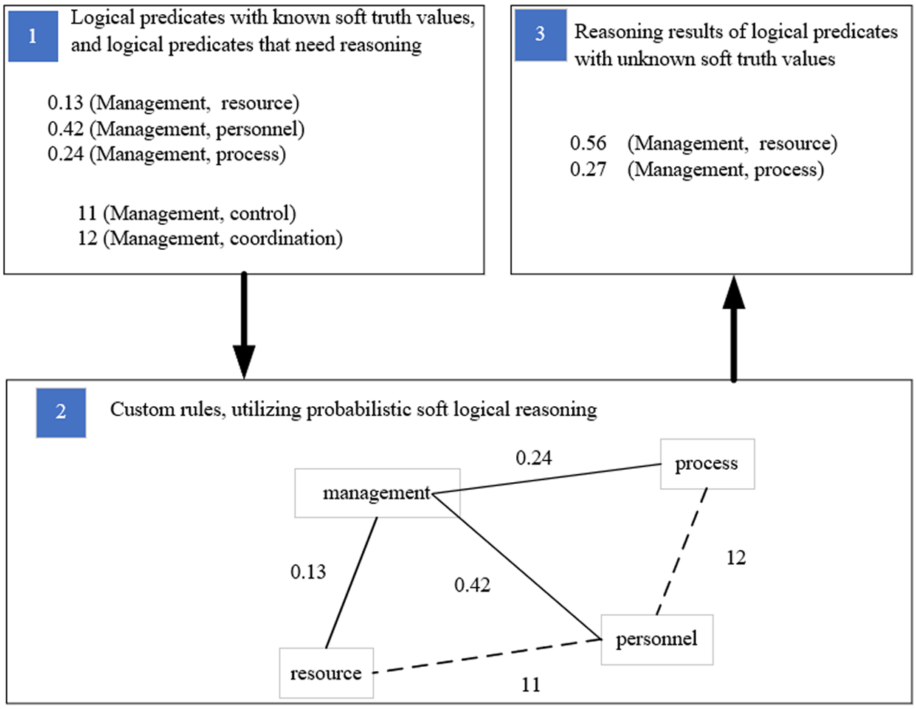

Fig. 2. Schematic diagram of probabilistic soft logical reasoning

3 CALCULATION OF ONTOLOGY TEXT SEMANTIC SIMILARITY

This study calculated the ontology text semantic similarity of the case knowledge base in an assisted instruction system, which aimed to more accurately match and recommend teaching resources and cases related to students' learning objectives, thereby enabling them to focus more on meaningful contents to improve learning efficiency. The semantic similarity calculation not only considered the literal similarity of words but also included deep semantic connections, which helped reveal the implicit correlations and patterns in the ontology, providing a richer perspective for understanding the internal structure of knowledge.

3.1 Calculation of concept semantic similarity

Taking the business administration major as an example, its core concepts, entities, attributes, and relationships should be considered when designing the domain ontology for the experimental subject of this study. The name of the domain ontology is the business administration ontology. The core concepts include: 1) Management: as a central concept, management includes the control and coordination of resources, personnel, and processes within the organization. 2) Organization, which involves enterprises, companies, departments, etc. and is the basic unit of management activities. 3) Strategy: a long-term plan and goal about how to achieve organizational goals and missions. 4) Market: the environment and mechanisms related to the transaction of goods and services. 5) Finance, which involves fund management, including budgeting, accounting, investment, etc. 6) Human resources, which involve recruitment, training, evaluation, and motivation of employees. 7) Operation, including production, supply chain management, logistics, etc. 8) Marketing, which is related to the promotion and sale of products and services.

This study calculated the concept semantic similarity of the business administration ontology through various methods. Various concepts mentioned in the ontology occurred at the same frequency, and some of them did not have attributes, which may lead to the inaccuracy of traditional similarity calculation methods based on frequency or attributes. This problem was overcome using a distance-based semantic similarity algorithm because this algorithm mainly relied on ontology structure rather than frequency or attributes. To more accurately reflect the semantic distance between concepts, the shortest path length between two concepts in the ontology tree organizational structure was considered, which represented the semantic distance between them more accurately. In a structured ontology, path length can be a good index of concept similarity. To provide deeper analysis and weight each relationship path edge, different levels of concepts in the ontology hierarchy tree can be given different importance, which helps highlight more important or core concepts and provides deeper analysis. In addition to considering the hyponymy relation between concepts, this study weighted the object attribute relationships between concepts, which captured richer relationship types in the ontology, thereby enhancing the precision and robustness of similarity calculation.

Let LE be the maximum depth of case ontology tree structure in an assisted instruction system, $DI(v_1,v_2)$ be the length of the shortest path between concepts v_1 and v_2, WE_u be the weight of the u-th edge on the shortest path connecting v_1 and v_2, and DE be the number of hierarchies where the superordinate concept connected by the edge is located. The equations of concept semantic similarity were as follows:

$$SI(v_u, v_k) = \frac{2*(LE-1) - DI(v_u, v_k)}{2*(LE-1)} \tag{6}$$

$$DI(v_u, v_k) = \sum_{u=1}^{b} WE_u \tag{7}$$

$$WE_u = \frac{1}{DE} \tag{8}$$

3.2 Text similarity calculation

The case ontology text similarity of the assisted instruction system, which was based on conceptual and non-conceptual feature words, was calculated according to the above three equations. Based on the calculation results, this study further proposed a similarity algorithm between texts suitable for case knowledge bases in assisted instruction systems. Let β be the regulatory factor, then the equation of text semantic similarity was as follows:

$$SI(f_u, f_k) = \beta * vo\_SI(f_u, f_k) + (1-\beta) * y\_SI(f_u, f_k) \tag{9}$$

4 HIERARCHICAL KNOWLEDGE CLUSTERING AND PROPAGATION EVOLUTION IN COMPLEX NETWORK

4.1 Hierarchical text clustering of ontology

This study organized relevant information according to different hierarchies and classes through hierarchical clustering of the knowledge base, which aimed to make the teaching system recommend relevant learning materials and cases to students more efficiently, thereby providing a personalized learning experience. This structured data organization method helped users find the information that they needed more quickly and intuitively. Hierarchical clustering revealed the potential correlations and patterns between different knowledge points and cases by analyzing the similarities and differences of text contents, which is valuable for discovering new knowledge or understanding the structure of complex themes. Figure 3 shows the basic pattern of the hierarchical text clustering algorithm used in this study.

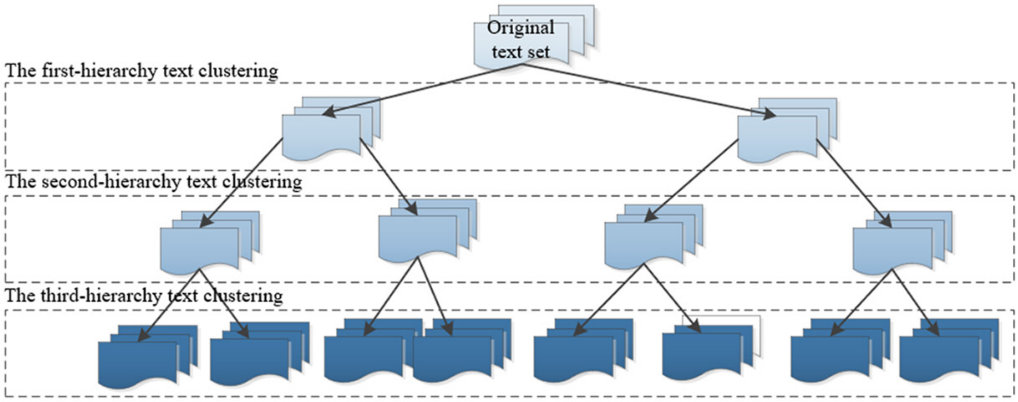

Fig. 3. Basic pattern of hierarchical text clustering algorithm

Thematic knowledge bases usually have a tree structure, which includes hierarchical information and knowledge. The complex relationships between texts may not be fully captured, if the single-hierarchy clustering method is used. However, hierarchical clustering discovers and utilizes these relationships at different abstraction levels. In addition, when text clustering is presented in a tree structure, users more clearly understand the relationships between different classes and their positions in the knowledge system. This study first revealed the potential patterns and correlations in text data through hierarchical clustering. An agglomerative hierarchical clustering algorithm was used in hierarchical text clustering. In the algorithm, let f_l be the texts in class CL_u, f_b be the texts in class CL_k, b_u and b_k be the total number of texts contained in classes CL_u and CL_k, respectively. The similarity between classes was then calculated using the following equation:

$$SI(CL_u, CL_k) = \frac{\sum_{l \in CL_u} \sum_{b \in CL_k} SI(f_l, f_b)}{b_u * b_k} \tag{10}$$

The algorithm's efficiency is crucial when dealing with large-scale ontologies. The fast community partitioning algorithm can process a large amount of data in a short period of time, reducing computational costs. At the same time, this algorithm maintains high clustering quality. In hierarchical clustering, this enables the clustering results obtained at each hierarchy to maintain consistency and accuracy. The hierarchical clustering steps of the ontology based on the fast community partitioning algorithm are given below:

Step 1: In the case of a knowledge base complex network in an assisted instruction system with b nodes and l edges, the network was initialized and divided into b knowledge communities. Let j_u be the degree of node u, r_{uk} be the total number of edges between communities u and k divided by that of the knowledge base complex networks, and s_u be the number of edges between community u and the outside world divided by that of the knowledge base complex networks. The initialization matrix $R = (r_{uk})$ was calculated to make r_{uk} and s_u meet the following conditions, respectively:

$$r_{uk} = \begin{cases} \frac{1}{2l}; \text{if there is any edge between nodes } u \text{ and } k \\ 0; \text{else} \end{cases} \tag{11}$$

$$s_u = \frac{j_u}{2l} \tag{12}$$

Step 2: The knowledge community with the maximum number of edges was merged, and the module variation $\Delta W = r_{uk} + r_{ku} - 2s_{uk}$ was calculated. After the merge was completed, the matrix R was updated by adding the u-th row of R to the k-th row and deleting the u-th row.

Step 3: The knowledge communities were merged continuously until all nodes of the knowledge base complex network formed a knowledge community. Then the community partitioning results corresponding to the iteration with the highest module variation were selected as the final partitioning results.

4.2 Analysis of knowledge dissemination evolution trend based on network structure

Knowledge constantly develops and evolves. By analyzing the dissemination evolution trend, the development and changes of teaching knowledge can be dynamically tracked, thereby better understanding the current situation and future direction of a certain field, which is very valuable information for educators and learners. By analyzing the dissemination evolution trend of knowledge, the assisted instruction system can provide learners with more personalized learning recommendations. The system recommends relevant learning resources and paths to learners based on their interests, and needs, and development trends in related fields.

Disruptive knowledge points often indicate significant changes in the knowledge field, while stable ones may represent long-term, stable development. By identifying these knowledge points, it is easier to identify key turning points in knowledge evolution. Provision of a dynamic perspective: in the knowledge dissemination network structure, disruptive knowledge points may lead to rapid changes in the network structure, while stable ones may be associated with more stable structures. By introducing this index, a dynamic perspective can be provided for analysis to reveal how knowledge networks change over time.

The case knowledge base complex network in the assisted instruction system was regarded as a directed three-part diagram $H = (C_1, C_2, C_3, R)$. The whole network was composed of three functional nodes of (C_1, C_2, and C_3) and an edge R. Let b_y be the total number of future related knowledge points within time constraints, and q_{uy} be the weight of each knowledge point. An index ZB_y was introduced to distinguish whether the knowledge point was disruptive or stable.

$$ZB_y = \frac{1}{b} \sum_{u=1}^{b} \frac{(-2d_{uy} n_{uy} + d_{uy})}{q_{uy}}, q_{uy} > 0 \tag{13}$$

where,

$$d_{uy} = \begin{cases} 1; \text{if } i \text{ is correlated with other central ontologies} \\ 0; \text{else} \end{cases} \tag{14}$$

and,

$$n_{uy} = \begin{cases} 1; \text{if } i \text{ is correlated with ontologies related to any other central ontology} \\ 0; \text{else} \end{cases} \tag{15}$$

The value range of ZB_y is [–1,1]. When ZB_y is greater than 0, it indicates that the ontology is more inclined towards disruptive knowledge. When it is less than 0, it indicates that the ontology is more inclined towards consolidated knowledge. The following weighting method was adopted to better identify the inclination degree of two types of knowledge:

$$lZB_y = \frac{l_y}{b_y} \sum_{u=1}^{b} \frac{(-2d_{uy} n_{uy} + d_{uy})}{q_{uy}}, q_{uy} > 0 \tag{16}$$

5 EXPERIMENTAL RESULTS AND ANALYSIS

Table 1. Experimental results of different methods for solving logical probability distribution

Methods		*P*	*R*	*F*1
Context-knowledge point	Maximum likelihood estimation	0.75	0.71	0.72
	Probabilistic soft logic	0.87	0.84	0.83
Problem-knowledge point	Maximum likelihood estimation	0.63	0.63	0.61
	Probabilistic soft logic	0.76	0.69	0.77

It can be observed from Table 1 that the probabilistic soft logic method is superior to the maximum likelihood estimation method in precision (P), recall (R), and F1-score (F) in the two tasks of "context-knowledge point" and "problem-knowledge point," indicating that probabilistic soft logic more effectively identifies and reasons relevant contexts, problems, and knowledge points when calculating the ontology relationship strength of the case knowledge base in an assisted instruction system. Specifically, in the context-knowledge point task, the probabilistic soft logic has 0.87 precision, 0.84 recall, and 0.83 F1-score. While the maximum likelihood estimation has 0.75 precision, 0.71 recall, and 0.72 F1-score, indicating that probabilistic soft logic has higher accuracy and coverage in context-knowledge point matching and reasoning. In the problem-knowledge point task, the probabilistic soft logic has 0.76 precision, 0.69 recall, and 0.77 F1-score. While the maximum likelihood estimation has 0.63 precision, 0.63 recall, and 0.61 F1-score, also indicating that probabilistic soft logic has higher accuracy and coverage in problem-knowledge point matching and reasoning.

Table 2 shows the text clustering results of different clustering methods and their entropy evaluation. Classes 1–8 in the table correspond to the following classes: 1. Marketing and strategy; 2. Human resource management; 3. Financial management and accounting; 4. Operation and supply chain management; 5. Organizational behavior and leadership; 6. Entrepreneurship and innovation management; 7. International business; 8. Project management.

It can be clearly observed from Table 2 that the method proposed in this study is significantly superior to the expected maximum clustering method in terms of entropy, both in the first- and second-hierarchy clustering results. The entropy is inversely proportional to the clustering effect, meaning that the proposed method has a better clustering effect. Specifically, the entropy of the proposed method and the expected maximum clustering are 176.64 and 301.36, respectively, in the first-hierarchy clustering results. Meanwhile, it can be noted that the sample distribution of the proposed method in both classes is more balanced, meaning that the proposed method is more accurate in distinguishing between classes "marketing and strategy" and "human resource management". The entropy of the proposed method and the expected maximum clustering are 62.72 and 188.63, respectively, in the second-hierarchy clustering results (Class Cluster 2, with four classes in total). Similarly, the sample distribution of the proposed method in the four classes is more balanced, meaning that the proposed method has higher accuracy in subdividing classes. In the second-hierarchy clustering results (Class Cluster 2 with eight classes in total, with 0.8 as the regulatory factor), the entropy of the proposed method further decreases to 36.97, while that of the expected maximum clustering is 121.63. The sample distribution of the proposed method in the eight classes is still relatively balanced, indicating that the proposed method still maintains high clustering accuracy when considering more classes.

Table 2. Text clustering results of different clustering methods and their entropy evaluation

The First-Hierarchy Clustering Results (Class Cluster 2)			
Evaluation Indexes	Entropy	Number of Samples Contained in the Class Cluster	
		Class 1	Class 2
Method proposed in this study	176.64	117	93
Expected maximum clustering	301.36	186	11

The Second-Hierarchy Clustering Results (Class Cluster 2, with Four Classes in Total)					
Evaluation Indexes	Entropy	Number of Samples Contained in the Class Cluster			
		Class 1	Class 2	Class 3	Class 4
Method proposed in this study	62.72	57	57	52	40
Expected maximum clustering	188.63	159	28	8	2

The Second-Hierarchy Clustering Results (Class Cluster 2 with Eight Classes in Total, with 0.8 as the Regulatory Factor)									
Evaluation Indexes	Entropy	Number of Samples Contained in the Class Cluster							
		Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8
Method proposed in this study	36.97	31	29	27	22	26	23	22	17
Expected maximum clustering	121.63	112	57	20	8	5	3	1	1

It can be seen from Figure 4 that the method proposed in this study is significantly higher than K-means clustering and the traditional hierarchical clustering method in each class in terms of the F1-score of Classes 1 to 8. A higher F1-score indicates that the clustering method has better performance in correctly classifying relevant knowledge points into corresponding classes because F1-score is the harmonic average of precision and recall. Specifically, the F1-score of the proposed method far exceeds that of the other two methods in Class 2 (human resource management) and Class 5 (organizational behavior and leadership), indicating that the proposed method has particularly outstanding clustering quality in these fields. In terms of average value, the average F1-score of the proposed method is 0.7, while that of K-means and traditional hierarchical clustering is 0.51 and 0.53, respectively, which further proves that the proposed method has higher clustering quality on the whole. Based on the data in Figure 4, it can be concluded that the proposed method exhibits a high F1-score in clustering knowledge points of business administration both in individual classes and at the overall average level, indicating the proposed method has significant advantages in clustering accuracy and accurate classification of relevant knowledge points. Therefore, the method proposed in this study is a more desirable and effective choice for ontology clustering in the business administration field.

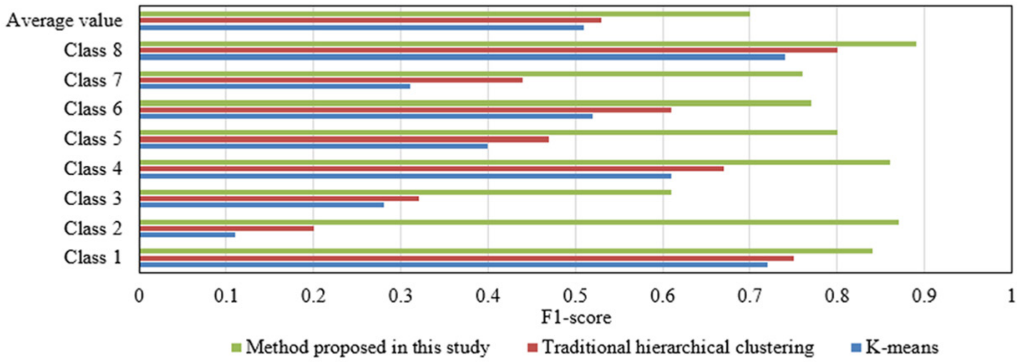

Fig. 4. F1-score of different clustering methods

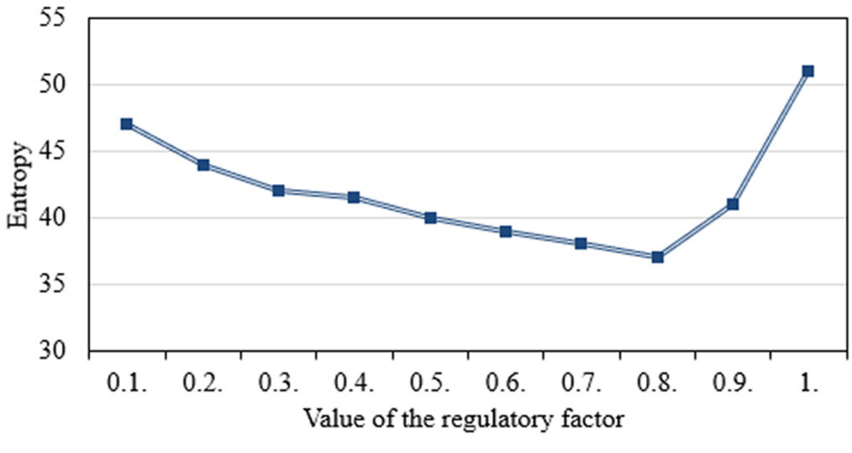

Fig. 5. Correspondence between regulatory factor and entropy

It can be seen from Figure 5 that the entropy shows an overall downward trend as the regulatory factor value gradually increases from 0.1 to 0.8. However, the entropy begins to rise as the regulatory factor value continues to increase to 0.9. Specifically, when the regulatory factor value ranges from 0.1 to 0.8, the entropy gradually decreases from 47 to 37, indicating that an increase in the regulatory factor value within this range improves the clustering effect and makes the clustering results more accurate. However, when the regulatory factor value increases to 0.9 and 1, the entropy increases to 41 and 51, respectively, indicating that excessively high regulatory factor values lead to over-regulation and poor clustering performance. As shown in the figure, the regulatory factor plays an important role in clustering analysis. The clustering accuracy can be improved by choosing an appropriate regulatory factor value. However, excessively high regulatory factor values have the opposite effect, leading to poor clustering performance. It can be observed from Figure 5 that the entropy reaches its lowest when the regulatory factor value is between 0.7 and 0.8, which is a relatively ideal value range, maintaining the clustering effect while avoiding the negative impact caused by over-regulation. Therefore, selecting an appropriate regulatory factor value is crucial when making clustering analyses.

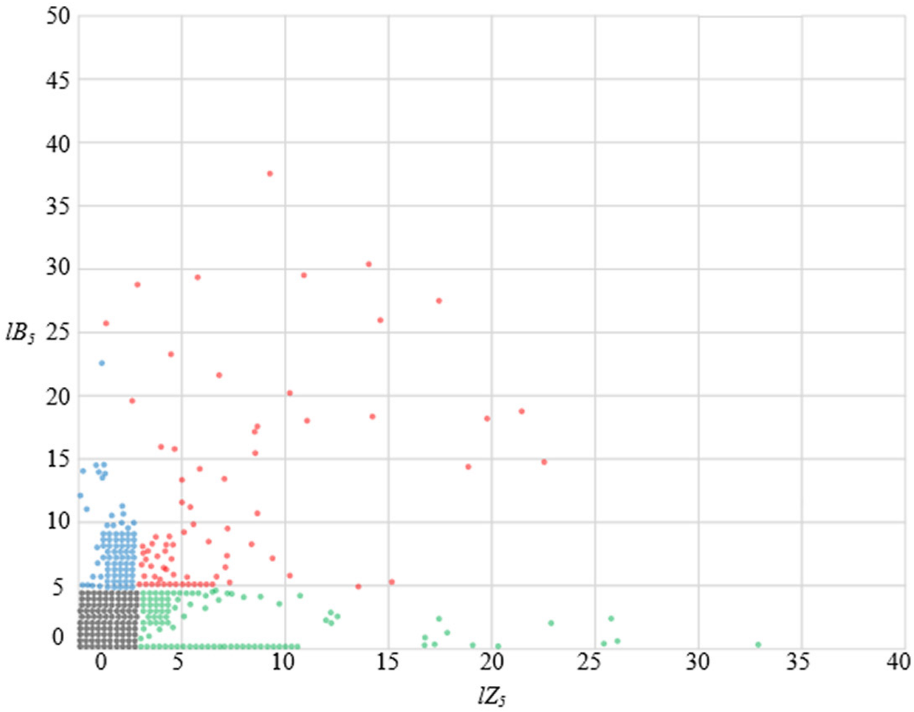

Fig. 6. Distribution of lZ_5 and lB_5 values of ontology

The index lZB_y, which is used for distinguishing whether the knowledge point is disruptive or stable, can be decomposed into two indexes: lZ_5 and lB_5. When lZ_5 or lB_5 of an ontology involving a teaching case is greater than 1, the case can be considered valuable and worth repeated use. This type of case only accounts for 1% of the total cases in the assisted instruction system. Figure 6 shows the distribution of lZ_5 or lB_5 values for all ontology patents involving the case. A knowledge point can be distinguished as disruptive or stable using the two indexes in Figure 6, and the value of the teaching case can be further judged. It can be seen from the distribution map that the index values of most ontologies are concentrated near the origin and spread outward in an exponential distribution, meaning that most knowledge points have low disruption and stability, while only a few of them are highly disruptive or stable.

6 CONCLUSION

This study mainly discussed the related issues of case knowledge base ontology in assisted instruction systems, such as hierarchical text clustering, semantic similarity calculation, and analysis of knowledge dissemination evolution trends. First, taking the business administration major as an example, this study constructed a domain ontology for organizing and managing relevant knowledge. The semantic similarity between ontology concepts was effectively calculated using the distance-based semantic similarity algorithm. This study improved the accuracy of semantic similarity results, by considering the hyponymy relation between concepts and object attribute relationships, and by weighting the relationship path edges. The hierarchical clustering method was used to classify texts in this study. Preliminary clustering was first performed based on text similarity calculations. Then more detailed clustering was performed by combining deeper text similarity results. This method better aggregated the texts with related topics and displayed them in a tree structure, which is conducive to the organization of a thematic knowledge base.

The experimental results showed that by analyzing the relationships between the regulatory factor and the entropy, it was concluded that the entropy decreased when the regulatory factor increased within a certain range, which improved the clustering effect. After analyzing the dissemination and evolution trends of ontology, this study introduced indexes to distinguish whether knowledge points were disruptive or stable, thus determining the value and reusability of teaching cases. The experimental results showed that the clustering method proposed in this study was significantly better than the traditional K-means and hierarchical clustering methods in F1-score, indicating that the proposed method had significant advantages in clustering accuracy and correct classification of related knowledge points.

In summary, by introducing a new algorithm and evaluation indexes, this study successfully solved the problems of the case knowledge base ontology of an assisted instruction system, such as semantic similarity calculation, hierarchical text clustering, and knowledge dissemination evolution trend analysis. The method proposed in this study showed significant advantages in clustering effect, knowledge representation, knowledge dissemination analysis, and other aspects.

7 REFERENCES

[1] S. Hua and Z. Ren, "'Online + Offline' Course teaching based on case teaching method: A case study of entrepreneurship education course," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 15, no. 10, pp. 69–85, 2020. https://doi.org/10.3991/ijet.v15i10.13999

[2] N. Emami, "Topology optimization: A case study of precast panels," *Technology | Architecture + Design*, vol. 7, no. 1, pp. 67–82, 2023. https://doi.org/10.1080/24751448.2023.2176143

[3] K.H. Huarng, C.F. Lee, and T.H.K. Yu, "Case study of a healthcare virtual community model," *Technological Forecasting and Social Change*, vol. 188, p. 122281, 2023. https://doi.org/10.1016/j.techfore.2022.122281

[4] B. Standl, T. Kühn, and N. Chlomske-Bodenstein, "Student-collaboration in online computer science courses – An explorative case study," *International Journal of Engineering Pedagogy (iJEP)*, vol. 11, no. 5, pp. 87–104, 2021. https://doi.org/10.3991/ijep.v11i5.22413

[5] L.S. Dhamande, V.P. Bhaurkar, and P.N. Patil, "Vibration analysis of induced draught fan: A case study," *Materialstoday: Proceedings*, vol. 72, no. 3, pp. 657–663, 2023. https://doi.org/10.1016/j.matpr.2022.08.329

[6] T.A. Eppes, I. Milanovic, and K. Wright, "Improving student readiness for inquiry-based learning: An engineering case study," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 16, no. 1, pp. 4–17, 2020. https://doi.org/10.3991/ijoe.v16i01.12051

[7] I. Zulaeha, Subyantoro, C. Hasanudin, and R. Pristiwati, "Developing teaching materials of academic writing using mobile learning," *Ingénierie des Systèmes d'Information*, vol. 28, no. 2, pp. 409–418, 2023. https://doi.org/10.18280/isi.280216

[8] A. Smiti and Z. Elouedi, "Dynamic maintenance case base using knowledge discovery techniques for case based reasoning systems," *Theoretical Computer Science*, vol. 817, pp. 24–32, 2020. https://doi.org/10.1016/j.tcs.2019.06.026

[9] Z. Li, G. Huang, X. Chen, Y. Zhou, and Q. Yang, "Wind-resistant design and equivalent static wind load of base-isolated tall building: A case study," *Engineering Structures*, vol. 212, p. 110533, 2020. https://doi.org/10.1016/j.engstruct.2020.110533

[10] L.A. Pineda, N. Hernández, I. Torres, G. Fuentes, and N.P. De Avila, "Practical non-monotonic knowledge-base system for un-regimented domains: A case-study in digital humanities," *Information Processing & Management*, vol. 57, no. 3, p. 102214, 2020. https://doi.org/10.1016/j.ipm.2020.102214

[11] D. Idnani, A. Kubadia, Y. Jain, and P.P. Churi, "Experience of conducting online test during COVID-19 lockdown: A case study of NMIMS university," *International Journal of Engineering Pedagogy*, vol. 11, no. 1, pp. 49–63, 2021. https://doi.org/10.3991/ijep.v11i1.15215

[12] Y. Fan, H. Cui, Z. Lou, J. Teng, Z. Tang, and J. Peng, "Base solder voids identification of IGBT modules using case temperature," *Microelectronics Reliability*, vol. 115, p. 113968, 2020. https://doi.org/10.1016/j.microrel.2020.113968

[13] J. Gao, P. Jin, Y. Sheng, and P. An, "A case study on crack propagation law of cement stabilised macadam base," *International Journal of Pavement Engineering*, vol. 21, no. 4, pp. 516–523, 2020. https://doi.org/10.1080/10298436.2018.1492135

[14] N.I.P. Pratama, G. Triyono, M.B. Hasanah, N.B. Kenzhaliyev, O.N. Kosherbayeva, A.K. Kassymova, and M.N.A. Azman, "Development of construction engineering teaching course by using VR-AR: A case study of polytechnic in Indonesia," *International Journal of Interactive Mobile Technologies*, vol. 16, no. 14, pp. 93–106, 2022. https://doi.org/10.3991/ijim.v16i14.30539

[15] A.N. Al Ahmad and R.A. Obeidallah, "Studying the effectiveness of a proposed methodology for teaching programming labs online and students' perspectives toward it during COVID-19: A case study of Hashemite university," *International Journal of Interactive Mobile Technologies*, vol. 16, no. 5, pp. 62–80, 2022. https://doi.org/10.3991/ijim.v16i05.27123

[16] T. Judiantono and M.I.P. Susanto, "Identification of transportation movement patterns based on the home base approach: A case study of Lembang District, West Bandung Regency," *IOP Conference Series: Earth and Environmental Science*, vol. 447, no. 1, p. 012074, 2020. https://doi.org/10.1088/1755-1315/447/1/012074

[17] R. Wu and J. Wang, "Digital management of teaching cases in colleges and universities based on cluster analysis," *International Journal of Emerging Technologies in Learning*, vol. 18, no. 10, pp. 264–279, 2023. https://doi.org/10.3991/ijet.v18i10.40241

[18] G. Sun, T. Sawaragi, Y. Horiguchi, and H. Nakanishi, "Knowledge-intensive teaching assistance system for industrial robots using case-based reasoning and explanation-based learning," *IFAC Proceedings*, vol. 47, no. 3, pp. 4535–4540, 2014. https://doi.org/10.3182/20140824-6-ZA-1003.01492

[19] Hua, Shuyang, and Zhongwei Ren, "'Online + Offline' course teaching based on case teaching method: A case study of entrepreneurship education course," *International Journal of Emerging Technologies in Learning*, vol. 15, no. 10, pp. 69–85, 2020. https://doi.org/10.3991/ijet.v15i10.13999

[20] Q.R. Xiao, "Evaluation model of case teaching effect of engineering cost based on data mining," In *e-Learning, e-Education, and Online Training: 6th EAI International Conference, eLEOT2020*, Changsha, China, pp. 16–28, 2020. https://doi.org/10.1007/978-3-030-63952-5_2

[21] L.J. Kan, J.G. Gao, and J.H. Xu, "Study on case-based teaching in military academies based on information technologies," In *2020 International Conference on Information Science and Education (ICISE-IE)*, Sanya, China, pp. 531–534, 2020. https://doi.org/10.1109/ICISE51755.2020.00119

[22] D. Ameta, S. Tiwari, and P. Singh, "A preliminary study on case-based learning teaching pedagogy: Scope in SE education," In *Proceedings of the 13th Innovations in Software Engineering Conference on Formerly known as India Software Engineering Conference*, Jabalpur, India, pp. 1–12, 2020. https://doi.org/10.1145/3385032.3385045

[23] H. Li, "Design of deep learning experiment teaching case based on emg signal analysis," In *Proceedings of the 4th International Conference on Advances in Artificial Intelligence*, London, United Kingdom, pp. 29–34, 2020. https://doi.org/10.1145/3441417.3441432

[24] G. Karalekas, S. Vologiannidis, and J. Kalomiros, "Europa: A case study for teaching sensors, data acquisition and robotics via a ROS-based educational robot," *Sensors*, vol. 20, no. 9, p. 2469, 2020. https://doi.org/10.3390/s20092469

8 AUTHORS

Shanshan Wang graduated from Hebei University of Technology with a master's degree. She is currently working at Shijiazhuang University of Applied Technology, focusing on business administration (E-mail: 2017010823@sjzpt.edu.cn; ORCID: https://orcid.org/0009-0001-7343-0197).

Xiangpeng Chang got his Master's degree in Engineering from Wuhan University of Technology and has been working at Shijiazhuang University of Applied Technology since 2001. His main fields of research include business management of tourism and hotels; and laws and regulations on tourism and hotels (E-mail: 2001000251@sjzpt.edu.cn; ORCID: https://orcid.org/0009-0003-3422-0351).

Dong Ding graduated from Hebei University of Science & Technology with a Master's degree. She is currently working at Shijiazhuang University of Applied Technology, and her research interest include business administration (E-mail: 2018010870@sjzpt.edu.cn; ORCID: https://orcid.org/0009-0002-5004-8611).