AIMS Mathematics, 8(8): 18491–18511.
DOI: 10.3934/math.2023940
Received: 11 February 2023
Revised: 18 May 2023
Accepted: 22 May 2023
Published: 31 May 2023

http://www.aimspress.com/journal/Math

Research article

Visual analysis of knowledge graph based on fuzzy sets in Chinese martial arts routines

Jun Jiang[1], Junjie Lv[1,*] and Muhammad Bilal Khan[2,*]

[1] Institute of Physical Education and Health, Yulin Normal University, Yulin 537000, China
[2] Department of Mathematics, COMSATS University Islamabad, Islamabad 4400, Pakistan

*** Correspondence:** Email: lvjunjie@ylu.edu.cn; bilal42742@gmail.com.

Abstract: There are many schools of Chinese martial arts routines and complex movements; research on this topic is more geared toward Taijiquan (a kind of traditional Chinese shadow boxing), which is a more well-known type of martial arts. Therefore, the purpose of this paper is to visually analyze the research of Chinese martial arts routines based on the knowledge graph method and to propose a knowledge graph method based on the fuzzy set theory, which is called the transF model throughout this paper. The transF model used the fuzzy relational operation of vectors to not only reduce the computational complexity, but to also better integrate multi-dimensional data, especially when the training set is not particularly sufficient. For the visual analysis of Chinese martial arts routines, this paper selected the 16-year data from 2005 to 2020 as the analysis sample, analyzed high-yield institutions and high-yield authors, and conducted a centrality analysis of the whole dataset. From the structure of the knowledge graph, traditional martial arts are the core part of Chinese martial arts, with a centrality of 0.14. Competitive martial arts are the main branch of Chinese martial arts and the third core after Tai Chi and traditional martial arts, with a centrality of 0.41, which is higher than that of traditional martial arts. This shows its importance in martial arts research.

Keywords: Chinese martial arts routines; visual analysis; knowledge graph; fuzzy set theory; martial arts routine visualization
Mathematics Subject Classification: 03E72, 90B50, 90C31

1. Introduction

Martial arts routines are an important part of traditional Chinese national sports. It is based on

offensive and defensive movements such as kicking, hitting, throwing, holding, and stabbing. A complete set of exercise forms was created by the law of movement changes. As a unique traditional sport in China, the only way to forward with the development of the times is to develop martial arts routines. Undoubtedly, theoretical research would play a huge role in guiding and promoting the development of martial arts routines [1]. The purpose of this paper is to intuitively analyze available research on Chinese martial arts routines based on the knowledge graph method and to propose a knowledge graph method based on fuzzy set theory, namely the transF model in throughout this article. This paper hopes that the establishment of a project on traditional Chinese martial arts will inevitably promote the research and organization of the inheritance, development, and protection of Chinese martial arts routines.

In this paper, the main significance of using the knowledge graph method to analyze martial arts routines lies in the following points: First, through a large number of documents, a detailed and objective description and analysis of the development status of martial arts routines research is carried out, which is conducive to grasping the current situation of martial arts routines research; second, the frontiers of martial arts routine research to provide a theoretical basis and scientific basis for future research and development of martial arts routines could be explored; third, grasping the knowledge base of martial arts routine research and understanding the historical context of martial arts routine development would help researchers grasp the trend of martial arts routine research, help researchers formulate scientific research plans, and provide future development planning and scientific research for martial arts routine research. The project layout provides support in decision-making.

Martial arts has a long history in China, and there are many studies on martial arts routines. Han [2] developed a scientific martial arts system training method by testing the anaerobic metabolism capacity of the upper limbs of male martial arts routine athletes. Sun [3] mainly discussed the real-time extraction and simulation algorithm of VR image targets for the difficult movement technology of martial arts routines. Gong [4] used martial arts routines to study the dynamic and static balance ability of university students. The study found that martial arts routines significantly improved the balance ability of university students when they closed their eyes and stood on one foot, increased the time to maintain their balance, improved the dynamic balance of balance beam walking, and improved long-term martial arts routines. Wehner [5] utilized a meta-examination technique to examine the impact system of Yoga preparation in hand-to-hand fighting schedules on their actual well-being. Song [6] believed that the final result of the martial arts routine competition was mainly determined by the difficulty of the movements and gradually determined by the level of practice. At the same time, the study also believed that in the case where both the quality and the difficulty of the action have obtained full marks, and it is difficult for the referee to distinguish and measure the final competition performance of each athlete through the level of the exercise. Therefore, it is necessary to formulate competition rules that keep pace with the times. Through the summary of relevant research, there is a lack of literature on quantitative research on martial arts routines, and most of these studies are on the development of martial arts routines in a certain field.

A knowledge graph is an important method of visualization and can play an important role in quantitative research. Many scholars have done research on the application of knowledge graphs. Zheng [7] analyzed the characteristics of China's sports power through knowledge graph drawing. Results from this study have an important significance on the construction of China’s sports power in the new era. Wang and Hu [8] analyzed the international organizations and organizations in the field of Olympic sports by using 1,196 documents about Olympic sports collected by scientific databases and by using the Cite Space knowledge graph. Wang and Liu [9] used the knowledge graph and Cite

space III software to visually analyze the international sports genes. The results showed that the United States is in a leading position in the field of sports genetics research. The continuous integration of sports science, agronomy, genetics, medicine, and physiology has created a new situation in international sports genetics research. Ji [10] proved that knowledge graphs are both vivid and informative for basketball game visualization, which can enhance the expressiveness of basketball game news. Chiu [11] used VOS viewer for the analysis, and this work contributed to an improved understanding of e-sports research, thereby further providing new perspectives for e-sports sustainability research. From the above literature, it can be found that the relevant authors have conducted various in-depth analyses of the relevant research directions of sports disciplines, including the construction of a strong sports country, the international Olympic movement, sports genes, basketball, e-sports, and other related fields.

Zhang [12] introduced a multi-scale dynamic convolutional network (M-DCN) model for knowledge graph embedding. This model has a first-class performance and can generate richer and more expressive feature embeddings than similar models. The subject entities and relationships in M-DCN are embedded in the input layer and combined in alternating patterns, which helps to extract additional feature interactions and to improve expressiveness. Li [13] proposed the simplified heterogeneous graph neural network (SHGNet), which is a universal framework that abandons two standard operations in graph neural networks (GNN), including transformation matrices and nonlinear activation. Notably, SHGNet only adopts the basic components of neighborhood aggregation in GNN and incorporates relational features into feature propagation. In addition, to capture complex structures, SHGNet uses a hierarchical aggregation architecture, including node aggregation and relationship weighting. Therefore, the proposed model could handle each relationship differently and selectively aggregate information features. Li [14] proposed a new heterogeneous graph neural network framework based on the hierarchical attention mechanism, including entity level, relationship level, and self level attention. Therefore, the proposed model could selectively aggregate information features and fully weight them. Then, the embedding of the learned entities and relationships can be used for downstream tasks. Numerous experimental results on various heterogeneous graphical tasks have shown that the proposed model has a superior performance compared to several state-of-the-art methods.

However, there is not much research on martial arts routines.

In short, only a very few scholars and researchers in China have used the theory and method of "knowledge graphs" to analyze any research on martial arts routines. Therefore, this paper will take "martial arts routine" as the basis of the relevant literature in the research field and use it as the analysis method of the knowledge graph. The innovation of this paper is to systematically sort out the mainstream academic groups and important academic documents of martial arts routine research, explore the research hotspots, cutting-edge issues, and knowledge base of Chinese martial arts routines, and reveal the gap in the development of Chinese martial arts routines. This paper also understood the development status of martial arts routines and found the problems and deficiencies in the development of martial arts routines, which provided a theoretical reference for the development and innovation of scientific research on martial arts routines.

2. Visualization of martial arts routines

2.1. Martial arts routines

Martial arts are one of the earliest sports disciplines in China, with a long history, profound

cultural heritage, and a solid foundation for the people. After the 21st century, due to the objective needs of society and changes within the discipline, the martial arts routine movement has been greatly developed, and a relatively complete theoretical system has been formed. Most of the current martial arts routines are based on performance. Some scholars have pointed out that although Chinese martial arts have made major breakthroughs, any external effects cannot be exerted due to a lack of internal motivation [15,16]. This paper summarizes research on martial arts routines in stages, comprehensively sorts out and analyzes the overall situation of martial arts routines research since the new century, and finds various hot spots and frontier issues. This paper considers the problems existing in the research and practice of martial arts routines, which would play an important role in promoting the future research and development of martial arts routines.

2.2. Knowledge graph method

This paper uses citation analysis, co-citation analysis, word frequency analysis, and other methods. Citation analysis is a study of the citation frequency of different research objects such as scientific journals, papers, authors, etc., so as to reveal their quantitative characteristics and internal laws. Citation analysis can help scholars rationalize the development of disciplinary knowledge. Co-citation analysis refers to the fact that two documents are cited by other documents at the same time, resulting in a co-citation relationship [17]. The number of co-citations refers to the number of times the two documents are cited at the same time. Word frequency analysis is a method that uses statistical data to perform statistical analysis on keywords [18]. Through the analysis of vocabulary frequency, it can understand the current research hotspots. This paper combines word frequency analysis with co-citation analysis to showcase various cutting-edge research fields of Chinese martial arts routines.

In the book and information industry, knowledge graphs are called knowledge domain visualization or knowledge domain mapping maps, which are a series of different graphs. It shows the knowledge development process and structural relationship. It uses visualization technology to describe knowledge resources and their carriers, mining, analyzing, constructing, mapping, and displaying knowledge and their interrelationships. The knowledge graph is the combination of theories and methods of applied mathematics, graphics, information visualization technology, information science, and other disciplines with a quantitative citation analysis, co-occurrence analysis, and other methods. Visual maps can be used to vividly display the core structure, development history, frontier fields, and modern theories of the overall knowledge structure of the discipline, so as to achieve multi-disciplinary integration.

2.3. Knowledge graph construction based on fuzzy set theory

The existing fully supervised deep learning relies on a large number of labeled data to train how to learn feature extraction. However, labeling video data requires a lot of labor costs. Self-supervised learning only requires video data without labeling information, which greatly reduces the requirements for the quality of training data and greatly increases the number of training data. 3. The existing self-supervised learning technology for video feature learning focuses on high-level semantic features and ignores low-level representations, that is, it only applies the self-supervised learning method to the high-level features extracted from deep networks to learn video features. However, the lower-level features can better represent the temporal relationship of the video, which is very important for universal video understanding.

With the increasing popularity of knowledge graphs in the network, more and more researchers have begun to study them [19]. The main idea of the TransF model is to use attention mechanisms to capture the contextual dependencies of each position in the input sequence, rather than relying on loop structures like recurrent neural networks such as Long Short-Term Memory (LSTM). On this basis, TransE technology has been widely used [20]. On this basis, this paper combines the fuzzy theory with the existing deep learning-based TransE model to establish the TransF model. In TransF, two fuzzy vectors, an entity, and a fuzzy graph are respectively constructed, and a head and tail entity is mapped to a fuzzy space composed of relations.

In the knowledge graph model, when describing the loss function of the pattern similarity, the L_P norm based on the Murkowski distance is commonly used [21], and its calculation formula is as follows:

$$L_P(A,B) = \left[\sum_{i=1}^{n} |a_i - b_i|^p\right]^{\frac{1}{p}}. \tag{1}$$

It can be seen from formula (1) that the value of the parameter p has a very important relationship with the norm, and different p corresponds to different distance algorithms.

Synthesis of Fuzzy Relationships

If the fuzzy sets U and V have

$$f: U \to F(V), \tag{2}$$

then f is called the mapping from U to V. The fuzzy relationship between A and R can be regarded as a fuzzy relationship from A to U. If R is a fuzzy relationship from U to V, then the synthesis of A and R is the ambiguity from A to V, that is, the fuzzy set on V [22]:

$$B = A \circ R. \tag{3}$$

Specifically, for $A, R \in F(U)$

$$A \circ R = \bigvee_{u \in U} (A(u) \wedge R(u)). \tag{4}$$

It can be seen from this that R is the mapping from the F set to the F set. Then: $F(U) \to F(V)$ is the transformation between U to V. By converting the F set A into T, the F set B: $T(A) = B$ can be obtained. At this time, B is called the pixel of A on F, and A is the original image of B. The operation of F is essentially the comprehensive operation of the fuzzy set A and the fuzzy relation R.

Some advantages of this are that it can improve the calculation efficiency of applications, increase the diversity of downstream application design, and use knowledge graph embedding as the pertaining vector input of downstream applications.

From the research results of this paper, the commonly used TransE-based models relies on deep learning and does not introduce the fuzzy correlation in fuzzy mathematics into the research of knowledge graphs. This section presents a TransF algorithm that optimizes the initial model for deep learning using fuzzy sets. Therefore, the associations between head-tail entities set in this section are maintained by fuzzy relations. In this section, the head-tail entities are projected to their corresponding subspaces of fuzzy relation R, respectively, and then the corresponding knowledge graph model is constructed with the TransE structure of TransF.

The goal of TransF is to handle one to many/many to one/many to many relationships without increasing the complexity and training difficulty of the patterns.

Its basic idea is to interpret relations as transformation operations on hyperplane. Each relationship has two vectors, as shown in Figure 1: The norm vector Wr of the hyperplane and the

translation vector (dr) on the hyperplane. Figure 1a shows the relationship space corresponding to TransE, and Figure 1b shows the relationship space corresponding to TransF.

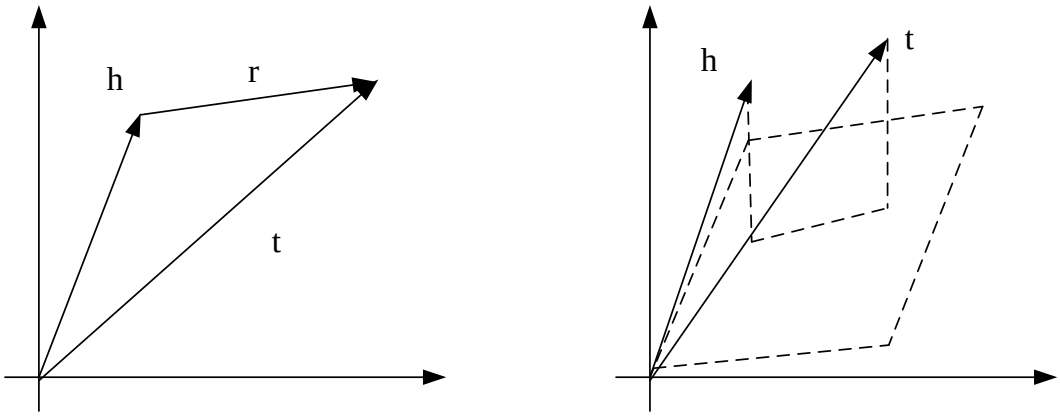


Figure 1. Simple illusrtation of TransE and TransF.

In this paper, each head vector (h) and tail vector (t) are projected onto the hyperplane to obtain new vectors (h $\vee$ and t $\vee$). There is a relation (d_r) in the hyperplane which we can train in a similar way to the TransE model. This can avoid two entities (head or tail) approaching during model training and handle one-to-many/many-to-one/many-to-many relationships. It is difficult for TransE to discern the complexity and training difficulty of the patterns and to obtain more accurate positions; TransF can accurately identify the relationship vectors corresponding to each cluster.

On this basis, this paper presents the basic structure of a fuzzy-based martial arts routine knowledge graph TransF model, as shown in Figure 2.

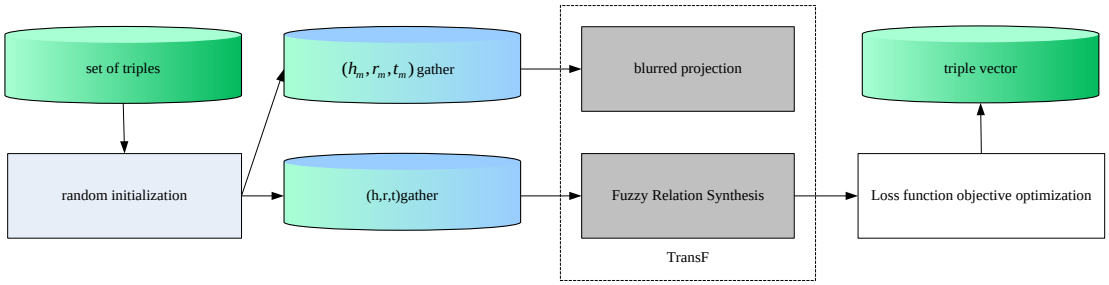


Figure 2. The basic framework of knowledge graph model based on fuzzy theory.

The central idea of the model is to randomly initialize the triples into two vectors: One for building the triples and one for building the fuzzy projections of the triples. Finally, the knowledge of triples can be constructed using the TransF method according to the loss function of the transformation model.

It is assumed that $X = (X_1, \mathrm{K}, X_n)$ is an observation sequence and $Y = (Y_1, \Lambda, Y_n)$ is a state sequence; the two correspond one-to-one, that is, X_i corresponds to the mark Y_i. The parameterized form $P(y|x)$ of the linear chain conditional random field is:

$$P(y|x) = \frac{1}{Z(x)} \exp\left(\sum_k \lambda_k T + \sum_i u_i S \right) \tag{5}$$

$$T = t_k(y_{i-1}, y_i, x, i) \tag{6}$$

$$S = s_i(y_i, x, i) \tag{7}$$

where T represents the transition characteristics of the marker position y_{i-1} and y_i of the observation sequence, and only involves the previous node and the current node; S is a state characteristic function, which is only related to the information of the current node; and λ_k and u_i represents the weighting of the individual characteristic functions, where Z(x) is the normalization factor described below:

$$Z(x) = \sum_y \exp\left(\sum_k \lambda_k T + \sum_l u_l S \right). \tag{8}$$

In the sequence labeling problem, the feature function has the greatest impact on the labeling accuracy and often starts from the characteristics of the sequence; in different situations, its manifestations are not the same, so the feature functions are also very rich.

In this paper, the conditional random field (CRF) method is used to observe the sequence labeling problem. The concept of the CRF model is derived from the Hidden Markov Model (HMM), which is a method where the target sequence based on the observation sequence [23]. In CRF, the most commonly used field is the linear chain state random airport, as shown in Figure 3.

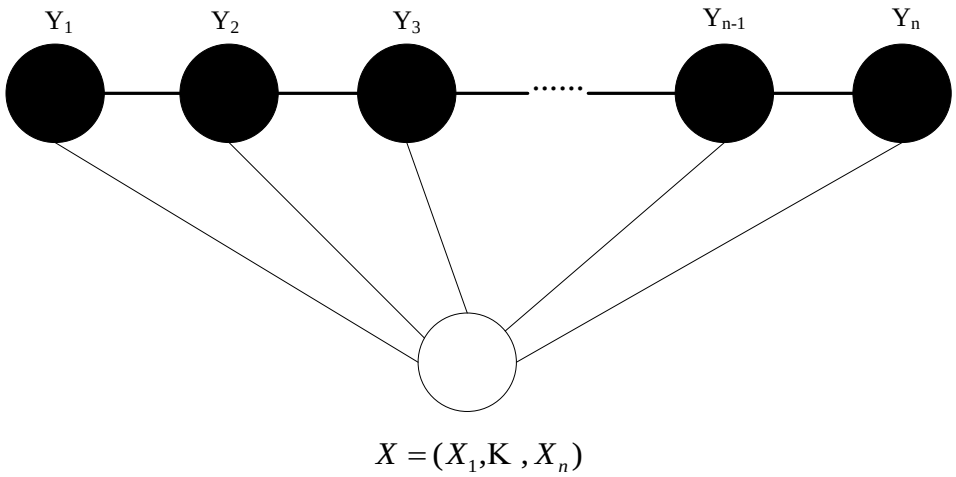


Figure 3. Conditional random field structure.

In the sequence labeling problem, linear chain conditional random fields can take advantage of the former and latter features. Combined with other external information, it can better solve the problem of the strong independence assumption of HMM and the shortcomings of the maximum entropy (ME) model in labeling.

Similarly, this paper utilizes canonical correlation analysis (CCA) to observe the degree of the correlation between two variables. The neural network must optimize a very complex nonlinear

model, and there is basically no global optimal solution, in which initialization plays a very important role. Among them, the selection of the initial point can sometimes determine whether the algorithm converges. When converging, the initial point can determine how fast learning converges and whether it can converge to a point with either a high or low cost. Too large of an initialization leads to a gradient explosion, and too small of an initialization leads to a gradient disappearance.

CCA is a statistical method to study the relationship between two groups of variables, which can test the degree of correlation between variables in different groups at the same time, such as the degree of correlation between two groups (X, Y) [24]. Assuming a random variable $(X, Y \in R)$, where X represents a word and Y represents a background, CCA can find the maximum correlation between these two variables in the k-dimensional projection vector.

The basic idea of CCA is to reduce the dimension of high-dimensional data from one dimension, and then use the correlation coefficient to analyze the correlation. The correlation of two one-dimensional datasets x and y is:

$$\rho(x, y) = \frac{\mathrm{cov}(x, y)}{\sqrt{D(x)}\sqrt{D(y)}}. \tag{9}$$

Here, cov(x, y) is the covariance of x and y, and D(x) and D(y) are the changes in x and y, respectively. The correlation coefficient ρ is in the range of [-1, 1]; the closer it is to 1 or -1, the stronger the correlation, and the closer it is to 0, the weaker the correlation.

In the case of high-dimensional data, a correlation analysis cannot be performed directly, and the high-dimensional data must be converted into one-dimensional data. For multi-dimensional variables X and Y, the eigenvectors a and b are introduced into one dimension, that is:

$$X' = a^T X; Y' = b^T Y. \tag{10}$$

The optimization purpose of CCA is to maximize the correlation coefficient $\rho(X', Y')$, so that it has the eigenvectors a, b:

$$\arg\max_{ab} \rho(X', Y') = \arg\max_{ab}[{}^{cov(X', Y')}\Big/_{(\sqrt{D(X')} \cdot \sqrt{D(Y')})}]. \tag{11}$$

Before the calculation, the original data must be normalized so that the mean value of X and Y changes from 0 to 1. Therefore, the X', Y'-related calculations are as follows:

$$\mathrm{cov}(X', Y') = \mathrm{cov}(a^T X, b^T Y) = E[(a^T X)(b^T Y)^T] = a^T E(XY^T)b \tag{12}$$

$$D(X') = D(a^T X) = a^T E(XX^T)a \tag{13}$$

$$D(Y') = D(b^T Y) = b^T E(YY^T)b \tag{14}$$

$$D(X) = \mathrm{cov}(X, X) = E(XX^T) \tag{15}$$

$$D(Y) = \mathrm{cov}(Y, Y) = E(YY^T) \tag{16}$$

$$\mathrm{cov}(X, Y) = E(XY^T) \tag{17}$$

$$\mathrm{cov}(Y, X) = E(YX^T) \tag{18}$$

According to the relevant calculation rule, the correlation coefficient $\rho(X', Y')$ can be

transformed into:

$$\arg\max_{ab}\rho(X',Y')=\arg\max_{ab}[{a^T\cdot cov(X,Y)b}/{(\sqrt{a^T\cdot cov(X,X)a}\cdot\sqrt{b^T\cdot cov(Y,Y)b})}]. \quad (19)$$

When the maximum value ρ of this optimization objective is obtained, the corresponding a and b are the feature vectors during the dimensionality reduction.

Additionally, CCA is a multivariate statistical method to study the correlation between two groups of variables, which can reveal the internal relationship between two groups of variables. In a univariate statistical analysis, the correlation coefficient Q is used to measure the linear correlation between two random variables. The multiple correlation coefficient is used to study the linear correlation between a random variable and multiple random variables. However, none of these methods can be used to study the correlation between the two groups of variables, so CCA is proposed. Its basic idea is very similar to that of a principal component analysis. First, it can find the linear combination of variables in each group of variables, so that the linear combination of two groups of variables has the largest correlation coefficient. Then, one can select another pair of linear combinations that are not related to the selected pair of linear combinations and maximize their correlation coefficient, and so on, until the correlation between the two groups of variables is extracted. The selected linear combination symmetry is the gauge variable, and its correlation coefficient is called the gauge correlation coefficient.

2.4. Knowledge graph based on fuzzy relation synthesis

To measure the quality of an algorithm, we can comprehensively evaluate it through the three aspects mentioned above. It can find algorithms with a short running time, low resource consumption, easy understanding, and easy implementation from several candidate algorithms. However, the reality is not satisfactory. Often, a seemingly simple algorithm runs much slower than a formally complex algorithm; however, an algorithm with a short running time often takes up more resources. Therefore, different algorithms need to be selected in different situations. In the real-time system, if the system response time is high, one can try to choose the algorithm with a smaller execution time. When the amount of data processing is large and the storage space is small, one can try to choose a space-saving algorithm. This paper mainly discusses the time characteristics of the algorithm, and gives the measurement index of the time complexity of the algorithm.

Suppose there is a triple $(h_i,r_i,t_i),i=1,\Lambda\ ,p$, where h_i represents the head entity, r_i represents the relationship, and t_i represents the tail entity. Similar to the classical fuzzy set theory, the dual fuzzy sets involved in this chapter are still expressed in the form of the product of fuzzy matrices when synthesizing fuzzy relations:

$$t_{f_r}=l_t\circ f_r=\vee(l_t\wedge f_r)\,. \quad (20)$$

When $-1\le a\le b\le 1$ is set, there are:

$$\begin{cases}a\wedge b=a\\ a\vee b=b\end{cases}. \quad (21)$$

In this section, each element (entity and association) in the triple is divided into two vectors: One vector represents the element and one vector constructs a mapping matrix. Specifically, the

vector corresponding to each triple (h, r, t), are $h\ （h_m）, r\ （r_m）, t\ （t_m）$. The m-subscript vector is the unit used to construct the mapping matrix, and the vector without the m-subscript refers to the vector of the unit itself, in a specific vector structure: $h\ （h_m）, t\ （t_m）\in R^k$, $r\ （r_m）\in R^n$, where k and n are the dimensions of the entity vector and relation vector. Figure 4 shows the basic principle of the whole algorithm.

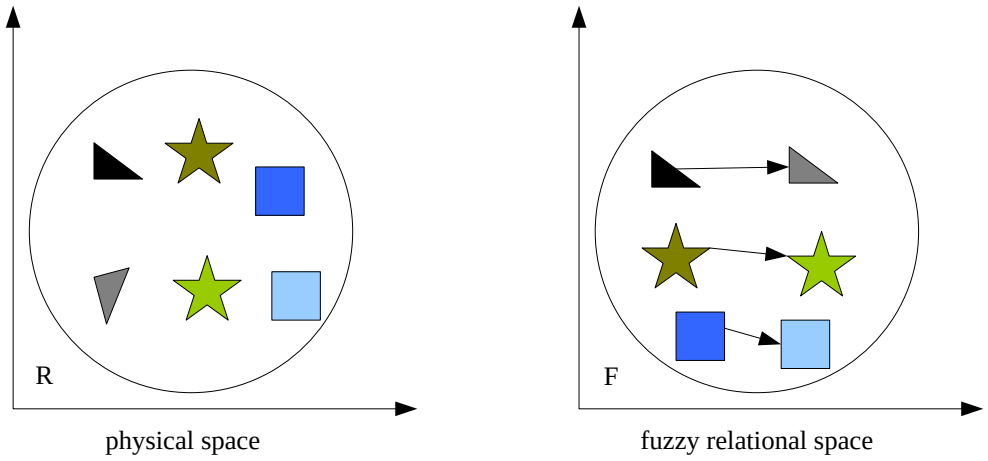


Figure 4. Basic principle of TransF.

When introducing the basic principle of this paper, the head entity h and tail entity t can be projected into the fuzzy space constructed based on relation r, respectively, so as to obtain a greater differentiation after mapping the similar elements in the original head entity and tail entity into the fuzzy space.

3. Visual analysis of Chinese martial arts routines

3.1. Objects, methods and data sources

3.1.1. Research object

In this paper, the relevant literature on martial arts routines collected by China National Knowledge Infrastructure (CNKI) and Chinese Social Science Citation Index (CSSCI) were used for analysis. In the National Library of China, the library of Beijing Sport University, CNKI, the Chinese Social Science Citation Database, etc., systematic collation, statistics, and analysis have been carried out, which has laid a solid foundation for future research work. Cite Space (i.e., Citation Space) is a visual citation analysis software. It focuses on analyzing the potential knowledge contained in scientific literature and gradually develops it in the context of scientific metrology and data visualization. The structure, regularity, and distribution of scientific knowledge are presented through visualization methods, so the visual graphs analyzed by such methods would also become scientific knowledge graphs. It is used to explore the research hotspots, research frontiers, knowledge base (key literature), main authors, and institutions of a certain research field, and to help predict the

future development trend of a certain research field.

3.1.2. Data sources

The data used in this article are from the CNKI database and the Chinese social science citation database. The main research materials in this paper are from CNKI, and other documents are from the Chinese Social Science Literature Retrieval Database.

3.1.3. Data processing tools

This paper adopts three types of knowledge graph software: Cite Space, Bibexcel, and Pajek.

Cite Space (Version 5.1 R6. SE) is a visual knowledge graph software for multiple, complex, and dynamic network analysis. The Cite Space system realizes the reading of various commonly used documents on the Java platform. This paper uses three visual graphs of cluster view, timeline view, and time segment to display the evolution of knowledge in different periods and uses mutation detection technology to detect knowledge development trends.

Bibexcel can read a large amount of data from a large number of databases such as Wos and Scopus, and uses different methods to simplify and normalize the data. It is a general data measurement and analysis tool.

Pajek (version 645.01) is a program running on Windows, which has a strong application value in large-scale network visualization. Based on the knowledge graph technology of fuzzy concentration, this paper studied the current Chinese martial arts routines.

3.1.4. Data handler

Through the comparison and analysis of various retrieval methods, this paper adopts the retrieval strategy of “martial arts”, “competitive martial arts”, “tai chi”, “martial arts routines”, and “traditional martial arts”. The retrieval period is from 2005-2020, and is regarded as an academic paper. By sorting out the relevant literature, it eliminated the relevant materials such as “martial arts sanda” and “journal cover” that are not related to this research and collected a total of 949 authentic and valid materials. Then, it imported the title, keywords, authors, journals, article abstracts, citations, etc., into the computer and stored them in a TXT format.

Stored data: It named 949 documents downloaded from CSSCI as Download_2017.txt.

Input data: It imported the saved literature data into Cite Space, Bibexcel, Pajek, etc. in the required order.

Visual analysis of knowledge graph: According to the requirements of this article, it set the time, nodes, screening criteria, cutting methods, etc. Through this software, the corresponding knowledge map can be obtained, and the obtained information can be counted and analyzed.

3.2. Literature distribution characteristics of martial arts routine

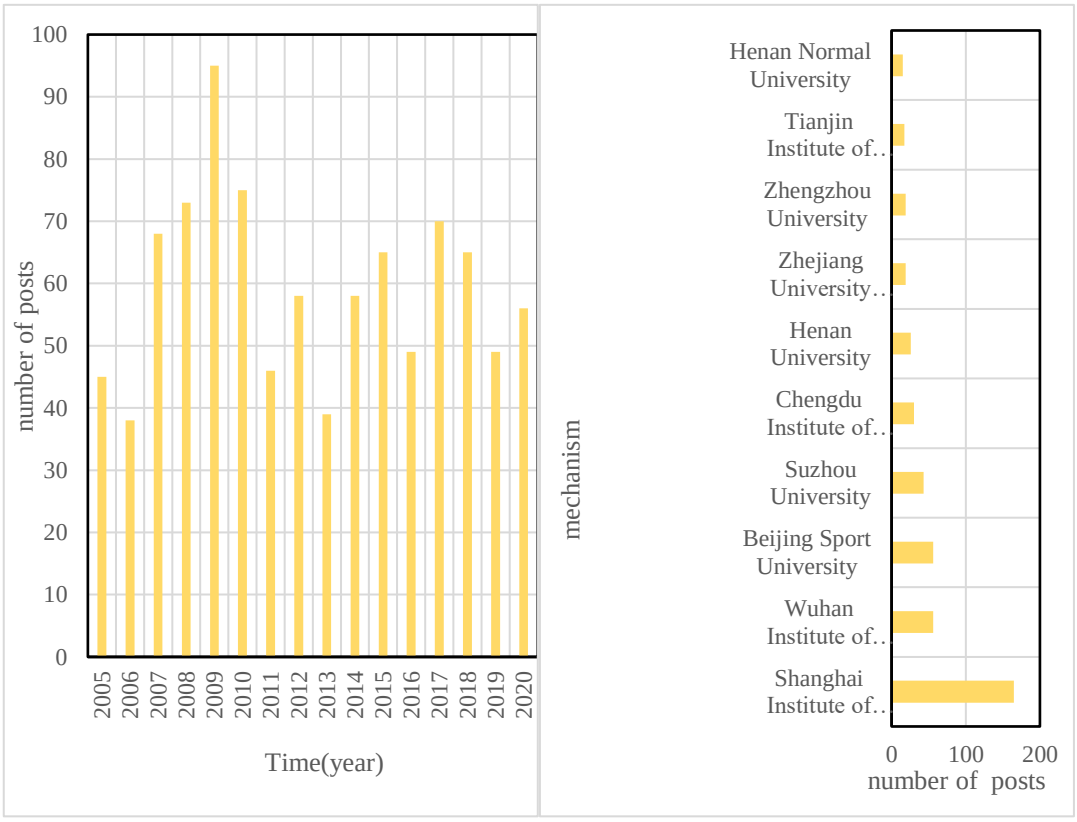
3.2.1. Number of documents and publishing institutions

The change in the number of documents is an important manifestation of the scientific knowledge of martial arts routines. By counting the amount of literature published each year, it can intuitively reflect the time distribution of the literature on martial arts routine research, and make

predictions about its development trend. Figure 5A shows the annual distribution of Chinese martial arts design research literature from 2005 to 2020.

Research institutions are an important platform for scientific research work and a powerful carrier to promote the development of disciplines. Through the analysis of martial arts design institutions, it can better understand the distribution of scientific research forces in this field. In order to better reflect the research results of major universities, this paper analyzed the names and volumes of 10 institutions that published papers from 2005 to 2020, as shown in Figure 5B. The scientific research units involved in this article include the Martial Arts Academy of Beijing Sport University, the Graduate School of Beijing Sport University, Beijing Sport University, and other secondary colleges of the university.

As shown in Figure 5, the number of literatures fluctuates, which is related to the research hotspots each year, and for the publishing institutions, the publication of Chinese martial arts routine literature is mainly concentrated in individual institutions.



A. Quantity of literature research on martial arts routines in the past 16 years;
B. Situation of high-yield research institutions

Figure 5. Number of documents and issuing institutions.

3.2.2. Highly prolific authors and highly cited articles

This paper selects 5 or more authors as high-yield authors, and obtains a high-yield author group, as shown in Table 1. Due to limited space, only the top five are shown. By comparison, one can see that in high-yield scientific research units, most of the high-yield writers come from

high-level institutions, which is also in line with the characteristics of high-yield writers.

Table 1. Statistics of high-yield authors (top five).

Serial number	Author	Affiliation	Post volume
1	Qiu**	Shanghai Institute of Physical Education	37
2	King*	Wuhan Institute of Physical Education	27
3	Young**	Zhejiang University of Technology	26
4	Yu**	Shanghai Institute of Physical Education	20
5	Guo**	Shanghai Institute of Physical Education	16

The difference between prolific and highly cited authors is that prolific authors are highly productive, high-level writers in a field, and highly cited authors are authorities in a field. The number of citations is an important measure of the influence of a researcher. The more citations, the higher the degree of attention from other scholars. According to the ranking of the number of citations, the authors with more than 20 citations were selected as the highly cited authors, leading to the information in Table 2.

Table 2. Statistics of highly cited authors (top five).

Serial number	Author	First Cited (year)	post volume
1	Qiu**	2000	154
2	king*	2002	104
3	temperature*	2000	80
4	week**	2004	76
5	National Sports martial arts Academy	2002	58

As shown in Table 2, Qiu and Wang's articles have been cited more than 100 times; they can be noted as the most influential writers on martial arts routine research in China. In addition to individual authors, the Martial Arts Research Institute of the National Sports Commission and the National Sports School Textbook Committee are also cited more frequently, which shows that their publications occupy a pivotal position in Chinese martial arts research.

3.2.3. Analysis of journals with articles

By analyzing the papers contained in a subject area, it is possible to determine the main distribution of the papers contained in it. Figure 6 shows the major academic journals in China from 2005 to 2020. Knowing the sources and journal distribution of Chinese martial arts routine research literature would help researchers to grasp their distribution, so as to formulate better retrieval and submission plans.

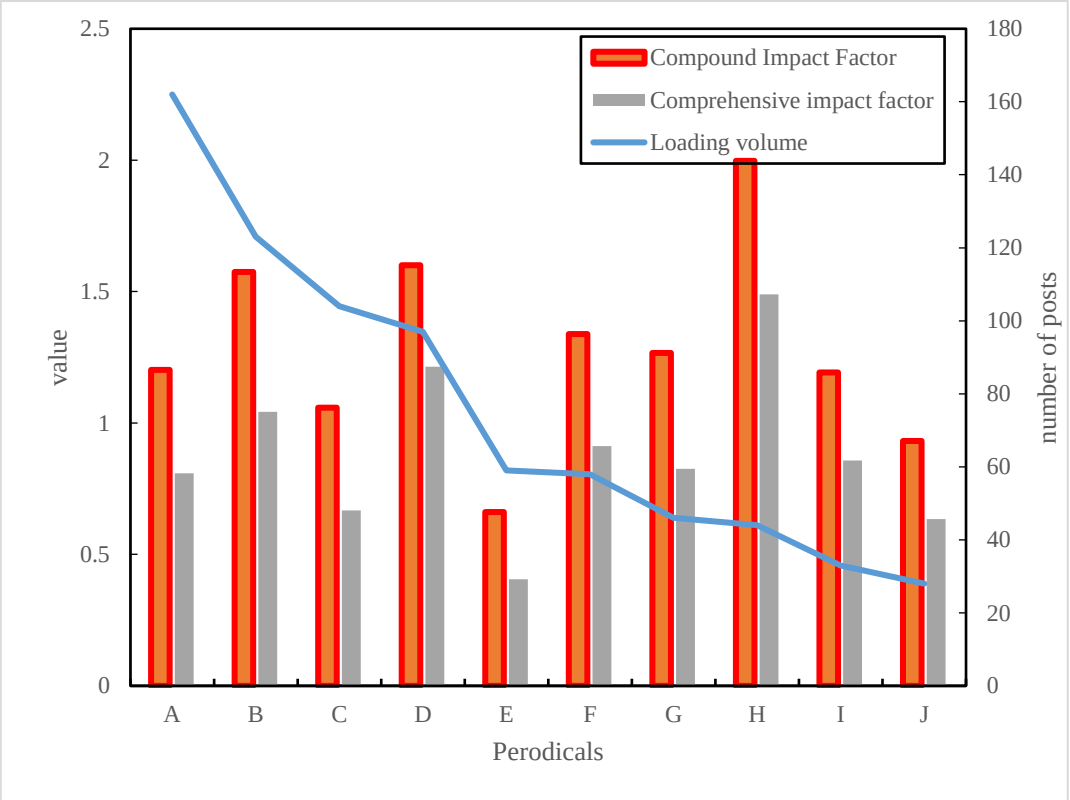


Figure 6. The main journals with articles on martial arts routine research.

Bars A–J in Figure 6 represent these ten journals, respectively: Journal of Wuhan Institute of Physical Education, Journal of Beijing Sports University, Journal of Chengdu Institute of Physical Education, Journal of Shanghai Institute of Physical Education, Sports Culture Guide, Sports Journal, China Sports Science and Technology, Sports Science, and Tianjin Sports Institute Journal of Xi'an Institute of Physical Education.

3.3. Scientific research cooperation in martial arts routine

The three levels of scientific cooperation are as follows: macro-level, inter-institutional, and inter-individual collaboration (micro). In view of the fact that this paper is based on research on Chinese martial arts routines, we discuss scientific cooperation from two aspects of inter-institutional cooperation and micro-cooperation.

Inter-agency research collaboration is the main way for researchers to carry out research activities, which has a significant impact on knowledge dissemination. It can be seen from Table 3 that the cooperation among various institutions in China is mainly concentrated in 2-3 institutions. There is little cooperation among more than 4 colleges and universities, which indicates a relatively low level of cooperation in this area.

Table 3. Cooperative papers of research institutions.

Number of institutions	Number of collaborative papers	proportion
1	458	48.261%
2	376	39.621%
3	87	9.168%
4	14	1.475%
5	12	1.264%
6	1	0.105%
8	1	0.105%

Centrality analysis is used to measure the status or rights enjoyed by either individuals or organizations in their own social networks. It is the earliest and most widely used method of social relationship analysis. This paper uses the Pajek software to analyze the centrality of the cooperation network of Chinese martial arts design institutions from three aspects: point centrality, intermediary centrality, and proximity centrality. The size of the point degree centrality determines the position of a node in the entire network, and can represent the influence of a node in the entire network. The size of between ness centrality refers to the minimum number of paths through this node. The higher the middle point of a node, the greater the role it plays in the connection between other nodes. Under the concept of closeness centrality, the larger the size of a node, the less control other nodes have over it. This means that this node can easily connect to other nodes. The results are shown in Table 4. In Table 4, only the top 4 institutions of the point degree centrality are shown.

Table 4. Centrality analysis results of institutional cooperation.

Serial number	Mechanism	Point centrality	Between ness centrality	Closeness centrality
1	Martial Arts College of Shanghai Institute of Physical Education	24	0.361	0.430
2	Henan Normal University School of Physical Education	8	0.209	0.403
3	Martial arts Institute of Wuhan Institute of Physical Education	7	0.192	0.379
4	Soochow University School of Physical Education	6	0.151	0.391

From the results of the centrality analysis, the point-degree centrality value of the Shanghai Institute of Physical Education is 24, and there is a certain connection between the Shanghai Institute of Physical Education and the cooperation network of 24 colleges and universities. This shows that the Shanghai Institute of Physical Education has a strong ability to spread knowledge in “martial arts routines” and has a greater influence.

The Martial Arts College of Shanghai Institute of Physical Education has the highest degree of

intermediary centrality, followed by the Physical Education College of Henan Normal University, which both have strong intermediary centrality in the network. They play a very important role in the whole network and can be compared to a bridge connecting the whole network.

From the closeness centrality, it can be seen that the centrality of Shanghai Institute of Physical Education is the largest, which indicates that it can easily connect with other scientific research units and plays a greater role in the network.

Points 1–9 in Figure 7 represent the Soochow University School of Physical Education, the Graduate School of Shanghai University of Physical Education, the School of Physical Education of Henan Normal University, the School of martial arts of Shanghai University of Physical Education, and the Graduate Division of Shanghai University of Physical Education, Martial Arts Department of Guangzhou Institute of Physical Education, Scientific Research Office of Shanghai Institute of Physical Education, Sports Department of Wuhan University of Technology, and Sports College of Hebei Normal University, respectively.

Figure 7 shows the network density of individuals in the overall network of Chinese martial arts Research Journal from 2005 to 2020. Size is the number of cooperation units that do not include the node itself; ties indicate the number of cooperation units that do not include nodes and are directly connected to the node; and pairs indicates the theoretical number of organizational collaborations that this node should achieve. It can see in Figure 7 that, in theory, the number of institutions theoretically linked by most nodes is much larger than the actual number. This shows that the degree of the cooperation network connection between scientific research units of Chinese martial arts research journals is not high, and there is a large space for cooperation between scientific research units.

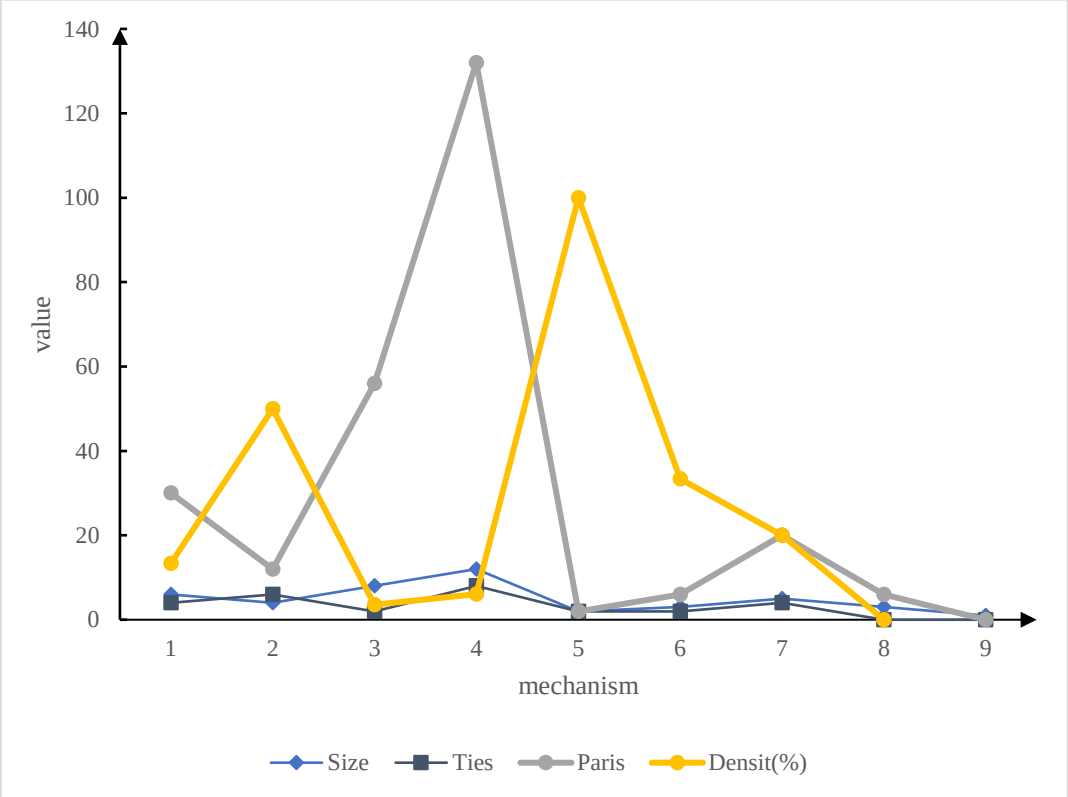


Figure 7. Individual network density values in the overall network of institutional cooperation.

3.3.1. Analysis of author cooperation in the study of martial arts routines

According to statistics, the proportion of cooperation is more than half, which shows that the author has a high level of cooperation in this area. Table 5 specifically shows the author's collaborative research.

Table 5. Cooperative completion of papers by authors.

Number of authors	Number of papers	Proportion
1	421	44.36%
2	317	33.40%
3	157	16.54%
4	29	3.06%
5	16	1.69%
6	6	0.63%
7	1	0.11%
8	2	0.21%

On this basis, one could draw a network diagram of the author collaboration with more than 5 authors. The Bibexcel software is used to draw the co-occurrence matrix of users, and stores it in the form of .net; then, the Pajek software is used to build a network diagram of author collaboration. As shown in Figure 8, the entire core author collaboration network is disconnected.

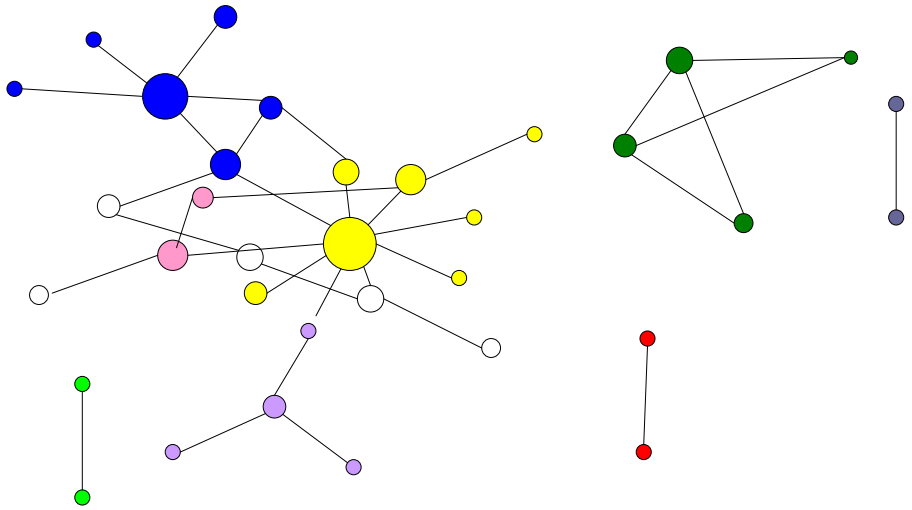


Figure 8. Collaboration map of core authors of Chinese martial arts routine research from 2005 to 2020.

Similarly, this paper uses three indicators to study the author's collaborative relationship, and the results are shown in Table 6. Table 6 shows the top 5 authors with point degree centrality.

Qiu **'s point degree centrality is 11, which means that he has a direct cooperative relationship with 11 core authors. In the entire network, he is the strongest.

The highest degree of intermediary centrality is Qiu **, which means that Qiu ** has more resources in this partnership and plays the role of the "intermediary".

From the perspective of closeness to the centrality, Qiu ** of Shanghai Sports Academy is the

closest to the centrality. This shows that he can easily collaborate with other core writers and plays an important role in a collaborative network with martial arts design researchers.

Table 6. Results of centrality analysis of author cooperation.

Serial number	Author	Point centrality	Between ness centrality	Closeness centrality
1	Qiu**	11	0.414	0.410
2	Yu**	5	0.116	0.243
3	Guo**	4	0.158	0.322
4	king*	4	0.068	0.199
5	Young**	3	0.059	0.176

3.3.2. Hotspots and evolution of martial arts routine research

Table 7 is the keyword list established according to Cite Space software. This article has made some changes. It shows the frequency, emergence rate, centrality of each keyword, and the geometric mean (Sigma) of each index. In the keyword co-occurrence network diagram, the keywords that are used more than 5 times are as follows: traditional martial arts, competitive martial arts, traditional national sports, development, routines, martial arts culture, Olympic Games, competitive martial arts routines, martial arts, etc. This article takes martial arts routines as the starting point. The keywords node martial arts, Chinese martial arts, martial arts routines, etc., do not reflect any practical values in this article, so they are excluded. Competitive martial arts has the highest centrality of 0.41, followed by traditional martial arts, development, and national traditional sports, with centrality values of 0.14, 0.13, and 0.11, respectively.

Table 7. 2005–2020 Chinese martial arts routine research keyword co-occurrence network list sorted by Sigma.

Frequency	Emergence rate	Centrality	Sigma	Key words
128	12.14	0.14	5.05	traditional martial arts
72	3.42	0.41	3.24	competitive martial arts
42	9.56	0.11	2.72	National traditional sports
51	8.05	0.13	2.69	develop
30	4.52	0.05	1.23	martial arts culture
30	4.68	0.04	1.21	culture
6	2.88	0.05	1.15	Olympic Games
13	5.05	0.02	1.12	Competitive martial arts routines
9	4.3	0.01	1.03	martial arts
7	3.2	0.01	1.03	standardization

Geometric mean (Sigma) is a comprehensive index based on frequency, centrality, and emergence rate, which can reflect the importance of nodes in the keyword network. This paper uses a software to create a keyword knowledge map, as shown in Figure 9. In the keyword co-occurrence

network, keywords with a larger geometric mean can reflect some current hot issues. Traditional martial arts, national traditional sports, and martial arts are important subjects in the development of traditional martial arts. Competitive martial arts routines, competitive martial arts routines, Olympic rules, competition rules, etc., are important topics in the current development of competitive martial arts routines. The importance of martial arts culture in Chinese martial arts sports can be seen from the two keywords "martial arts culture" and "culture".

Figure 9. Knowledge graph of hotspots in martial arts routine research.

4. Discussions

The article explores the current research hotspots of martial arts routines from three perspectives: traditional martial arts routines, competitive martial arts routines, and martial arts culture. Martial arts culture has been a popular theme in recent years, and the high frequency of the words "culture" and "martial arts culture" is a good example. From the perspective of either technology or the ontological research of sports projects, there is a gradual shift towards deeper cultural research, reflecting the current era of cultural globalization and the construction of a socialist cultural power. People have a high interest in martial arts culture and a relatively wide research scope. The second part of this paper is the study of traditional martial arts culture, which is analyzed from two aspects: first, the dissemination of martial arts culture, and second, research on regional martial arts culture. Additionally, the connotation and value of martial arts culture, the development of martial arts culture, and the tourism of martial arts culture were analyzed.

5. Conclusions

Accordingly, in order to have a more in-depth and comprehensive understanding of the research situation and development of Chinese martial arts routine theory since the new century, this study used visual knowledge graph software to systematically analyze the research progress of Chinese martial arts routines in combination with literature, logical analysis, and mathematical statistics. According to the current situation, this paper explored the knowledge base of its development and analyzed the cooperation between mainstream academic institutions and authors. This can reveal the hot spots and frontier problems of martial arts routine research, so as to promote the healthy

development of martial arts routine in the future. However, there are some shortcomings in this article. First, TransF has a greater number of parameters than TransE. Due to its complexity, TransF is difficult to apply to large-scale knowledge graphs. Second, both TransE and TransF models assume that the entities and relationships are vectors in the semantic space, so similar entities can be very close in the same entity space and sometimes difficult to distinguish. In the future, this we will continue to search for a more effective method to express knowledge, which may be a challenging research topic.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Conflicts of interest

Dr. Muhammad Bilal Khan is the Guest Editor of special issue "Decision Theory Analysis and Fuzzy Sets Extensions" for AIMS Mathematics. Dr. Muhammad Bilal Khan was not involved in the editorial review and the decision to publish this article.

All authors declare no conflicts of interest in this paper.

References

1. S. F. Li, F. G. Li, Evolvement and prospect of martial arts in the developing course of China, *China Sport Sci.*, **33** (2013) 84–88.
2. Y. J. Han, H. B. Wang, H.M. Zhang, G. L. Cai, A test study on upper limbs' anaerobic metabolism ability of male martial arts routine players, *Liaoning Sport Sci. Technol.*, **41** (2019), 37–41.
3. N. N. Sun, R. Y. Sun, S. H. Li, X. L. Wu, Martial arts routine difficulty action technology VR image target real-time extraction simulation, *IEEE Access*, **8** (2020), 155811–155818. https://doi.org/10.1109/ACCESS.2020.3014459
4. H. P. Gong, Influences of martial arts routine training on dynamic and static balance ability of university students, *J. Guangzhou Sport Univ.*, **41** (2021), 75–78.
5. C. Wehner, C. Blank, M. Arvandi, C. Wehner, W. Schobersberger, Effect of Tai Chi on muscle strength, physical endurance, postural balance and flexibility: A systematic review and meta-analysis, *BMJ Open Sport Exerc.*, **7** (2021). https://doi.org/10.1136/bmjsem-2020-000817
6. W. Song, Y. P. Liu, L. N. Wang, Correlation analysis on competition results of high level chinese martial arts routine athletes, *J. Beijing Sport Univ.*, **38** (2015), 135–140.
7. S. Z. Zheng, An analysis of the scientific knowledge map of the study of sports power in China, *Sichuan Sports Sci.*, **40** (2021), 12–17.
8. Q. Wang, Z. G. Hu, Visualization analysis for the research fronts of the international Olympic movement, *J. Xi'an Phys. Educ. Univ.*, **28** (2011), 433–436.
9. Z. Wang, Q. Liu, Knowledge map analysis in genetic research based on cite space, *J. Southwest China Norm. Univ. (Natural Science Edition)*, **43** (2018), 143–150.
10. N. Y. Ji, Y. Gao, Y. B. Zhao, D. G. Yu, S. W. Chu, Knowledge graph assisted basketball sport news visualization, *J. Comput.-Aided Des. Comput. Graph.*, **33** (2021), 837–846. https://doi.org/10.3724/SP.J.1089.2021.18590

11. W. S. Chiu, T. C. M. Fan, S. B. Nam, P. H. Sun, Knowledge mapping and sustainable development of esports research: A bibliometric and visualized analysis, *Sustainability*, **13** (2021). https://doi.org/10.3390/su131810354
12. Z. L. Zhang, Z. F. Li, H. Liu, N. N. Xiong, Multi-scale dynamic convolutional network for knowledge graph embedding, *IEEE T. Knowl. Data Eng.*, **34** (2022), 2335–2347. https://doi.org/10.1109/TKDE.2020.3005952
13. Z. Li, Q. Zhang, F. Zhu, D. Li, C. Zheng, Y. Zhang, Knowledge graph representation learning with simplifying hierarchical feature propagation, *Inform. Process. Manag.*, **60** (2023). https://doi.org/10.1016/j.ipm.2023.103348
14. Z. F. Li, Y. Zhao, Y. Zhang, Z. L. Zhang, Multi-relational graph attention networks for knowledge graph completion, *Knowl.-Based Syst.*, **25** (2022).
15. Z. X. Fang, S. Y. Wang, S. H. Li, Reconsideration of Chinese martial arts going global, *J. Cap. Univ. Phys. Educ. Sports*, **33** (2021), 564–569.
16. X. D. Wang, Challenges, opportunities and countermeasures of the international communication for Chinese martial arts culture under the background of ant globalization, *J. Phys. Educ.*, **29** (2022), 20–24.
17. R. Kleminski, P. Kazienko, T. Kajdanowicz, Analysis of direct citation, co-citation and bibliographic coupling in scientific topic identification, *J. Inform. Sci.*, **48** (2020), 349–373. https://doi.org/10.1177/0165551520962775
18. G. H. Feng, Y. X. Kong, Subject hotspot research based on word frequency analysis of time-weighted keywords, *J. China Soc. Sci. Tech. Inform.*, **39** (2020), 100–110.
19. Z. G. Ma, R. Y. Ni, K. H. Yu, Recent advances, key techniques and future challenges of knowledge graph, *Chinese J. Eng.*, **42** (2020), 1254–1266.
20. W. J. Chen, Y. Wen, X. Zhang, N. Zhang, S. Zhao, An improved TransE-based method for knowledge graph representation, *Comput. Eng.*, **46** (2020), 63–69.
21. J. C. Figueroa-Garcia, M. A. Melgarejo-Rey, G. Hernandez-Perez, Representation of the Minkowski metric as a fuzzy set, *Optim. Lett.*, **14** (2020), 395–408. https://doi.org/10.1007/s11590-018-1290-6
22. F. X. Wang, (U, V)-interval-valued fuzzy subsemigroups of semigroups, *J. Chongqing Univ. Technol. (Natural Science)*, **3** (2019), 225–230.
23. L. Mao, M. Xie, Image semantic segmentation based on higher-order CRF model, *Appl. Res. Comput.*, **30** (2013), 3514–3517.
24. Y. M. Li, X. Liang, Survey on canonical correlation analysis, *J. China Univ. Metrol.*, **28** (2017), 113–118.

AIMS Press