Analysis of Knowledge Structure on Trade and Environment Issue

SHen Jun[1, a], Jiang Huan[2,b], GU Zu-sha[3,c]

[1, 2, 3] Business School of Shandong University（Weihai）,Weihai China,264209

[a]sdshj5018@sdu.edu.cn, [b]jianghuan09@163.com, [c]gzs@sdu.edu.cn

Keywords: Trade environment effect; Knowledge structure; Co-citation analysis; Strategic diagram.

Abstract. This paper, studying on the basis of existing research, taking the issues of trade and environment as the research object, applying the analytical models of document co-cited, combined with Cluster analysis and Strategic diagram analysis, made empirical analysis on the related literature from CSSCI database, intuitively showing the research status as well as research hotspots and knowledge structure in this field. The research suggests that the subjects such as the Coordination of Trade Policy and Environmental Policy and so on, are in the core position among the whole research field.

INTRODUCTION

The relationship between trade and environment has gradually attracted people's attention since the United Nations Conference on the Human and Environment held in Stockholm in 1972. But in the 1980s, environmental problems were temporarily ignored because of the world economic recession. Since the beginning of the 1990s, the world economic began to recover, it has once again become the focus of attention. People adhering to the theory of free trade and environmental protection engaged in a fierce debate on the issue of trade and environment. Those who insist the theory of free trade believe that trade liberalization is beneficial to the improvement of the environment. They think that the root cause of environmental degradation is the market and government failure, not trade liberalization [1] [2]. However, environmentalists argue that trade growth runs counter to environmental protection objectives. They believe that free trade is one of the important factors of destruction in the environment. Unrestricted trade will cause the destruction of the ecological environment, especially in countries with loose environmental policies, trade liberalization is far more damaging to the environment [3][4].This paper, on the purpose of intuitively showing the domestic research status, hotspots and knowledge structure on the trade and environment problems, taking published papers from Chinese Social Science Citation Index (CSSCI) database as the research object, is going to carry out a bibliometrics and scientometrics study on relevant literature.

DATA SELECTION AND STANDARDIZATION

With the help of the research methods and indexes of bibliometrics, applying the document co-cited analysis, this study intends to intuitively show the research status as well as research hotspots and knowledge structure of trade and environment problems in China. So we select the CSSCI database, set the retrieval time for 1998-2012, and do the retrieve in the advanced search in the search type = "trade * environment |@ALL= carbon emissions * trade" . Retrieval results show that we obtain 1753 literature in total, but it contains conference notice, meeting review, book review and so on. Besides, there are some in the search results, but they have nothing to do with the research subject. Therefore, after removing above literature, this paper finally gained 926 valid data, with download and update time in July, 2013.

THE MAIN MEASURING INDEXES AND RESEARCH METHODS

A. Main indexes

(1) Cosine index: it is a kind of index that measures the co-occurrence strength and this research is mainly to measure the strength of co-occurrence of co-citation literature by using this index. The formula is as follows:

$$Cosine = \frac{F(A,B)}{\sqrt{F(A)F(B)}} \qquad (1)$$

Where F(A) is the frequency of the cited literature A appeared in a given collection of literature; F(B) is the frequency of the cited literature B appeared in a given collection of literature; F(A,B) is the frequency of appearance of both cited literature A and cited literature B. The value range is between 0 and 1, the greater the value, the higher co-occurring strength between the cited literature.

(2) Density: it is used to measure the close degree of cited literature in each internal cluster and reflects the cluster's ability to maintain and develop themselves [5]. Density calculation has a variety of measurement methods, such as mean, median, square and so on. In this study, we use the average cosine index between cited literature in clusters for measurement.

(3) Centrality: it is used to measure the close degree of the cited literature in each cluster with other cluster. It reflects the degree of mutual influence between one research subject and the others. If the number and intensity of connection between one research subject and others are larger, then this research subject is the more central position in the field of the whole study [6].The more the number of edges of a cluster connected to other clusters, the greater the centrality of this cluster compared with other clusters.

B. Main research methods

(1) Citation Analysis:Citation analysis is an important part of bibliometric and scientometrics. By using a variety of mathematics and statistics method and logical methods, such as comparison, induction, abstract, generalization and so on, it analyzes and studies the quoted or cited phenomenon of scientific journals, papers, authors and other analysis objects in order to reveal quantitative characteristics and inherent laws, and achieve the purpose of evaluating and predicting scientific development trend. To some extent, literature citation reflects the degree of influence and quality level. Co-citation analysis of literature refers to the Co-citation relationships between two papers when they are cited together by later one or more papers, some scholars also call it relations with cited. Co-citation analysis also includes co-citation analysis of journals, authors, subjects and so on.

(2) Cluster analysis:Cluster analysis is an active method of bibliometric and visualization in data mining. It analyzes co-occurrence strength among keywords, and then assigns the keywords which have relatively greater strength into various types of clusters. Since traditional clustering algorithm is to cluster words with shortest distance, this method has several drawbacks: having no central concept, the clustered words may not likely to express the same content, etc. This paper, on basis of clustering algorithms, borrowing from Callon's method of constituting sub-sets, i.e. having no more than 10 blocks of knowledge in a constituted sub-set, takes a pair of technical terms which have the highest index of cosine in co-occurrence matrix (the greatest co-occurrence strength) as blocks of knowledge, in order to reflect the research content and orientation of this cluster.

(3)Strategic diagram analysis:Strategic diagram, drawn by the deviation from mean of density and centrality of each cluster, is a two-dimensional planar graph to show the position of research subject. Among them, the horizontal axis represents Centrality, the longitudinal axis represents density, and origin of coordinates is determined by the mean of Centrality and density of all clusters. We separate the plane into four quadrants(Fig. 1):

The cluster in Quadrant I has higher density and higher centrality. Higher Density indicates that the internal of cluster has strong internal connection and shows that the researches represented by these clusters receive comparatively higher attention; Higher Centrality indicates that the cluster with

other clusters have stronger connectivity, and the researches represented by these clusters are the core of whole research domain. So clusters in Quadrant I are known as the core research subject.

The cluster in Quadrant II has higher density and lower centrality. Higher Density demonstrates that the cluster has been well studied and was once the center area of research subjects. But lower Centrality shows that it is no longer the center and has been gradually marginalized and peripheral and we have less and less interest in the research. So clusters in Quadrant II belong to a secondary research subject.

The cluster in Quadrant III has lower density and lower centrality which illustrates that the researches represented by the clusters are the marginal zone in the whole research network. Therefore, it is called marginal research subject. This kind of cluster has two trends of development: one possibility is to obtain adequate research into the second quadrant, and then enter Quadrant I as a core subject, or contact with other subjects more and more closely and enter Quadrant IV, then enter the first quadrant as the core research subject. Another possibility is that the researches represented by the clusters receive less attention, and have less tightly contacts with other research subjects, finally the researches disappear from the whole research network.

The cluster in Quadrant IV has higher density and lower centrality. Higher Density shows that although the cluster has higher connectivity than other clusters, lower density shows its internal connectivity is not very strong and has Low degree of polymerization. Therefore the researches represented by clusters in Quadrant IV are called fundamental research subjects. For the researches represented by these clusters, we should consider its strategic position. They may be the object of study in other fields, or may be emerging and will become a hot field in the future research center. But they have not yet become significant objects of study, and they are becoming important and mature. As long as these research subjects obtain adequate research, they will enter into Quadrant I, thus become the core of the whole research area.

EMPIRICAL ANALYSIS AND DISCUSSION

A. Co-citation matrix

CiteSpace is visual analysis software based on JAVA platform and was developed by Dr. Chen Chaomei from College of information science and technology in Drexel University in the United States. This study used CiteSpace to generate a co-cited matrix[7], did the main settings in the set interface of CiteSpace, such as time interval was set to one year, co-cited literature, the threshold value was (3,3,0; 4,4,0; 3,3,0), etc. We ran the CiteSpace software, got 182 highly cited literature and generated a matrix of cited literature. The size of the co-occurrence strength of cited literature in a matrix is measured by value of cosine index: the greater the strength, the higher the value. According to co-occurrence relationship and cosine index among cited literature, a 182×182 co-cited matrix is constituted.

B. Cluster partition and Strategic diagram analysis

(1)Classification principles of Cluster

1)In the co-occurrence matrix （182×182） which is generated by CiteSpace, the pair of cited literature with the highest cosine index are taken as the initial literature of the first cluster. 2) Sequence cosine index of the 182 cited literature and one of the pair of cited literature in the matrix descending order, then select 10 cited literature (including the pair of cited literature as the initial literature) from higher to lower from the sequence (if there are less than 10 literature whose cosine index is more than 0, only select the literature whose cosine index is more than 0). If there are more than 10 cited literature whose cosine index is more than 0, choose only 10 cited literature to join in the cluster, for that cluster has reached saturation value (10 cited literature).3) When the first cluster has formed (or saturated, or there are less than 10 cited literature whose cosine index is more than 0), delete the literature which joined the cluster from the matrix (deletion should include both the row and the column), so as to ensure that these literature would not join in the following other clusters. 4) Go over repeatedly step 1 to 3,then clusters could be formed one by one, until all cited literature with

co-occurrence relationship have joined in clusters. If there are still some cited literature with no co-occurrence relationship, that is to say that the co-occurrence strength among all cited literature is 0 (cosine index=0), the clustering generation has been finished, the cited literature left would not ever joined in any cluster.

Among the 182 high-frequency cited literature, the ones concerning "Is Free Trade Good for the Environment?" enjoy the highest frequency (Freq=669), but it is still unable to identify research content and direction only with help of individual high-frequency cited literature. Therefore, this paper uses Callon's cluster analysis for reference to identify research content and direction. According to the above clustering algorithms and principles, 182 cited literature are distributed into 17 clusters, each cluster has its name summarized on account of the cited literature included, the 17 names are just the research contents and direction in that field.

From the 17 cluster names and constitution, we can be seen that research hot spots in the field of trade and environment in our country present the characteristics of diversification and the vast majority of cited literature was published after entering the 1990s. This is because, since the 1990s, economic globalization and trade liberalization wave engulfing the world have greatly promoted the commodity and capital flow internationally. The rapid growth of the international trade has improved the national welfare of every trading nation and stimulated economic development in each county. However, at the same time, the global ecological environment is progressively deteriorating: climate warming, energy crisis and environment pollution and so on, are frequently warning human and challenging the sustainable development in today's world economy. Therefore, the issues of trade and environment have caused the wide attention of academic community in our country, and scholars have carried out studies on this issue from different perspectives.

(2)Strategic diagram analysis

The Centrality is taken as lateral axis, and the density is taken as longitudinal axis. According to meaning of these two indicators and position of quadrant, we can clearly see that research hot spots in the field of trade and environment present the characteristics of diversification, but the positions of each research contents in the whole research field are very different.

Subjects with higher density and higher Centrality include the coordination of Trade policy and environmental policy, Pollution haven hypothesis and so on. Since the 1990s, these two subjects have aroused wide attention of scholars in our country. They have carried out long-term and in-depth researches, made them become research focus in trade and environment field and at the core of the entire research field.

Subjects with higher Centrality but lower density mainly include "The Environmental Kuznets curve ", "Environmental effect of trade liberalization ", " The environment standard and international competitiveness " ," Trade, economic growth and environmental pollution ". These research subjects are closely related with other subjects in the whole field of trade and environment, the researches to themselves are not sufficient. Especially "The environmental Kuznets curve" has the highest Centrality. Since the Environmental Kuznets curve hypothesis was put forward, many scholars have been inspired on the research of the relationship between economic growth and environmental quality. Chinese scholars have also used the Environmental Kuznets curve as analysis tool. But so far, the EKC theory analysis is still lack of unified theoretical framework, and the empirical analysis results also have very big difference.

Located in Quadrant II are fields with lower Centrality but higher density including "Trade and sustainable development", "Environmental regulation and comparative advantage", "Economic growth and environmental protection", "Green barrier", the theme was the core problem of trade and environment research field, but in recent years, scholars have less interest in this research.

Located in Quadrant III are fields with lower Centrality and lower density including "The relationship between trade and environment", "The environmental impact of foreign direct investment", "Legal coordination of trade and environment", "Carbon emission effect of trade", "Energy cost of foreign trade" , "Environmental policy and trade patterns", "The issues of trade and environment in WTO". These subjects are at the marginal zone in the research network, and have

not caused the attention of researchers. Many clusters gathered in Quadrant III indicate that research hot spots in the field of trade and environment in our country present the characteristics of diversification, but many subjects are in the periphery of the whole research network, and haven't got sufficient research, so they need researchers more attention.
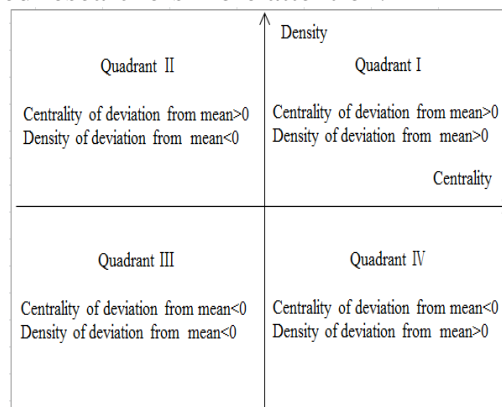

Fig. 1 Strategic diagram

CONCLUSION AND PROSPECT

This paper, through measurement indicators of bibliometrics and scientometrics, adopts co-cited analysis, cluster analysis and strategic diagram to intuitively illustrate research status, hotspots and knowledge structure of the issues of trade and environment. Through the comprehensive analysis of sub cluster and large cluster, this paper summarizes the researches in this field in China are mainly concentrated in the four major themes: the related research literature have been divided into four categories: theory basis of trade and environment relation, the effects of trade liberalization on the environment, the effects of environmental regulation on trade, and the coordinated development of foreign trade and environment protection. This study is an attempt that applies scientometrics and bibliometrics in the field of the issues of trade and environment. However, the non-standard data brought a lot of obstacles to empirical research, and also may affect the empirical results in a certain extent. How to make full use of literature database, and how to normalize the literature data of the current database, will be an important research subject in the future.

REFERENCES

[1]Anderson K, "Agricultural trade liberalization and the environment: a global Perspective"[J].The World Economy ,1992，15(1): 153- 171.
[2] Stevens R, "Environmental regulation and international competitiveness"[J]. Yale Law Journal.1993, 102(8): 2039-2139.
[3]Daly H, "The Perils of Free Trade" [J].Scientific American, 1993, (269):24-29.
[4]Jenkins R, "Environmental regulation and international competitiveness: a review of literature and some European evidence" [J].Discussion Paper No.1, United Nations University, Institute for New Technologies, 1998.
[5] Callon M, Courtial J P, Laville F. Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemistry [J]. Scientometrics, 1991, 22(1): 155-205.
[6] Bauin S, Michelet B, Schweighoffer M G, et al. Using bibliometrics in strategic analysis: "understanding chemical reactions" at the CNRS[J]. Scientometrics, 1991, 22(1): 113-137.
[7] Chen C. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature [J].Journal of the American Society for Information Science and Technology, 2006, 57(3): 359-377.