Neoteric Developments in Management Science in Engineering

Neoteric Developments in Management Science in Engineering:

Perspectives from Scientific Journals

By

Jiuping Xu

Cambridge Scholars Publishing

Neoteric Developments in Management Science in Engineering:
Perspectives from Scientific Journals

By Jiuping Xu

This book first published 2023

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-5275-0474-3
ISBN (13): 978-1-5275-0474-5

Contents

Chapter 1
Introduction

Last year, we published the scientific journal report, titled “Scientific Progress in Management Science in Engineering”, and presented it at the opening ceremony of the online 16th International Conference on Management Science and Engineering Management in Turkey. Participants gave a positive evaluation of the report at the conference. They believed that the report could provide an important reference and assistance for editors to improve the influence of their journals by evaluating them based on the impact factor and comprehensive index. These two aspects highlight the main differences between the journal report and other journal rankings, and affirmed the contribution of the report to: (1) determine the relationship dynamics between the seven management science engineering (MSE) journal categories and the frontier journals and emerging journal categories; (2) introduce five-year journal performance indicators as evaluation indicators; (3) define and classify the MSE journal categories for the first time; and (4) analyze and compare the contributing performance of different regions and countries for journals over the past three years for the first time. Encouraged by those positive responses, we continued to devote ourselves to the MSE journal report to develop and publish the 2021 edition.

Originating from the 2018 edition, the purpose of this report is to identify the main MSE research categories and identify the most representative journals in each category.

In this report, a comprehensive approach is proposed, including literature mining, cluster analysis, and expert systems. The data of the proposed method mainly come from Web of Science (WoS) and Journal Citation Reports (JCR), and the results identify the seven categories most relevant to MSE: civil engineering, engineering management, industrial engineering, energy engineering, environmental engineering, information engineering and agricultural engineering. Using Note-Express and CiteSpace software to identify the most relevant journals in each category from the literature mining, it was found that there are 137, 67, 49, 114, 54, 86 and 14 journals in each category, respectively. Then, all the identified journals in each category were divided into four zones, labeled A to D, where zone A comprises the journals that have had the greatest impact on the development of the topic; zone B comprises the journals with high-quality papers, high submission rates, and low acceptance rates; zone C comprises recognized journals but with limited impact; and zone D comprises the journals with a low citation rate and of low importance. The total data of journals in A, B, C and D zones in recent three years has been analyzed, and the main components of journal evaluation indicators are discussed. Moreover, contributions by countries and regions to zone A journals are recognized. The main advantage of the report is that it can adjust the journals' influence and citation differences in various disciplines over a long time, and provide a reference for the future development direction of the journals.

This report was developed with the joint efforts of members of the scientific committee, many of whom are editors of the most relevant journals, academics, researchers from different countries, and members of professional associations. I would like to sincerely thank them for their contributions in writing the report. Compared with the previous version released in 2020, in this version some journal categories have been replaced according to the degree of topic concentration, and other journal categories have been added. In addition, it is worth noting that the newly added three-year zoning data comparison and contributions of zone A journals further provide reference for the future development and endeavors of journals in different regions. Since this collection of journals usually conforms to JCR's SCI journals, those that are excluded from WoS due to paper quality are not included in this report.

Chapter 2
Support Committee

Support for this proposed MSE journal report was received from both the expert committee and the organizational committee. In the expert support group, the expert-in-chief, professional expert board, and working group all worked together. The expert-in-chief was Prof. Jiuping Xu, an outstanding MSE researcher, the professional expert board was made up of fellows from different national academies of sciences and editors from international journals, and the main members of the working group were PhD candidates with MSE majors. An introduction to the expert support is given in the following sections.

2.1 Expert Committee

The expert support for this journal report came from the expert-in-chief and the professional expert board.

2.1.1 Expert-in-chief

Jiuping Xu
Assistant Vice-President of Sichuan University and Dean of the Business School;

Most cited Chinese researcher in the field of Decision Science (2018-2019) and Management Science (2021), Elsevier;
Honorary Academician, Academy of Sciences of Moldova, 2016; Academician, Mongolian National Academy of Sciences, 2016;
Lifetime Academician, International Academy for Systems and Cybernetic Sciences (IASCYS), 2010;
President, International Society for Management Science and Engineering Management (ISMSEM), 2007 to the present.

2.1.2 Professional Expert Board

Benjamin Lev
Editor-in-Chief of Omega, the International Journal of Management Science;
Professor, LeBow College of Business, Drexel University, Philadelphia PA, USA.

Asaf Hajiyev
Secretary General of Parliamentary Assembly of the Black Sea Cooperation (PABSEC);
Academician at Azerbaijan National Academy of Sciences.

Gheorghe Duca
President, academician, the Academy of Sciences of Moldova;
Chemist, founder of the Research School on Ecological Chemistry.
Editor-in-chief, Chemistry Journal of Moldova.

Mitsuo Gen
Tokyo University of Science, Research Institution for Science and Technology Fuzzy Logic Systems Institute, Japan.

Fausto Pedro García Márquez
Full Professor at the University of Castilla-La Mancha (UCLM), Spain;

Honorary Senior Research Fellow, Birmingham University, United Kingdom.

Fulya Altiparmak
Professor of Industrial Engineering, Faculty of Engineering, Gazi University, Turkey.

Mohamed Hag Ali Hassan
President of The World Academy of Sciences, Professor and Dean of the School of Mathematical Sciences, University of Khartoum, Khartoum.

2.1.3 Working Group

Ziqiang Zeng
Researcher at Sichuan University, Business School, China.
Member of the American Institute of Electrical and Electronics Engineers (IEEE).
Associate Member of the American Society of Civil Engineers (ASCE).

Zongmin Li
Professor, Sichuan University, China.
Managing editor of International Journal of Management Science and Engineering Management (IJMSEM).

Yi Lu
Professor, Sichuan University, China.
Guest editor of Environmental Hazards.

Liming Yao
Professor, Sichuan University, Business School, China.
Member of the Institute for Operations Research and Management Sciences (INFORMS).

Zhibin Wu
Professor, doctoral tutor at Sichuan University, China.
Visiting Scholar at the Department of Industrial and Systems Engineering, University of Washington.

Chengwei Lv
Assistant Professor, Sichuan University, China.
Visiting scholar at Joseph M. Katz Graduate School of Business, University of Pittsburgh.

Other members of the working group were: Hongyan Tao, Fengjuan Wang, Yidan Huang, Chuandang Zhao, Kejing Shu, Keru Fan, Yalou Tian, Liqing Yao, Zongze Wu and others.

2.2 Organizing Committee

The International Academy for Systems and Cybernetic Sciences (IAS-CYS);
International Society of Management Science and Engineering Management (ISMSEM);
Sichuan University.

Chapter 3
General Methodology

The academic journal analysis is one of the most valuable vehicles for promoting academic interaction. Moreover, exploring the co-citation relationships between academic journals can display the relationships between disciplinary subject development. Systematic theoretical analysis, which typically encompasses concepts such as paradigms, theoretical models, phases, and quantitative or qualitative techniques, is vital to illustrate the methods and principles associated with a certain branch of knowledge and decide the co-citation relationships. Therefore, to profoundly assess journals, a systematic but general methodological classification mean is applied to identify the journal rankings and categories. Therefore, in this report, a specific literature analysis approach was taken in which: (a) the Web of Science database was employed for the literature analysis; and (b) CiteSpace was used to visualize the scientific literature patterns and trends. Compared to purely qualitative methods, this classification method was regarded to be fairer and more objective and could more accurately reflect the actual situation.

3.1 Data Acquisition

The data adopted for identifying a category and recognizing a journal mainly come from the Web of Science (WoS) database, which is an online subscription based scientific citation indexing service that was initially created by the Institute for Scientific Information (ISI) and has been preserved by Clarivate Analytics from December, 2017 [1]. The database links multiple databases that reference cross-disciplinary research, and therefore supports a complete exploration of the specific subfields within an academic or scientific discipline [4]. The WoS platform is comprised of specialized subject indices, regional citation indices, patent data, and research data set indices, which comprise over 21,100 high-quality global peer-reviewed academic journals (including open-access journals) covering 250 disciplines since 1900 [2, 3].

To promote the accuracy of the analysis, the WoS Core Collection Indexes have been applied for the data acquisition. The Web of Science Core Collection-Citation index has six main online databases: the Science Citation Index Expanded (SCI-Expanded); the Social Sciences Citation Index (SSCI); the Arts & Humanities Citation Index (A&HCI) [8]; the Emerging Sources Citation Index (ESCI); Conference Proceedings Citation Index Science (CPCI-S); and Conference Proceedings Citation Index Social Science & Humanities (CPCI-SSH) [2, 3], as shown in Fig. 3.1.

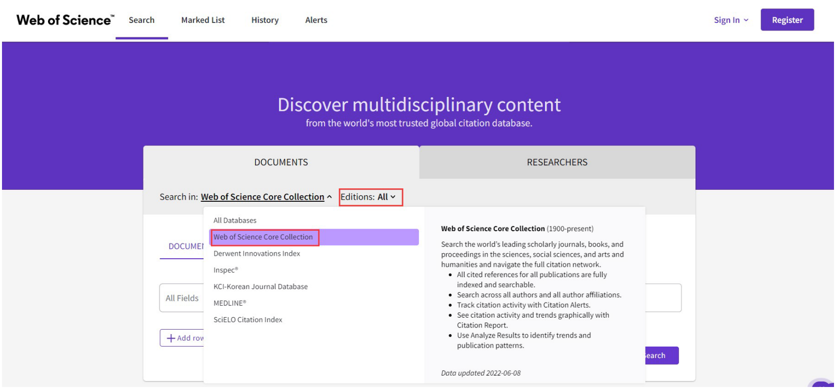

Fig. 3.1: Web of Science Core Collection Database

Advanced search

The advanced search of WoS was discovered to be a significant improvement on the more traditional finite state algorithm. It was selected to include a finite state pattern-matching algorithm in the search program, which recognizes all titles, and citation indexes that satisfy Boolean keywords. An advanced search query consists of one or more field identifiers and search strings, allowing the use of Booleans and wildcards. The search for records involves using field identifiers, set combinations, or a combination of the two. Common abbreviations in WOS are: TS=Topic, TI=Title, AU=Author, AI=Author Identifiers, GP=Group Author, ED=Editor, SO=Publication Name, ZP=Zip/Postal Code, FO=Funding Agency, FG=Grant Number, FT=Funding Text, SU=Research Area, WC=Web of Science Category, DO=DOI, PY=Year Published, CF=Conference, AD=Address, OG=Organization-Enhanced, OO=Organization, SG=Suborganization, SA=Street Address, CI=City, PS=Province/State, CU=Country/Region, IS=ISSN/ISBN, UT=Accession Number, PMID=PubMed ID, and ALL=All Fields.

The Boolean Operators in the Web of Science Advanced Search are: **AND**, to find records with all terms; **OR**, to find records with any of the terms; **NOT**, to exclude records with certain words; **Near**, to find records with all terms within a certain number (n) of each other; and **SAME**, to search terms that must exist in the same sentence. To control the retrieval of complex sums with different spellings, Truncation and wildcard characters are used; Item *****, to retrieve words with a variation from zero to many characters, Item **?**, to retrieve words with the replacement of 1 character; Item **$**, to retrieve zero or one character, and Item **“”**, to search for exact phrases.

To attain the search results, the following processes for the “advanced search” were followed: (1) Click on the Web of Science website, choose the Web of Science Core Collection database and enter the Advanced Search item; (2) Set the search command as “TS=(management science)”, and restrict results by setting the language as “English” and document types as “Article OR Proceeding Paper OR Review”. Other items are set by default, including “Timespan=All years (1900-2021)”, and “Web of Science Core Collection: Citation Indexes=SCI-Expanded, SSCI, A&HCI, CPCI-S, CPCI-SSH, ESCI”;
(3) Click the “Search” item and find the search results at the bottom of the page in “Search History”, shown in Fig. 3.2.

A search history is a search query or multiple search queries saved to a server or hard drive. There is no limit to the number of search histories

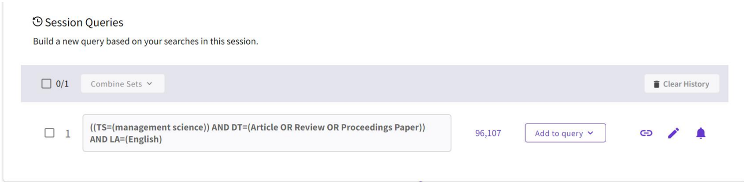

Fig. 3.2: Search history in the Web of Science Core Collection Database

that can be saved, and combined search sets can be created from the Search and Advanced Search functions and are listed in the Search History table in reverse numerical order, with the most recently created set at the top. The operational steps for this are as follows: (1) Click the AND or the OR option; (2) Select the check box for each set that is to be combined; (3) Click the Combine button; and (4) Click the link in the Results column to view the results of your search.

Data Export

In the Web of Science, the search results are displayed with the full title, author names, and source as a list of 10, 25, or 50 items per page. When the full text is available, the option to "view free full text" appears. Related records can be found and sorted by the latest date, times cited relevance, first author, publication year, and source title. The results can also be analyzed (i.e., by author, country/territory, or document type), and a citation report presented with a labeled bar chart. The results can also be further refined, and the records viewed or excluded.

At the bottom of the Results pages, output options are available (as shown in Fig. 3.4), including Save to Endnote Online, Save to EndNote Desktop, Save to Email, Save to my Publons profiles, and Save to Other reference software. As CiteSpace software was used for the follow-up analysis, the Save to plain text file was chosen.

This option supports the output selection of all search results (no more than 500 at a time) or specific record(s) such as: (1) Selected records on one page by checking the box for each desired record; (2) All records on a page by selecting a value in the Show 10, 25, or 50 per page list; and (3) Records,

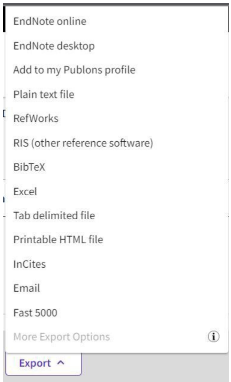

Fig. 3.3: Data Export in the Web of Science Core Collection Database

to select a range of records. In this case, 500 items at a time were selected until all of the records had been exported.

Four Record items are provided: (1) Author, Title, Source; (2) Author, Title, Source, Abstract; (3) Full record; and (4) Full Records and Cited References. The CiteSpace analysis was based on the "Full Records and Cited References", as shown in Fig. 3.4. The following file formats are available for Saving to Other File Formats: Other Reference Software, HTML, Plain Text, Tab-delimited (Win), Tab-delimited (Mac), Tab-delimited (Win, UTF-8), and Tab-delimited (Mac, UTF-8). For this chapter, Plain Text was chosen. Click "send" and the system saves the document as a text file (for example, savedrecs.txt). The selected data include the following information: (1) Bibliographic Fields include the author, title, and source information; (2) Bibliographic plus Abstract includes the bibliographic fields and author abstract; (3) Full Record includes all data on the Full Record page; and (4) Full Record plus Cited Reference includes all data on the Full Record page as well as the cited references.

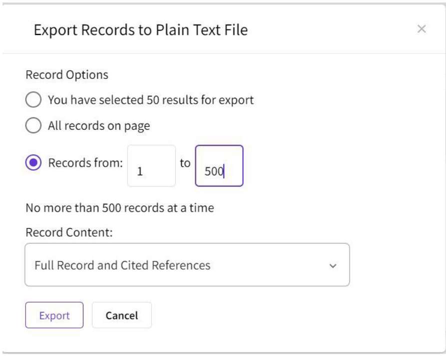

Fig. 3.4: Send to File in the Web of Science Core Collection Database

3.2 Category Identification

A thoroughgoing literature review supports the recent findings to be discussed about a particular research question, as scientific knowledge based on text usually has its evolving life cycle. Therefore, analyzing the evolutionary patterns has been an increasingly valuable text mining research direction in the last few years [9, 10]. With the literature analysis expanding, the literature mining scope should categorize the journals and rank the related journals under consideration. In this edition, full MSE-connected journals were identified and categorized by CiteSpace's cluster.

Developed by Dr. Chen [15], CiteSpace is a Java language information visualization software through citation analysis theory. Inspired by Thomas Kuhn's scientific revolution structure, the CiteSpace design has the central idea that research foci change over time, sometimes incrementally and sometimes drastically. Therefore, scientific developments can be traced by studying the footprints revealed in scholarly publications. With contributions made by contemporary scientific community members, a dynamic, self-organizing knowledge system that embodies consensus, disputes, uncertainties, hypotheses, mysteries, unsolved problems, and unanswered questions has been formed and it broadens the scope that a specific topic is related to other topics.

The basic function of CiteSpace is network investigation and visualization. Network modeling and visualization build up the intellectual landscape of a knowledge domain, figure out the questions researchers have been trying to answer, and identify the methods and tools that have been developed to find solutions. In addition, CiteSpace integrates information visualization methods, bibliometric methods, data mining algorithms and web algorithms, all of which allows research data to be converted into a scientific knowledge map from which knowledge generation and interpretation, key evolutionary research domain paths and knowledge inflection points can be identified, allowing detailed exploration of citations in management science journals [5, 6].

Parameter Setting

Divided into two parts, the user interface consists of the project control and the progress report windows on the left, and several panels that support the process to be configured by various parameters on the right-hand side. Particularly, the CiteSpace process obtains the current project input dataset, builds a network model of bibliographic items, and visualizes trends and patterns extracted from the dataset in the network.

This is primarily, to build a new project so that the relevant research is submitted to CiteSpace. The step of the literature database entry into CiteSpace is shown in Fig. 3.5.

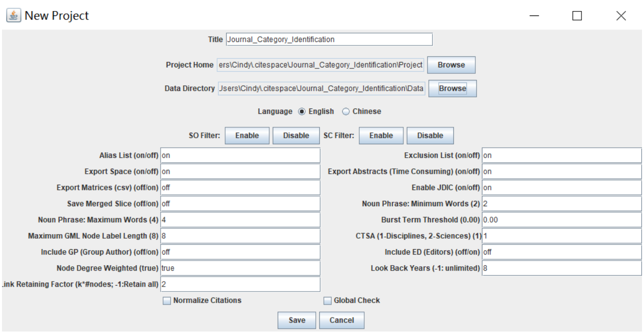

Fig. 3.5: Establishment of the literature database for entry into CiteSpace

Next, setting the basic parameters using the right-hand side panels and selecting the Node Types chosen based on the research object for the cited journals, are shown in Fig. 3.6.

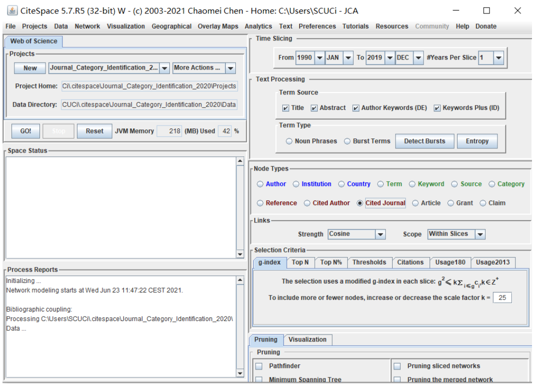

Fig. 3.6: CiteSpace interface for configuring analysis

CiteSpace possesses abundant functions: a collaboration atlas (author, institution and country), a co-occurrence atlas (feature word, keyword, subject category), and a co-citation atlas (literature, author and journal). A cited journal refers to the journals cited in the same document and reflects the relevance of the various journals and disciplines. The knowledge base distribution in a research field is formed from a cited journal analysis, thereby further revealing the classifications for the cited periodicals.

Third, click on the **GO** button to start the process. CiteSpace reads the data files in the current project (Demo) and reports the progress in the two windows on the left-hand side of the user interface. Once the modeling process is complete, three options are given: **Visualize**; **Save As GraphML**; or **Cancel**. When Visualize is selected, it moves to the visualization window for further interactive exploration.

Visualization results

After pressing the Visualize button, the Visualization Window appears, which primarily has movements with a black background. Once the movements have settled, the background color changes to white. Here, the initial visualization is examined and the additional functions explained. First, CiteSpace shows a merged network visualization of the developments in the field based on networks that correspond to consecutive years from 1990 through to 2021 and shows the most vital footprints for the related research activities. Each dot represents a node or a cited journal in the network. Also, CiteSpace generates networks for other types of entities. The lines that connect the nodes are the co-citation links. Similarly, CiteSpace can generate networks for other types of links. The line colors indicate when a connection was made for the first time.

Fig. 3.7 is the original graph, in which each node represents a different journal and the lines indicate the connections between the journals. The node size is related to the number of journals, that is, the greater the frequency, the larger is the node. A control panel is shown on the right-hand side of the Visualization Window, in which the node labels are displayed and can be changed using the sliders and a combination of threshold values; the node size can also be changed using the node size slider.

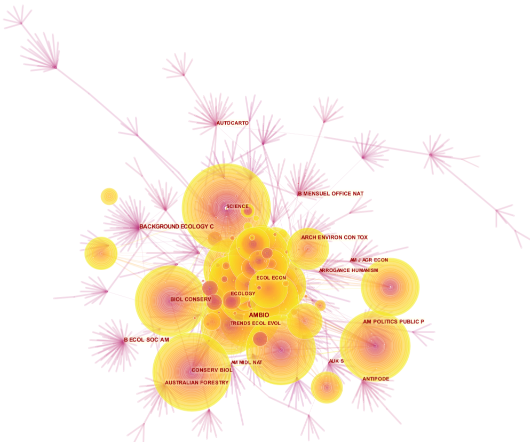

Fig. 3.7: Visualization window: Journal cooperation and affiliations

Clustering

Although it is possible to identify some prominent groupings by studying the visualized network, the clustering function in CiteSpace provides greater precision when seeking to identify the groupings or clusters. To start the clustering function, the Find cluster icon is simply clicked. To characterize the nature of an identified cluster, CiteSpace extracts noun phrases from the titles (T in the following icon), keyword lists (K) or abstracts (A) of articles that cite a particular cluster; here, T was chosen. Once the process is finished, the chosen labels are displayed. By default, labels based on one of the three selection algorithms are shown; that is, TF*IDF. Cluster labels are displayed once the process is completed, and the "# Clusters" are shown in the upper right corner of the canvas. The clusters are numbered in descending order of cluster size, starting from the largest cluster # 0, the second-largest # 1, and so on.

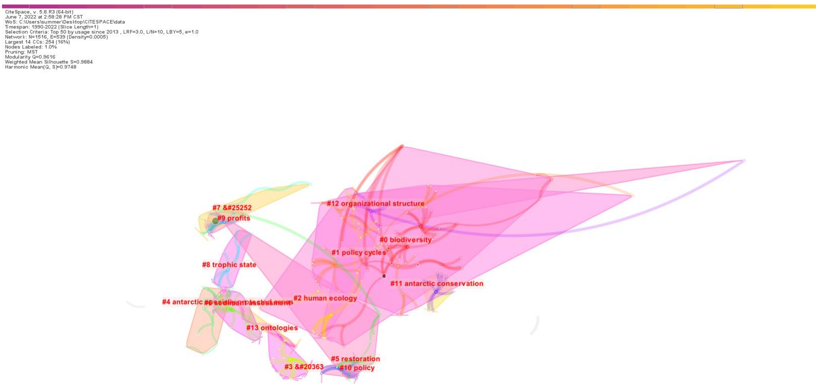

Fig. 3.8: Members of different clusters are shown in different colors

Fig. 3.8 shows the clustering from an analysis of relevant literature from a cited management science journal, with each cluster corresponding to an underlying theme, a topic, or a line of research. Kuhn's paradigm involves time period clustering, with the cluster colors indicating the average years

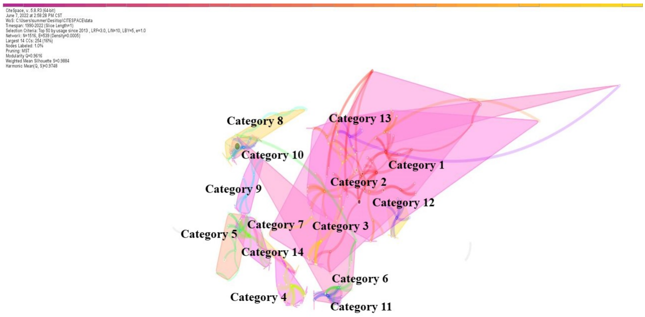

Fig. 3.9: Cluster view for category identification (C-A)

for that cluster [12]. From this visualization, a deeper understanding of the connections from one cluster to another cluster can be gained.

There are two important metrics for describing the overall structural properties of the network and modularity and the mean silhouette scores. As seen in Fig. 3.8, the higher the modularity values, the more reasonable is the cluster clarified. To decide whether the data are appropriately clustered, the Silhouette measure is applied. The silhouette measure closing to 1, represents the data suitable for clustering, and closing to minus one, suggests the data are not appropriately clustered; however, when the silhouette measure is close to zero, the data are on the border of two natural clusters [13].

There were identified clusters in auto-labeling; however, to determine which of these were the most appropriate, silhouette and size measures were used. *n* clusters with a high silhouette (ID: ID1, ID2, ..., IDn; silhouette: S1, S2, ..., Sn) and a large size (SZ1, SZ2, ..., SZn) were extracted and the cluster labels (TF*IDF) summarized as “category 1”, “category 2”, ..., “category n”. The category identification cluster view is shown in Fig. 3.9.

3.3 Journal Recognition

To identify the related journals in each of the journal categories identified in the last section, a new data analysis system was designed. To determine the number of citations and the relationships between the cited journals, first the data analysis system was used to identify and review thousands of management science research papers, after which literature mining, the process of extracting unknown, comprehensible and available knowledge, was applied to organize the information [7, 14].

The scientific knowledge map analysis had word frequency analysis, co-word analysis, co-citation analysis, and cluster analysis, which together illustrated the span of scientific knowledge and scientific research directions. A knowledge-domain-based knowledge map illustrates the development and structure of scientific knowledge by identifying the complex relationships between the knowledge units or knowledge groups, the network structure, the interactions, the evolution, and the new knowledge. The CiteSpace visualization process has four stages: document retrieval and preservation; data guide and tuning; map customization; and map analysis.

Data recollection and tuning

The analyzed data, which included SCI-EXPANDED, SSCI, A&HCI, and other citation libraries were also extracted from the WoS core collection [11]. An advanced search was then used after which articles, proceedings papers, and reviews were selected. After unnecessary research was filtered out, the related articles were downloaded in Plain Text format, with each data record including author, title, abstract, and research citation, and the retrieval period being from 1990 to 2021. When searching, the search command was set at X_i AND Y_i, for which "X_i" represented "management" and "Y_i" represented "category n". When the search was complete, all related articles were downloaded and saved in CiteSpace, after which it was vital to coordinate the data.

(a) Time slicing

The time span specifies the range of citations published over a certain time, the value of which is dependent on the citation year distribution and the period the analyst is interested in. The time division length refers to the complete time span, which is divided into equal yearly time intervals.

(b) Term source and term type

The cluster tags are derived from the Title, Abstract, Author Keywords, and Keywords Plus (ID). The keywords are divided into noun phrases and emergent words; however, CiteSpace automatically extracts specific noun phrases as the cluster labels to reflect the specific research focus.

(c) Node types

Depending on the subject analysis, the cited journal is selected as the node type, which then determines the type of object that is represented in the graph. Cited journals are used as important indexes for the terms extracted from the published articles in the journals or databases.

(d) Set threshold

CiteSpace controls the number of network nodes based on a threshold value in a single time zone. The citation that satisfies the threshold condition is then visualized under four settings: Top N; Top N%; Threshold Interpolation; and Select Citers. Top N selects the highest cited N citations, Top N% selects first by the number of citations and then by the percentage of citations (N%) in the same time threshold, and Threshold Interpolation sets the threshold from the citation counts, co-citation counts, and co-citation coefficients (c, cc, ccv), the formula for which is:

$$CC_{\text{cosine}}[i, j] = \frac{cc(i, j)}{sqrt(c(i), c(j))}. \tag{3.1}$$

A threshold anchor is set in the time span for the first, middle and last partitioned times, and a linear interpolation algorithm is used on the rest of the time partition threshold to calculate the time needed to achieve a different personalized partition threshold. The citation literature was selected (Select Citers) based on the TC value in the citation records, and then one of the three methods (Top N, Top N%, or Threshold Interpolation) was used to screen the references in the citation literature. A reasonable threshold can be calculated and compared depending on the number of selected citations, nodes, and connections in the lower left corner of the CiteSpace interface (Space Status and Process Reports).

(e) Pruning and visualization

The Minimum Spanning Tree is used to control the number of network connections, reduce connection density, and reduce cross connections. Using the Pruning sliced network and the Pruning merged network tools, the earliest connections are found and retained and the entire time span for the merged network structure is displayed. After setting the parameters, GO is clicked to start the document screening, from which a statistical analysis and a results graph that shows the nodes and lines are produced.

Atlas interpretation

The goal for analysis is to explore the themes and discover the graph's meaning. The node object relies on the type of node-set and node ring thickness, with the corresponding time partition being proportional to the number of cited journals for which the ring color represents the cited year; the greater the nodes across the time, the higher is the number of citations. Therefore, the size of the node is proportional to the number of citations, with the number of citations being standardized by publication date. The connections between the nodes indicate common citation relationships, and the connections between the line length and thickness indicate the connection strength between the two nodes. The length and width of the edges are proportional to the same coefficient. A visual view of the category identification is shown in Fig. 3.10.

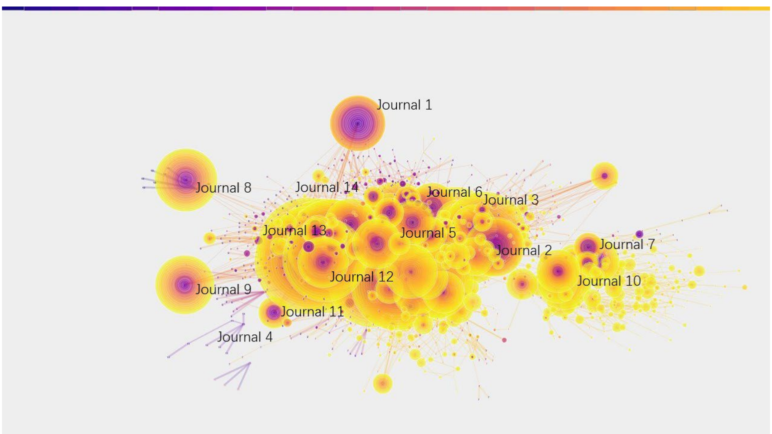

Fig. 3.10: Visual view for category identification (V-B)

As shown in Fig. 3.10, Journal 8 was the biggest node, indicating that it was the most cited. Further, the connection degree between Journal 2 and Journals 3, 5, and 6 was also the highest. Fig. 3.10 only shows a selection of the higher frequency journals. It should also be noted that “m” is less than or equal to the total number of retrieved journals.

Frequency selection

CiteSpace visualization is clear and easy to read. The different colors, node sizes, and node positions make it easy to identify the high citation frequency nodes, the distribution of the cited journals, and the citation years; that is, the cited journal analysis identifies the relevance and academic influence of the journals. Based on the cited frequency of the core journals, a list is generated, which can be sorted by citation frequency, centrality or year, that gives the name of the journal, the publication time and other information. A node in the list can also be shown or hidden. Fig. 3.11 shows the journal frequency list for management science.

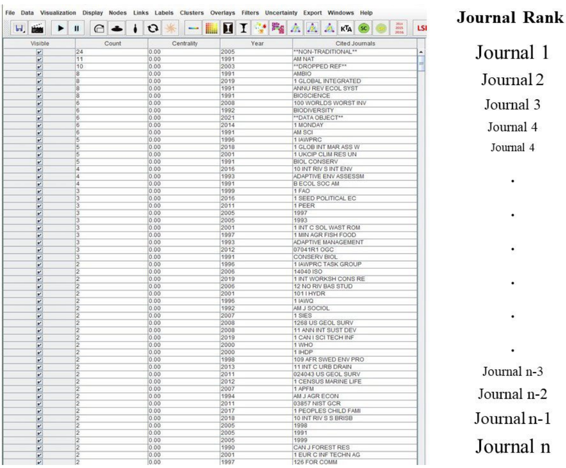

Fig. 3.11: Journal ranking I for category identification

Frequency refers to how often an event is repeated per unit of time and frequency distribution is a table that displays a sample of the frequencies from various outcomes by summarizing the data groupings and occurrences in a class. Generally, low-frequency journals are not selected as sorting objects as journals that publish fewer articles each year are cited less frequently. Therefore, it is necessary to compare the number of selected journals with the total number of articles published each year, as shown in Eq. 3.2. If F is still small, the journal is not selected.

$$F = \frac{\text{Frequency}}{\text{Annual Total of Articles}}. \tag{3.2}$$

3.4 Journal Classification

At present, the main evaluation system used for mainstream journals in China is the Chinese Academy of Sciences Journal partition table and Clarivate Analytics JCR Journal Ranking, and while both are based on SCI journal impact factors, there are many differences. Understanding the quality of a journal is important when seeking to study a subject or field from published articles.

(1) Clarivate Analytics Journal Ranking

JCR divides journals into 254 different categories as of 2021, with the subjects first classified according to the impact factor (IF) in that year and then divided into four quartiles; Q1 for the top 25%; Q2 for those between 50% and 25%; Q3 for those between 75% and 50%; and Q4 for those in the bottom 25% of the IF distribution; therefore, the journals are evenly divided into four. This method has been widely recognized by most universities and research institutions in the world.

(2) Chinese Academy of Sciences Journal Ranking

The Chinese Academy of Sciences first separates all JCR journals into 13 categories: mathematics, physics, chemistry, biology, earth science, astronomy, engineering, medicine, environmental science, agricultural science, social science, management science, and comprehensive journals. Then, four districts are generated based on the average influence factors for each journal in the previous 3 years; the top 5% is District 1; from 6% to 20% is District 2; 21%-50% is District 3, and the remaining area is District