

The Application of Cluster Analysis in the Classification of China's Modern Service Industries

Weiwei Yang, Ran Miao

School of Economics and Management, Beijing University of Technology, Beijing, China
Email: janeyang99@hotmail.com

Received 1 October 2014; revised 27 October 2014; accepted 24 November 2014

Copyright © 2014 by authors and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY).
<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The current classified researches on the modern service industries are mainly based on the state statistical standards and for the practical consideration. However, these researches which based on qualitative analysis and experience are not enough for the further study of the industrial development and internal structure relationship of the modern service industries. By further defining the connotation and characteristics of the modern service industries, this paper selects the quantitative indexes reflecting the development of the modern service industries and adopts cluster analysis to make a quantitative research on the classification of the modern service industries from the position of the academic research and the perspective of the development of the industry internal structure, hoping to lay some foundation for the further theoretical studies of modern service industries.

Keywords

Modern Service Industries, Classification, Cluster Analysis, K-Means

1. Introduction

With the rapid development of economy and the deepening of industrial structure adjustment, to speed up the development of modern service industries has always been the important content for government departments at all levels to formulate and implement the economic and social development strategies. In 2011, China's service industry accounted for the 43.4% proportion of the national economy, which closed to the proportion of the secondary industry. In particular, the proportion of part of the developed areas has nearly reached to 80%. The proportion of China's service industry in national economy has been more than 40% since 2001. However, China's

service industry proportion rose by only 2.9% after ten years of development, and even seriously decreased in some years, such as 2003 and 2008. The reasons for this result, on the one hand, come from the international financial crisis influence, which has larger negative effects on the modern service industries having a close relation to the economic situation. On the other hand, the development of China's service industry much more depends on that of the traditional service industry. While, the development of the modern service industries is still relatively backward, which is the main reason why the service industry has always accounted for much lower proportion in the national economy in recent years.

To accelerate the development of modern service industries to a higher level, the report of the 18th Chinese National People's Congress further put forward the requirement of "accelerating the transformation and upgrading of traditional industries and promoting the rapid development and growth of service industry, especially modern service industries", which raises guiding opinion to further study the development of modern service industries and find the problems in the economy development as well. The classified study of modern service industries is the basis of researching modern service industries. Especially, in the study of industry development and its internal structure, it clearly interprets the connotation, characteristics and classification of modern service industries, which will bring great help to the research of modern service industries. The quantitative analysis based on the classified study of modern service industries in this paper is a new exploration and supplement for modern service industries classification method.

2. Research Status of the Classification of Modern Service Industry Classification

The international organizations and national statistics departments still do not give a specialized standard for the classification of modern service industries. The current research on modern service industries classification made by part of the organizations and scholars are as follows:

The countries of Organization for Economic Co-Operation and Development (OECD) listed the financial insurance services, information services, education services, professional and technical services, and health care services as "knowledge-intensive" services, which are also known as the strategic services [1].

At present, the most recognised classification in China is what is proposed by Xu Guoxiang and Chang Ning (2004). They divide modern service industries into eight groups with a number of small classes: 1) The logistics and express delivery enterprises: Including logistics, the express delivery industry; 2) The information transmission industry, computer service industry and software industry: Including telecommunications and other information transmission service industry, computer service industry and software industry; 3) E-commerce: Including wholesale trade, retail industry; 4) The financial insurance industry: Including banking industry, securities industry, insurance industry and other financial activities; 5) The real estate industry; 6) Leasing and commercial service industry: Including leasing industry, consulting service industry, conference and exhibition industry; 7) Scientific research and technology service industry: Including research and experimental development, professional technical service industry, science and technology exchange and promotion services; 8) Remote education: Including academic education and non-academic education [2]. This classification method is mainly to satisfy the needs of statistical calculation of modern service industries, as well as statistical data resource sharing.

By adopting a qualitative analysis of the index weight analysis approach and a quantitative analysis of expert discriminant, Beijing statistics bureau has defined the scope of modern service industries in Beijing. Firstly, the three selected indicators of application of informationization, technology and knowledge-intensive degree, industry efficiency and benefit are given scores quantitatively in accordance with the weight of 3:4:3. After utilizing the expert judgment method to get rid of and increase certain industries, the scope of modern service industries is then defined. Then, on the basis of statistical standards, the modern service industries can be divided into nine categories: information transmission, computer service and software industry, financial industry, real estate, leasing and commercial service industry, scientific research, technical service and the geological prospecting industry, water conservancy, environment and public facilities management, health, social security and social welfare, culture, sports and entertainment [3].

Wang Zhiming, Zhang Bin and Fang Mingshan (2009) [4] think that the system standard of modern service industries statistical classification should be a dynamic system. And on the basis of two sources brought up by modern service industries: the emerging modern industry and the traditional service industry, they divide the modern service industries into the transformed and improved traditional service industry (modern financial industry, modern commerce service industry, real estate industry, modern logistics industry, modern logistics and

transportation industry, modern social services) and the emerging knowledge-intensive service industry (information transmission, computer services and software industry, leasing and commercial service industry, scientific research and technical service industry) [4]. Besides, there are also some scholars put forward different classification methods with different emphasis in the light of their own research needs and different academic viewpoints and their classification results are different as well.

Through the above analysis, the present research on the classification of modern service industries is not only mainly based on the practicability but also it relies on the statistical standards to classify with a qualitative way. Due to the scope of modern service industries has not been reached a consensus, this kind of classification method on defining the scope of modern service industries filterers modern service industries mainly on the basis of experience, which may lead to possible deviation between the scope definition and attributive character of modern service industries; Secondly, it is difficult for the qualitative classification method to reflect the inherent trends of the industry development and the relevance among industries. Thus, the classification results cannot reflect the internal structure development differences of modern service industries and the more microscopic relations among the small categories. If you want to stand in the position of the academic research and make great contributions to the government's policy, it is requested that classification design should reflect the understandings of modern service industries' internal structure. However, the existing classification design cannot meet this request [5].

Therefore, in order to apply cluster analysis to make a quantitative analysis on modern service industries from the perspective of industrial internal structure development, this paper will firstly interpret the connotation and characteristics of modern service industries, then establish the characteristic index system in accordance with the characteristics of modern service industries and select the scope of modern service industries by the data calculation. This research will introduce Cluster Analysis Method of Data Mining Theory into the classification research of modern service industries, which is not only a new attempt to the classification method of modern service industries but also is the supplement to classification methods of the national economy industry in economic research. Furthermore, this research has great reality and theoretical significance on laying the foundation for further deeply studying the internal structure of modern service industries and the industry development in China. The international organizations and national statistics departments still do not give a specialized standard for the classification of modern service industries. The current research on modern service industries classification made by part of the organizations and scholars are as follows:

The countries of Organization for Economic Co-Operation and Development (OECD) listed the financial insurance services, information services, education services, professional and technical services, and health care services as "knowledge-intensive" services, which are also known as the strategic services [1].

At present, the most recognised classification in China is what is proposed by Xu Guoxiang and Chang Ning (2004). They divide modern service industries into eight groups with a number of small classes: 1) The logistics and express delivery enterprises: Including logistics, the express delivery industry; 2) The information transmission industry, computer service industry and software industry: Including telecommunications and other information transmission service industry, computer service industry and software industry; 3) E-commerce: Including wholesale trade, retail industry; 4) The financial insurance industry: Including banking industry, securities industry, insurance industry and other financial activities; 5) The real estate industry; 6) Leasing and commercial service industry: Including leasing industry, consulting service industry, conference and exhibition industry; 7) Scientific research and technology service industry: Including research and experimental development, professional technical service industry, science and technology exchange and promotion services; 8) Remote education: Including academic education and non-academic education [2]. This classification method is mainly to satisfy the needs of statistical calculation of modern service industries, as well as statistical data resource sharing.

By adopting a qualitative analysis of the index weight analysis approach and a quantitative analysis of expert discriminance, Beijing statistics bureau has defined the scope of modern service industries in Beijing. Firstly, the three selected indicators of application of informationization, technology and knowledge-intensive degree, industry efficiency and benefit are given scores quantitatively in accordance with the weight of 3:4:3. After utilizing the expert judgment method to get rid of and increase certain industries, the scope of modern service industries is then defined. Then, on the basis of statistical standards, the modern service industries can be divided into nine categories: information transmission, computer service and software industry, financial industry, real estate, leasing and commercial service industry, scientific research, technical service and the geological prospecting industry, water conservancy, environment and public facilities management, health, social security and

social welfare, culture, sports and entertainment [3].

Wang Zhiming, Zhang Bin and Fang Mingshan (2009) think that the system standard of modern service industries statistical classification should be a dynamic system. And on the basis of two sources brought up by modern service industries: the emerging modern industry and the traditional service industry, they divide the modern service industries into the transformed and improved traditional service industry (modern financial industry, modern commerce service industry, real estate industry, modern logistics industry, modern logistics and transportation industry, modern social services) and the emerging knowledge-intensive service industry (information transmission, computer services and software industry, leasing and commercial service industry, scientific research and technical service industry) [4]. Besides, there are also some scholars put forward different classification methods with different emphasis in the light of their own research needs and different academic viewpoints and their classification results are different as well.

Through the above analysis, the present research on the classification of modern service industries is not only mainly based on the practicability but also it relies on the statistical standards to classify with a qualitative way. Due to the scope of modern service industries has not been reached a consensus, this kind of classification method on defining the scope of modern service industries filterers modern service industries mainly on the basis of experience, which may lead to possible deviation between the scope definition and attributive character of modern service industries; Secondly, it is difficult for the qualitative classification method to reflect the inherent trends of the industry development and the relevance among industries. Thus, the classification results cannot reflect the internal structure development differences of modern service industries and the more microscopic relations among the small categories. If you want to stand in the position of the academic research and make great contributions to the government's policy, it is requested that classification design should reflect the understandings of modern service industries' internal structure. However, the existing classification design cannot meet this request [5].

Therefore, in order to apply cluster analysis to make a quantitative analysis on modern service industries from the perspective of industrial internal structure development, this paper will firstly interpret the connotation and characteristics of modern service industries, then establish the characteristic index system in accordance with the characteristics of modern service industries and select the scope of modern service industries by the data calculation. This research will introduce Cluster Analysis Method of Data Mining Theory into the classification research of modern service industries, which is not only a new attempt to the classification method of modern service industries but also is the supplement to classification methods of the national economy industry in economic research. Furthermore, this research has great reality and theoretical significance on laying the foundation for further deeply studying the internal structure of modern service industries and the industry development in China.

3. The Connotation and Characteristics of the Modern Service Industries

As stated earlier, the scope of modern service industries in the theoretical circle has not reached a consensus, which has brought troubles to the theory study of modern service industries. So, making the scope definition of modern service industries is the precondition of the research on modern service industries classification. Besides, this paper hopes to research the classification of modern service industries through a quantitative analysis method. Thus, it is essential to establish a data index system for modern service industries to reflect the developmental level and internal structure of different modern service industries. Therefore, before classifying modern service industries, it is necessary for this paper to make a clear definition for the dispute of connotation and the characteristics of modern service industries based on its characteristics to establish an index system to determine the scope of modern service industries.

3.1. The Connotation of the Modern Service Industries

“Modern service industries” is a concept with significant Chinese characteristics, which is seldom used in the world. Instead, there are several similar concepts, such as “Knowledge-based service industry”, “Knowledge-intensive business services”, “Strategic service industry”, “Modern producer services” and “Knowledge-based economy” and so on which are commonly used internationally. The concept of “Modern service industries” originated from Browning and Singelman's classified study on the functionality of the service industry in 1975 [6]. In the division of the three stages of the service industry, it is believed that the third stage of the development of

service industry in Western countries is just the stage of the development of modern service industries, which is also the universal law of the development of service industry in Western countries. After the publish of the report of the 15th National People's Congress of the Communist Party of China in September 1997, the concept of modern service industries began to be paid greater attention and modern service has become the focus of the national economic structure adjustment and theoretical research.

Within the current contentious world, for the definition of modern service industries connotation, different people hold different views and it still has not been reached a consensus. But in the development process of the theory study of modern service industries and the economic life, the understandings of modern service industries are gradually deepened and some consensuses about the nature of the modern service industries have been reached, which can be mainly summarized as the following several aspects:

Firstly, it is believed that producer services are the core of modern service industries. Modern service industries came about in a relatively well-developed industrialized stage. With the development and osmosis of information technology and knowledge-based economy, the industrial production mode has turned to the complicated and top development and the social production division of labor has been further deepened. As an intermediate link of social production, the producer services have been separated from the production process, and then entered into the production area in the form of a single factor, which is the result of the further production division of labor [7]. Because of the information technology penetration and knowledge influences, the producer services possessed higher technical properties and surplus of science and technology. The rapid development of this service industry is bound to make productivity in social production rapidly improved.

Secondly, due to the infiltration and the role of information and knowledge, modern service industries are greatly different from traditional service industry. On the one hand, the emerging service industry originated from the influences of the scientific and technical revolution. Because of the service demands of the social production and living, the service industry can exist. In addition, under the influences of the information technology revolution, the social production and living brought new and high level service demands that stimulated the emergence of the emerging service industry. On the other hand, the development of information technology caused the change of production ways and means. Part of the traditional service industries adopted higher efficient technologies and modern modes of operation and organization forms, thus, realizing the promotion and transformation of traditional service industry to modern service industries.

Thirdly, the development of modern service industries is dynamic and the meaning of modern service industries will change because of the different technological development levels. There is different advanced technology in different ages. With the development of the scientific technological level and the rapid change of the social technology and the mode of operation, the service demand is also dynamic. As a result, the meaning of modern service industries is continuously changing as well [8].

From the above analysis, the following conclusion is made: Modern service industries are different from traditional service industry from the meaning of information and knowledge. And it came about with the industrialization process. Relying on the information technology, the modern management concept and organizing form, modern service industries has been a high value-added and knowledge-intensive production and life service industry which not only includes the new service industry but also includes the traditional service industry that has been transformed and improved by the new technology, new commercial activities and new service mode.

3.2. The Characteristics of Modern Service Industries

With the development of modern service industries, its characteristics are also continuously changing. With the increasing research on modern service industries characteristics, its understandings have been deepened. Combining the above analysis of the connotation of modern service industries, the characteristics of modern service industries can be mainly manifested in the following aspects:

A. The essential attributes of modern service industries:

Modern service industries are high technological, high knowledge, high value-added and opening industries. The development of modern service industries, especially producer services, is the result of the deepening of the social production division to a certain extent. This division of labor is accompanied by the information technology and the osmosis of knowledge that makes this kind of industry with high technical attributes and high knowledge content. The operation and management of this industry should be schemed, managed and operated by people with specialized knowledge. Besides, people working in this industry should not only be familiar with

the industry's business but also should be equipped with abundant theoretical knowledge, superior skills and rich work experience. The production chain, associated with high technology content and intensive knowledge, occupies much higher profits and additional value in the process of production. It not only makes modern service industries become a high value-added industry but also is an important way to improve the social economic efficiency and enhance the competitiveness of modern service industries. The high aggregation of modern service industries is reflected in the agglomeration of the cities which is a kind of geographic concentration phenomenon associated with the industrial development. Both the manufacturing industry and high-technology and information industry will have industrial agglomeration effect in their development process. However, modern service industries have a great relationship with the agglomeration of the cities. Therefore, only when the population gathers to a certain scale can modern service industries be managed as an industry. In addition, the urban market has broad hinterlands and the superiority of quickly exchanging information, which can not only make modern service industries expand market easily and approach customers at the greatest extent but also expand the spillover effects of knowledge and learning effects in the production process. Furthermore, there are numerous scientific and technological talents and scientific-research bases with solid research foundation that can help modern service industries obtain the talent and technical support conveniently.

B. The attributes of the industrial development:

Firstly, from the market structure, modern service industries have the characteristics of low barriers to entry, and its internal structure is given priority to competitive market, which has higher openness for the competitiveness. Service industry belongs to the human capital-intensive industry, so scale effects of the mechanized and mass production cannot play significant roles in modern service industries. The key of the development of modern service industries depends on talents rather than machines and a high-quality talent team is the fundamental guarantee for the development of modern service enterprises instead of excessively relying on capital accumulation. Therefore, with lower capital barriers to entry, the entry of new enterprises is relatively easy. Moreover, as the products of service industry have personalized characteristics, with the diversity of demand, middle and small-sized enterprises can transform their strategies in flexible ways in the light of the requirement in competitive environment to adapt to their market demands. Thus, the competitiveness from the industrial structure of modern service industries is more significant and is more open to the market.

Secondly, from the industrial perspective, modern service industries should have the features of high industrial relevancy. Modern service industries is an intermediate link of production, consumption and investment as well as an important link to fuse or contact the various industries and it also has the effect of adhesive in the national economy. As modern service industries have galvanizing impacts on the high-technology industry which has higher requirement on knowledge, technology and capital, it should have higher industrial relevance. However, at present, the degree of the industrial relevance of service industry in China is still not high, which is even far lower than that of manufacturing. This result has a close relationship to the low developmental level and the inappropriate internal structure of China's modern service industries, which is also one of the reasons why China should pay more attention to the development of modern service industries.

3.3. Equations

The equations are an exception to the prescribed specifications of this template. You will need to determine whether or not your equation should be typed using either the Times New Roman or the Symbol font (please no other font). Equations should be edited by Mathtype, not in text or graphic versions. You are suggested to use Mathtype 6.0 (or above version).

Equation numbers, within parentheses, are to position flush right, as in (1), using a right tab stop. To make your equations more compact, you may use the solidus (/), the exp function, or appropriate exponents. Italicize Roman symbols for quantities and variables, and Greek symbols. Do not italicize constants as π , etc. Use a long dash rather than a hyphen for a minus sign. Punctuate equations with commas or periods when they are part of a sentence, as in

$$\alpha + \beta = x. \quad (1)$$

Note that the equation is centered. Be sure that the symbols in your equation have been defined before or immediately following the equation. Use "Equation (1)", not "Equation (1)" or "(1)", and at the beginning of a sentence: "Equation (1) is ..."

4. Study Method Design of Cluster Analysis

Based on cluster analysis, this paper tries to classify modern service industries. Cluster analysis is an important part of data mining technology and can find interesting data distribution patterns in potential data under the condition without prior experience. Different from the qualitative classification method that mainly relies on experience and specialized knowledge to do qualitative classification for things, cluster analysis is a kind of unsupervised classification method and identifies similar things by establishing a classification standard and analyzing the different characteristics of the things [9]. In this paper, K-means is adopted to classify modern service industries. K-means, also called fast cluster analysis method, is one of the most classical clustering methods, which has the characteristics of high efficiency, high speed, simple operation etc., and is one of the most commonly used methods [10].

4.1. The Selection of Indicators

To establish an index system, the principles of scientificity, systematicness, comprehensiveness and independence should be observed. In the process of building classification index system of modern service industries, the following tips should be noticed: 1) Indicators must fully reflect the characteristics of modern service industries. Because the social economy is changing all the time and the connotation of modern service industries is also dynamic, only when indicators are selected on the basis of the characteristics of modern service industries can scholars accurately define and screen out the scope of modern service industries and then classify modern service industries and ensure the pertinence of every index; 2) Indicators must have availability. On accounting of various kinds of the service industry, it is essential to consider the availability of indicators to take into account industry selection index of every industry, which fully reflects the principle of operability; 3) The selection of indicators should be as simple as it can be. Cluster analysis method is a kind of unsupervised classification method, thus, the result of choosing too much indicators will not only make cluster analysis less accelerate but also will cause difficulties in analyzing the results of classification.

Therefore, according to the analysis about the characteristics of modern service industries and the principles of establishing index system in the previous chapter, this paper chooses the following aggregative indicators as the indicator to classify modern service industries (Table 1):

Table 1. Indicator system of modern service industries.

| First grade indicators | Second grade indicators | Third grade indicators | Calculation Method | | |
|------------------------|--|------------------------------------|--|--|--|
| Development level | Industry scale | Number of employees | Number of employees | | |
| | | Informatization level | The per capita amount of computer | Capita amount of computer/Number of employees number of employees | |
| | Essential attributes | Technology and knowledge-intensity | The proportion of college staff | Number of college staff/Number of employees | |
| | | | The proportion of intermediate above title | Number of intermediate above title/Number of employees | |
| | | | The proportion of senior workers and above | Number of senior workers and above/Number of employees | |
| | | Efficiency | Productivity | Annual operating revenue/Number of employees | |
| | | | Input and output efficiency | Annual operating profit/Annual operating cost | |
| | | | Herschman-Herfindahl index | | |
| | | Market structure | | The average staff size | Number of employees/Number of enterprises |
| | | | | The proportion of non-public enterprises | Number of non-public enterprises/Number of enterprises |
| | The proportion of foreign-funded enterprises | | Number of foreign-funded enterprises/Number of enterprises | | |
| Development attributes | Industrial relevancy | Induction coefficient | | | |
| | | Influence coefficient | | | |

Here, it should be noticed that Herself-Heyerdahl index is used to measure the indicators of geographic agglomeration degree of industrial development:

$$HH_i = \sum_{j=1}^m (s_{ij} - x_j)^2 \quad (i = 1, 2, \dots, n; j = 1, 2, \dots, m)$$

s_{ij} indicates the proportion of employment or production of i in the area of j ; x_j indicates the proportion of the whole employment or production in the whole country; m indicates the number of province. If the distribution of one industry is consistent with the distribution of overall economy of one region in the nationwide, then $HH = 0$. Besides, Heyerdahl, Entropy index, Heil index and Gini Coefficient can also be employed to measure the indicators of geographic agglomeration degree of industrial development [11]. Considering the larger differences of economic development degree, population and territory among every province in China and to avoid these differences having negative effects on the industrial geographic agglomeration degree, Heyerdahl is chosen to measure the degree of industrial concentration in this paper.

Induction coefficient and influence coefficient are the core indicators to measure the industry relevant effect in the industry relevant theories:

$$\text{Induction coefficient of } i \text{ industry} = \frac{n \sum_{j=1}^n q_{ij}}{\sum_{i=1}^n \sum_{j=1}^n q_{ij}} \quad (i = 1, 2, \dots, n)$$

$$\text{Influence coefficient of } j \text{ industry} = \frac{n \sum_{i=1}^n q_{ij}}{\sum_{j=1}^n \sum_{i=1}^n q_{ij}} \quad (j = 1, 2, \dots, n)$$

q_{ij} is the element of Leontief inverse matrix: $(E - A)^{-1}$. Induction coefficient measures the spreading influences from other industries on a certain industry, which embodies that when a certain industry provides supplies for other industries, it can have spreading effects on that industry; Influence coefficient measures the spreading influences from a certain industry on other industries, which embodies that when a certain industry absorbs the supplies provided by other industries, it will have spreading influences on these other industries.

4.2. The Processing of Index Data

In this paper, according to the standard of *National Industries Classification* (GB/T4754-2002), the industrial category and meaning of modern service industries are defined. The index data of 339 subclass industries comes from the 179 service industries of modern service industries comes from the industry data of *The Chinese Economic Census Yearbook in 2008* and *Chinese Input-Output Tables in 2007*. As the number of computer owners has not been given general investigation, the number of computer owners in this paper is calculated on the basis of the economic census data; Heyerdahl is derived in the light of the *Chinese Basic Unit Statistical Yearbook in 2012*. Besides, part of the financial industry data is from *Chinese Finance Statistical Yearbook*, *Chinese Securities Statistical Yearbook* and *Chinese Insurance Industry Statistical Yearbook*.

Cluster analysis takes the distance between sample points as its classification basis. Though different cluster analysis methods have different distance definition formulas and different clustering algorithm, the calculation of distance will always have great influences on the clustering results. Consequently, the order of magnitude differences caused by different units of measurement between different indexes will have larger impacts on the distance between the data points in the process of clustering. Meanwhile, the indexes with larger orders of magnitude can conceal the differences between the indexes with minimal orders of magnitude, which will make the classification results tend to be classified in accordance with the indexes of larger orders of magnitude. In order to avoid this useless deviation having negative influences on the classification result, every index data should be standardized with the following formula:

$$x'_j = \frac{x_j - \min x_j}{\max x_j - \min x_j} \times 100 \quad (j = 1, 2, \dots, n)$$

As for the index which is inversely proportional to the index data and their pre-reflected characteristics, for example, the average enterprise scale of a certain industry is inversely proportional to its degree of openness, they should be standardized by the following formula:

$$x'_j = \frac{\max x_j - x_j}{\max x_j - \min x_j} \times 100 \quad (j = 1, 2, \dots, n)$$

With the help of standardization, the minimal index data is 0, the maximal one is 100 and the other data is distributed among 0 - 100.

4.3. The Scope Definition of Modern Service Industries

In this paper, cluster analysis is adopted to classify modern service industries and the data in the previous chapters are the data of the whole service industries. In the process of cluster analysis, the industries involved in the cluster must be simplified to get rid of the samples which do not belong to modern service industries and reduce the influences on the classification results from these samples. Meanwhile, it is helpful to adjust and analyze the classification results. Consequently, it is essential to define and filtrate the scope of modern service industries.

About the definition of the scope of modern service industries, it is generally acknowledged that there are two kinds of standards, that is, the strict standard and broad standard. From the perspective of the strict standard, modern service industries should have the characteristics of all modern service industries. If some industries lack of any characteristic of the whole characteristics, they cannot be classified as modern service industries. On the contrary, according to the broad standard, modern service industries should be defined with more broad scope. The industry that has one or more characteristics of modern service industries, it then should be classified as modern service industries.

According to the discussion on the connotation of modern service industries, it is generally considered that the scope of modern service industries should be defined from the perspective of the relatively broad standard. Firstly, this definition can ensure the characteristics of modern service industries not too distinct. The industries having greater effects on social economy, especially on improving social productive forces, should remain in the definition scope of modern service industries to analyze their state of development. Secondly, as a result of the different industry development situation, the characteristics of some emerging modern service industries and the services which are transformed and upgraded because of information and technology are not very significant. If these industries are excluded from the scope of modern service industries, the research results will not be reasonable.

In this paper, the authors have carried on quantitative and qualitative analyses on the scope of modern service industrial in accordance with the relatively broad standard to determine the scope of modern service industries. The analyses are concretely made as follows: Firstly, according to the industry data made in the previous chapters, 13 kinds of three-level index data in every industry is calculated and ranked on the basis of average weight with the proportion of 1:1. The 40% of top industries are selected as modern service industries with priority. The service industries with the ranking of 40% - 60% are defined based on the definition result of other industries. Secondly, according to the definition based on the scores, the industries, such as transportation industry, accommodation and catering industry, public administration and social organization industry, which obviously neither belong to the modern service industries nor the government public administrative departments are excluded. Besides, the other industries, such as posting and delivering services, the other commercial service industries and other financial activities not mentioned, for which include logistic industry, the conference and exhibition industry and financial industry will be defined in accordance with their general industrial attributes to determine whether they should be included in the scope of modern service industries. As a result, 67 kinds of industries are finally classified as the subclasses of modern service industries, which can be seen below.

4.4. Cluster Analysis on the Modern Service Industries

After preliminarily defining the modern service industries, the authors use the statistical software SPSS 21.0 to make a cluster analysis on the 67 kind of industries including modern service industries though K-means. When K-means makes a cluster analysis, it will automatically confirm the initial clustering center and gradually iterate and revise the clustering center in its calculation process, but the number of clusters needs to be determined be-

forehand by K-means, that is to determine the value of K. In the application of K-means, it is always rather difficult to determine the value of K. And different scholars put forward different methods to solve this problem. It is generally considered that the value of K should be close to \sqrt{n} . To make it clear, several examples are given here. When K = 8, Herschman-Herfindahl index from the clustering result is not very significant. When K = 9, all the indexes have quite reasonable significance. Consequently, when the authors choose K = 9, 67 kinds of modern service industries are divided in to 9 categories though a cluster analysis and the results can be seen below (Table 2, Table 3):

Table 2. Cluster results of K-means when K = 9.

| Category | Industries included |
|----------|--|
| 1 | Trade brokers and agents, machinery equipment leasing |
| 2 | Transportation agency services, the other posting and delivery services, telecommunications, Internet information services, computer system services, data processing and other computer services, public software services, the other software services, consulting and investigation, the advertising industry, medical research and experimental development, other science and technology services, nature conservation, environmental governance |
| 3 | Finance Company |
| 4 | Life insurance, property management, enterprise management services, other commercial services |
| 5 | Natural science research and experimental development, engineering and technology research and experimental development, research of agricultural science and experimental development, social humanity research and experimental development, earthquake services, marine services, surveying and mapping services, technical detection, environmental monitoring, engineering technology and planning management, other professional technique services, science and technology services, mineral geological exploration, broadcast, film, literary and artistic creation and performance, journalism, basic geological exploration, geological exploration technology services, radio and television transmission services, other real estate activities, non-life insurance, insurance auxiliary services, legal services, intellectual property services, travel agency, audiovisual production |
| 6 | Satellite transmission service, meteorological service, technology extension service, publishing, television, the other culture and arts, arts and culture brokerage |
| 7 | Commercial bank |
| 8 | Other banks, securities market management |
| 9 | Securities brokerage and trading, securities investment, securities analysis and consultation, financial trust and management, financial leasing, financial activities not classified, real estate development |

Table 3. K-means cluster centers when K = 9.

| Category | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|--|------|------|------|------|------|------|------|------|------|
| Number of employees | 7.9 | 14.3 | 0.2 | 67.9 | 5.4 | 3.9 | 100 | 1.8 | 12.7 |
| The proportion of college staff | 26.3 | 55.3 | 91.4 | 19.1 | 59.4 | 62.8 | 74 | 86.7 | 85.5 |
| The proportion of intermediate above title | 12.6 | 19.6 | 88.2 | 9.7 | 41.7 | 32.8 | 48.1 | 63.1 | 44.6 |
| The proportion of senior workers and above | 9.2 | 15.6 | 4.3 | 11.8 | 26.3 | 13.5 | 3.7 | 7.8 | 6.6 |
| The per capita amount of computer | 0.7 | 7.2 | 1.2 | 0.4 | 2.6 | 3.6 | 2.9 | 34.6 | 27.5 |
| Productivity | 27.7 | 6.3 | 100 | 5.9 | 4.7 | 4.9 | 30.4 | 69 | 33.3 |
| Final cluster centers | | | | | | | | | |
| Input and output efficiency | 6.3 | 7.5 | 100 | 7.6 | 9.4 | 6.6 | 8 | 23.4 | 9.7 |
| The average staff size | 99 | 94.8 | 97.2 | 70.8 | 95.3 | 95.8 | 17.6 | 78.8 | 91.2 |
| The proportion of non-public enterprises | 78.5 | 92.8 | 91.7 | 77.2 | 69.8 | 70 | 68.4 | 14.6 | 92.3 |
| The proportion of foreign-funded enterprises | 15.6 | 56.9 | 7.3 | 16.6 | 9.4 | 6.7 | 3.8 | 4.2 | 17.8 |
| Induction coefficient | 94.7 | 15.8 | 40.7 | 42.2 | 17 | 8.7 | 92.9 | 70.9 | 43.2 |
| Influence coefficient | 56 | 72.2 | 3.5 | 81.5 | 67.8 | 61.7 | 12.7 | 5.8 | 3.4 |
| Herschman-Herfindahl index | 6.3 | 9.1 | 8.2 | 5.6 | 10.5 | 59.6 | 9.7 | 15.1 | 6.2 |

According to the above results of the cluster analysis, it can be seen that the 5th category consist of 27 kinds of industries. In term of the index data, most of the industries are industries with average index data. In order to make the classification results not too aggregated, the authors try to make a cluster analysis on the industries in the 5th category and then reanalyze the results and adjust the categories of these industries on the basis of their clustering center. Then they make use of K-means and ensure $K = 5$ to make a cluster analysis on the industries in the 5th category. The final results are as follows (Table 4, Table 5):

4.5. The Classification Results of the Modern Service Industries

Based on the above results of cluster analysis, the authors make the following adjustment and classification to the classification results to ensure them more reasonable:

1) To deal with the results of cluster analysis on the basis of the figure characteristics of clustering center. The scattered industries, such as commercial banks, finance companies, are merged with the industries which have similar index data with them. The industries with distinctly different industry attributes and index data are separated and then merged into the industries that have similar industry attributes and index data with them to ascertain the number of classification.

2) As the classification results of cluster analysis are made by distance similarity in accordance with the

Table 4. Result of K-means cluster separately for the fifth category when $K = 5$.

| Category | Industries included |
|----------|--|
| 1 | Natural science research and experimental development, engineering and technology research and experimental development, research of agricultural science and experimental development, social and human science research and experimental development, seismic services, marine services, surveying and mapping services, technical testing, environmental monitoring, engineering and planning management other professional and technical services, technology intermediary services, mineral geological exploration, radio, film, literary and artistic creation and performance |
| 2 | Journalism |
| 3 | Basic geological prospecting, geological exploration technology services |
| 4 | Radio and television transmission services, other real estate activities |
| 5 | Non-life insurance, insurance ancillary services, legal services, intellectual property services, travel agencies, audiovisual production |

Table 5. Centers of K-means cluster separately for the fifth category when $K = 5$.

| Category | 1 | 2 | 3 | 4 | 5 |
|--|-------|-------|-------|-------|-------|
| Number of employees | 5.2 | 0.15 | 0.79 | 12.38 | 5.89 |
| The proportion of college staff | 57.64 | 90.37 | 48.12 | 30.63 | 72.48 |
| The proportion of intermediate above title | 49.28 | 34.22 | 70.09 | 17.26 | 21.34 |
| The proportion of senior workers and above | 27.14 | 16.98 | 84.54 | 24.99 | 6.62 |
| The per capita amount of computer | 1.8 | 2.74 | 4.19 | 6.42 | 2.83 |
| Productivity | 3.87 | 8.45 | 2.96 | 4.34 | 6.99 |
| Final cluster centers | | | | | |
| Input and output efficiency | 8.4 | 8.53 | 6.51 | 9.5 | 12.95 |
| The average staff size | 96.28 | 89.99 | 93.57 | 93.46 | 94.96 |
| The proportion of non-public enterprises | 71.37 | 0 | 38.28 | 67.29 | 88.57 |
| The proportion of foreign-funded enterprises | 10.8 | 4.7 | 1.35 | 12.49 | 7.92 |
| Induction coefficient | 11.48 | 2.48 | 1.17 | 26.94 | 36.03 |
| Influence coefficient | 64.81 | 66.12 | 81.96 | 15.53 | 88.81 |
| Herschman-Herfindahl index | 10.38 | 7.28 | 13.95 | 14.94 | 8.68 |

characteristic indexes, it is difficult to define and analyze the industry similarities and attributes among the sub-classes of each industry. Thus, the classification results are too scattered, which will have negative effects on the application and analysis of the classification results. Therefore, it is necessary to adjust the industry classification results on the basis of cluster analysis.

3) For some emerging industries or the industries whose statistical yearbook has not been calculated, such as the conference and exhibition industry, logistics and express delivery industry, e-commerce industry and other industries, should be classified in the light of their corresponding industry nature.

The final result of the modern service industry classification in this paper is shown in **Table 6** as follows.

5. Summary

In this paper, K-means cluster analysis is introduced into the classification of modern service industries. By analyzing the index data, modern service industries in China is classified. From the classification results, it can be seen that the industries from the current modern service industries show different development characteristics due to their different industrial nature and development conditions. Moreover, each industry in modern service industries shows some similarities on quantity, which will not only lay the foundation for further studying the development trends and internal structure of modern service industries from the perspective of the industrial development but also provide some advantages for scholars to study the relationship between each industry from modern service industries and other industries.

With the growth of national economy and the industry development, the quantity index will continuously change. While, the scope of modern service industries will also be a dynamic development process because of its dynamic characteristic. Therefore, the classification method in this paper will offer a research framework and lay the foundation for further research on modern service industries development and industrial structure.

Table 6. Modern service industry categories.

| Industry categories | Industrial property | Industrial subcategories |
|---|---------------------------------------|---|
| High industrial relevance Industry | Intermediate of production | Trade brokers and agents, machinery equipment rental |
| High-impact and open industry | Production support | Transportation agency services, other service delivery concerns (logistics), telecommunications, Internet information services, computer system services, data processing and other computer services, public software services, other software, services, consulting and investigations, the advertising industry, medical research and experimental development, other technology services, nature conservation, environmental governance |
| Large-scale, high-impact industries | Commercial and residential service | Life insurance, property management, business management services, other business services (MICE), e-commerce, other real estate activities |
| High knowledge, high-tech industry | Scientific and professional Technical | Natural science research and experimental development, engineering and technology research and experimental development, agriculture, scientific research and experimental development, social and human science research and experimental development, seismic services, marine services, surveying and mapping services, technical testing, environmental monitoring, engineering and planning management other professional and technical services, technology intermediary services, mineral geological exploration, radio, film, literary and artistic creation and performance, based on geological prospecting, geological exploration technology services |
| Open, high-impact industries | Business services | Non-life insurance, insurance ancillary services, legal services, intellectual property services, travel agencies, audiovisual production |
| High-tech, high aggregation industry | Cultural Industries | Radio and television transmission services, satellite transmission services, meteorological services, technical extension services, journalism, publishing, television, art of other cultures, arts and culture brokerage |
| High knowledge, high value-added industries | Management of financial | Other banks, stock management, finance companies |
| High quality of employee, open industry | Finance and real estate | Commercial banking, brokerage and trading, securities investment, securities analysis and consultation, financial trust and management, financial leasing, financial activities not elsewhere classified, real estate development and management |

Competing Interests

Authors have declared that no competing interests exist.

Acknowledgements

This research was supported by the Philosophy and Social Science Post-Funded Project of The Ministry of Education (13JHQ052) & Youth talent cultivation plan of universities in Beijing in 2014 (CIT&TCD201404058).

References

- [1] OECD. The Service Economy 2010. <http://www.oecd.com>
- [2] Xu, G.X. and Chang, N. (2004) The Design of Statistical Standard of Modern Service Industry. *Statistical Research*, **12**, 10-12.
- [3] Anonymous (2005) Definition of Standards of Modern Service Industry Surfaced. *Data*, **9**, 46-47.
- [4] Wang, Z.M., Zhang, B. and Fang, M.S. (2009) Classification and Connotation Definition of Modern Service Industry. *Shanghai Business*, **6**, 8-12.
- [5] Pan, H.L. (2007) Ponder of the Connotation of Modern Service Industry. *Economic Review*, **12**, 51-53.
- [6] Liu, C.L. (2007) Modern Service Industry Development Theory and System Research. Ph.D. Thesis, Tianjin University, Tianjin.
- [7] Muller, E. and Doloreux, D. (2009) Managing the Protection of Innovations in Knowledge-Intensive Business Services. *Technology in Society*, **31**, 64-72. <http://dx.doi.org/10.1016/j.techsoc.2008.10.001>
- [8] Zhang, Z. (2011) Discrimination of the Essential Characteristics of Modern Service Industry. *City*, **10**, 25-28.
- [9] Fang, K.T. and Pan, E.P. (1982) Cluster Analyze. Geological Publishing House, Beijing.
- [10] Yang, W.W. (2010) China's Civil Airport Investment Research. China Water and Power Press, Beijing.
- [11] He, C.F. and Pan, F.H. (2007) Geographical Concentration and Agglomeration of Industries: Measurement and Identification. *Progress in Geography*, **2**, 3-15.