Analysis of spatial distribution characteristics of financial services industry clusters based on big data

Langchen Liu 1*

1Huazhong Agricultural University

Abstract. With the development of the times, the financial system is getting bigger and bigger, and the links between the various industries are getting closer, so we need to cluster the financial industry. But how to deal with is a problem, after thinking about comparison we found that we can make some treatment between them. So the purpose of this article is to analyze the spatial distribution characteristics of financial services industry clusters based on big data. Based on the experimental principle of data security, this paper processes some data that is known on the market and unknown within the enterprise, and simulates the experimental process by using the 4-model based $V^+$ on big data evaluation, and then the experimental results are drawn. The experimental results show that our model can analyze the spatial distribution characteristics of industrial clusters by analyzing some characteristics of financial services enterprises.

1 Introduction.

Because of the demand for information in the current era, our whole life has undergone earth shaking changes. Therefore, as a vital field of modern life, financial services have also obtained great development. Because the significance of financial services is that the whole financial industry plays its multiple functions to promote economic and social development, which plays an important role in the circulation of financial and monetary system in the current era. Therefore, in order to better develop financial services, we need to cluster distribution. To this end, we use big data to develop our understanding of this.

From the perspective of the layout of China's financial centers, the eastern part covers the major domestic financial centers, among which Beijing, Shanghai, Guangzhou and Shenzhen are the most typical [1]. Due to the national natural regional conditions, natural resource endowment characteristics, historical development track and other reasons, the eastern region has also formed its own financial service industry [2]. This systematic comparison of cluster models will help to explore the development of financial industry. The cluster development of financial service industry has the effect of technological innovation. Innovation is the driving force. As the source of economic development, financial industry cluster helps to promote the development of regional economy in spatial geography, and innovation driven development [4]. As we all know, Beijing, Shanghai and Shenzhen are the major financial institutions in China. At the same time, these fields are the birthplaces of the major technological innovation fields. Because in areas where financial factors are relatively concentrated, a large number of high-end product talents and rich market information can also provide support for technological innovation in the region [5]. It can be seen that in the cities with high concentration of financial industry, technological innovation is widespread, and this field also has good congenital advantages. In addition, if the financial enterprises in a region are more competitive in the market, the environment also encourages all kinds of enterprises and financial departments to continuously carry out technological innovation and service innovation. The innovative management system and business model promote the development and innovation development of the whole region [6]. In addition, the high concentration of financial enterprises will also stimulate more financial cooperation, and the emergence of a large number of financial auxiliary enterprises has led to the rise of emerging industries in the region [7].

Financial globalization relies on financial institutions (banking, insurance, securities and financial information management) and these financial conglomerates form the effective allocation of space-time resources, and through this way of distribution, it becomes a medium for transferring transnational investment or avoiding risks [8]. These financial activities are combined with the flow of international trade, and their root causes are accompanied by commerce, and services and Global trade turn into international trade activities [9]. Therefore, we analyze the spatial distribution characteristics of financial services industry clusters based on big data [10].
2 Algorithms and modeling processing

2.1 4 models for big data reviews V^+

This article builds a data evaluation model based on big V^+ data 4 features. First, the definition of 4 V^+ usability evaluation is given, and finally, the big data availability evaluation model is given.

Define 1 to make S=D, V> a big data volume evaluation system, where D is a big data collection, a randomly selected data sample, V is a volume measurement for a data sample; d_k V is a collection of volume measurement benchmarks for the dataset and a \( \forall v_0, v_1 \in V \) standardized data measurement baseline. Order:

\[
V_k = \begin{align*}
&v_0, v_k \geq v_0 \\
v_1, v_0 > v_k \geq v_1 \\
v_2, v_1 > v_k \geq v_2 \\
&\ldots \\
v_n, v_{n-1} > v_k \geq v_n
\end{align*} \tag{1}
\]

The availability of big data D's data volume V is defined as. Among them, is the probability that the volume of randomly selected sample data is higher than a certain benchmark measure, and the ratio of normalized base values, the higher the value, the higher the availability of D. Among them, is the probability that the volume baseline data metric is higher than that of a certain benchmark measure, and the ratio of normalized base values, the higher the availability of F.

In the above definition, the data volume measurement benchmark V_k can be given according to the characteristics of the industry, or according to the user's preference traces automatically ranked. For example, the volume baseline of a video data sample for a ping-pong match can be determined according to the length of the file byte corresponding to each game: 1st inning —- v_0, 2nd inning —- v_1, 3rd inning —- v_2, ..., 7th inning —- v_6.

Here's a definition of the availability of big data diversity.

Defining makes S=D, M> is a big data format diversity evaluation system, where D is a collection of big data, \( \forall d_k \in D \) a randomly selected data sample, a measure \( M_k \) of the diversity of data samples; d_k M is a collection of data format diversity metrics, a \( \forall m_i \in M(i = 1, \ldots, n) \) standardized data metric baseline. Order:

\[
M_k = \begin{align*}
m_0, M_k \geq m_0 \\
m_1, m_0 > W_k \geq m_1 \\
m_2, m_1 > W_k \geq m_2 \\
&\ldots \\
m_n, m_{n-1} > W_k \geq m_n
\end{align*} \tag{2}
\]

The availability definition of data format diversity M for big data D. Among them, the probability that the diversity measure of randomly selected sample data is higher than that of a certain benchmark measure, and the ratio of normalization, the higher the value, the higher the availability of M. \( P_m = P[M_k = m_k] \times (m_k/m_0)P[M_k = m_k] \times d_k m_k(m_k/m_0) \)

Definition 3 is a big S=D, T> data processing urgency evaluation system; which D \( \forall d_k \in D \) , \( T_k \), d_k , is a collection of big data, a randomly selected datasets, a measure of the urgency of a data sample, and T is a volume measurement baseline for the dataset, which is a \( \forall t_i \in T(i = 1, \ldots, n) \) standardized data measurement baseline. Order:

\[
T_k = \begin{align*}
t_0, T_k \geq t_0 \\
t_1, t_0 > T_k \geq t_1 \\
t_2, t_1 > T_k \geq t_2 \\
&\ldots \\
t_n, t_{n-1} > T_k \geq t_n
\end{align*} \tag{3}
\]

The availability of big data D's data processing urgency T is defined as. Among them, the urgency measure of the randomly selected sample data is lower than that of a certain benchmark measure, and the ratio of normalization, the smaller the value, the higher the availability of T. \( P_t = P[T_k = t_k] \times (t_k/t_0)P[T_k = t_k] \times t_k(t_k/t_0) \)

The fourth feature of big data is that the effective value of data is relatively low, which is reflected in the low freshness of the content of the data, such as many technical movements in sports games do not change much, reflected in the data content is not very different. To this end, this paper uses the difference of data content to depict and measure the availability of this feature.

Defining 4 makes S=D, F> is a big data differential evaluation system, where D is a collection of big data, \( \forall d_k \in D \) a randomly selected data sample, F is a measure of the variability of the data sample; d_k F is a collection of dataset differentiation measures, and is a \( \forall f_i \in F(i = 1, \ldots, n) \) standardized set of data variance measures. Order:

\[
F_k = \begin{align*}
f_0, F_k \geq f_0 \\
f_1, f_0 > F_k \geq f_1 \\
f_2, f_1 > F_k \geq f_2 \\
&\ldots \\
f_n, f_{n-1} > F_k \geq f_n
\end{align*} \tag{4}
\]

The availability of the data differential F for big data D is defined as. Among them, it is the probability that the difference evaluation of the randomly selected sample data is higher than that of a certain benchmark measure, and the ratio of normalization, the smaller the value, the higher the availability of F. \( P_F = P[F_k = f_k] \times (f_k/f_0)P[F_k = f_k] \times f_k(f_k/f_0) \)

In addition to the availability defined above based on the big data 4V characteristics, other factors that can be used to evaluate data availability, such as data authenticity and data cleanliness, can be condensed from domain application requirements. The big data availability evaluation model, which combines multiple domain factors, is called the 4 V^+ model.

Definition 5 makes S=D, \( 4 V^+ \) is a big data availability evaluation system, and the big data availability evaluation model is:

\[
V_0 = a_1 \times P_v + a_2 \times P_m + a_3 \times P_t + a_4 \times P_f \tag{5}
\]

Where, called a value-weighted factor \( \sum a_i = 1(i = 1, \ldots, 4), a_i \)

In the above definition, you can adjust the values based on the attention of each big data characteristic. Although
When it is calculated by algebra for expressions, it still has probability properties.\( a_iV_D \)

### 3 Experiment.

#### 3.1 Data source

We have visited a number of enterprises in the province to conduct surveys, surveys and data access, and other methods to collect data. Then carry on the statistical integration modeling input. Select from companies that meet the standards in each city and collect them through visits, access to information, and field testing. Through this group of people to investigate, our team formed a number of investigation teams to carry out this task, the data collected because of the operator's differences will be errors, we finally through the model, sampling survey methods of secondary unity to reduce the experimental error. The resulting data can only be entered into the established model, so that our data is more accurate, the results will be more accurate.

### 4 Evaluation results

#### 4.1 Simulation algorithm results

The following experimental results are derived from simulations using random and real data:

<table>
<thead>
<tr>
<th>Table1. Simulations According to the results of the experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>The amount of data introduced</td>
</tr>
<tr>
<td>The amount of data valid</td>
</tr>
<tr>
<td>Percentage of valid data</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table2. Real Data Experimental Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>The amount of data introduced</td>
</tr>
<tr>
<td>The amount of data valid</td>
</tr>
<tr>
<td>Percentage of valid data</td>
</tr>
</tbody>
</table>

By comparing the data in the two graphs above, we can see that the error in the simulation experiment and the real experiment is small and can be approximated as real modeling, i.e. the experiment is successful.

#### 4.2 Financial Services Industry Distribution Status Map Identification

![Figure 1. Financial Services Platform Data Confidence Status Map](image1)

Figure 1 shows the validity of the data obtained when we compare it with the data obtained on the credit, and from this we can determine that most of the data is true, i.e. that our experimental results are valid.

![Figure 2. Financial Services Industry Competition Map](image2)

Figure 2 can be seen, in a long time of development, many financial services industry was financed and acquired, a small number of forced to transform into other enterprises, and ultimately left a few enterprises, and finally also mostly become industry giants, the status of more and more stable, so we should change this status quo, vigorously support the development of small and medium-sized enterprises to form a good market management situation.
4.3 Key features of financial services companies

Features: Financial institutions are generally responsible for monitoring the company's global operations and technical risks. The Chief Executive officer implements a variety of long-term strategic measures to enhance operational risk monitoring by optimizing global information systems and databases. General precautions include supporting the company's business to multi-entity, multi-montification, multi-time zone development, and improving the control of complex cross-entity transactions. Promote the standardization of technology and operating procedures, improve the alternative utilization of resources, eliminate redundant regional request principles, reduce technical and operational costs, effectively meet the needs of market and regulatory changes, so that the company's overall operational risk control in the most appropriate range. As a tool to help real enterprises, finance can reach the benefit and small and medium-sized enterprise groups, which is the significance of innovative finance. For example, Dayi has plastic innovation of the Internet industry chain financial model, a wide range of joint offline credit guarantee institutions, so that investment users through Dayi plastic platform loaned to good physical operations, can provide full collateral, loan demand for high-quality plastic enterprises. Let professional institutions do professional things is the core concept of P2I model P2I industrial chain financial services aimed at rapidly integrating social stock capital, to promote the plastic industry cash high-quality liquid assets circulation, to achieve the circulation of money and goods closed loop, is committed to building a reduced circulation costs, improve circulation efficiency, ensure transaction safety of the industrial value chain service system, so as to optimize the industrial chain productivity structure, promote the upgrading and transformation of the entire plastic industrial chain, promote the ecological development of the industrial chain. In its model of innovation, Daye plastic in the letter-shen capacity, post-loan supervision, wind control measures and guarantee capacity, industry supervision has obvious advantages, can effectively balance the relationship between investment returns and risk aversion, in order to protect investment this high return, to solve the problem of financing difficulties for small and medium-sized enterprises, is the existing traditional financial market to improve and complement, but also to promote the smooth transformation of the real economy the most important driving force.

As far as financial services are concerned, they also have some significant characteristics compared to other industrial sectors, such as:

4.3.1 Less investment

Financial services have less physical capital investment and it is difficult to find a suitable physical unit to measure the quantity of financial services, which makes it impossible to define their prices accurately and thus to produce accurate price indices and quantitative indices, thus making it difficult to determine and measure the output of the financial services industry.

4.3.2 Financing intermediaries

The function of traditional financial services is the intermediary of capital and finance, while modern financial services have more and more functions related to the production, transmission and use of information, especially because of the increasing "financialization" of economic activities, so financial information has become one of the important resources of economic activities.

4.3.3 Labour-intensive

Financial services have traditionally been labor-intensive industries, and with the increasing complexity and informatization of financial activities, financial services have gradually become knowledge-intensive and human capital-intensive industries, human capital intensity and information resources in the modern financial services industry have become an important factor to determine the ability of financial enterprises to create value and the survival and development prospects of financial enterprises.

4.3.4 Liberalization

In today's era of increased domestic and international competition, financial services are in the midst of a process of major change, and the effects of information technology, deregulation and liberalization have changed and are continuing to reshape the financial services sector, and this trend will continue.

5 Conclusion.

The trend of the development of the times is vast soup, non-human can be blocked, so we should try our best to develop and change this era. Since the financial sector, one of the pillars of the national economic system, needs to be reformed, we should make adequate reforms to adapt to this era while ensuring security, so we have chosen to use the more sophisticated big data for modeling. That's all the author's point of view, if there's a mistake, and it's right.

References


