

Retraction

Retracted: The Development Trend of China's Agricultural Product Cold Chain Logistics from the Perspective of Big Data

Advances in Multimedia

Received 15 August 2023; Accepted 15 August 2023; Published 16 August 2023

Copyright © 2023 Advances in Multimedia. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] S. Wang, "The Development Trend of China's Agricultural Product Cold Chain Logistics from the Perspective of Big Data," *Advances in Multimedia*, vol. 2022, Article ID 2253448, 13 pages, 2022.

Research Article

The Development Trend of China's Agricultural Product Cold Chain Logistics from the Perspective of Big Data

Shenxiang Wang 

College of Intelligent Cold-Chain Industries, Guangzhou College of Technology and Business, Guangzhou, Guangdong, 510850, China

Correspondence should be addressed to Shenxiang Wang; v11314014@stu.ahu.edu.cn

Received 26 August 2022; Revised 14 September 2022; Accepted 26 September 2022; Published 17 November 2022

Academic Editor: Tao Zhou

Copyright © 2022 Shenxiang Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the gradual prosperity of social economy, people's demand for agricultural products consumption has also gradually increased. Through the efficient introduction of big data technology, we can optimize and innovate the management system of agricultural cold chain logistics enterprises, realize the modernization of enterprise logistics management, and effectively improve the management efficiency of agricultural cold chain logistics enterprises. Using big data technology can improve the efficiency of information docking, improve the stickiness of enterprise customers, and speed up the value-added of cold chain logistics data. This study introduces the common clustering algorithms in the field of cold chain logistics, with emphasis on the spectral clustering algorithm. The spectral clustering algorithm not only reduces the error rate but also significantly reduces the time spent in the clustering process, which shows the effectiveness and feasibility of the application of spectral clustering algorithm in China's agricultural cold chain logistics industry.

1. Introduction

Agricultural products are a broad concept, which not only refer to vegetables and fruits but also include the primary products and primary processed products of various animals and plants produced by agricultural industries such as agriculture, forestry, animal husbandry, and fishery; the specific varieties include grains, vegetables, and fruits in the planting industry; rubber, rosin, and lacquer in forestry; meat, eggs, milk, etc. in animal husbandry; marine products and freshwater products in fisheries. With the continuous development of China's socialist economy, the people's demand for agricultural products is increasing [1, 2]. The demand mainly includes the gradual improvement of the requirements for the distribution of agricultural products and great attention to the freshness of agricultural products. Cold chain logistics of agricultural products refers to a special supply chain system, in which fresh products such as meat, poultry, aquatic products, vegetables, fruits, and eggs are always in a suitable low-temperature control environment in the links of product processing, storage, transportation, distribution, and retail after being harvested,

slaughtered, or fished from the origin, so as to maximize the quality and quality safety of products, reduce losses, and prevent pollution. The development of China's cold chain logistics industry is mainly divided into three stages. In 1988, the cold chain industry of agricultural products was just in its infancy, and resources were very scarce. There was no concept of "cold chain logistics" for enterprises engaged in the logistics industry in China. Supermarket chains in first-tier cities such as Beijing, Shanghai, and Guangzhou began to sprout, and supermarkets began to use refrigerators in large quantities to ensure in-store sales, promoting the development of food cold chain. However, due to the backwardness of cold chain equipment and technology, many enterprises with ideas could only rely on modified refrigerated trucks for transportation. At the same time, because the whole market was in its infancy, the profits from cold chain logistics are also very rich, and the whole industry was in the stage of resource shortage. Since 2007, China has begun to pay attention to the necessity of cold chain logistics of agricultural products. The national development and reform commission has issued the first cold chain plan. At the same time, many TV programs and news have begun to

popularize the concept of cold chain logistics. Therefore, in the decade from 2008 to 2017, China's frozen food industry and cold chain logistics have developed rapidly. Since 2018, the cold chain market of agricultural products logistics has further changed, the cold chain demand for fresh agricultural products has exploded, the infrastructure system has been increasingly improved, and new technologies have driven the industry strongly.

China is a traditional agricultural country with a large population, and the annual consumption of agricultural products is also very much. According to the China cold chain logistics development report (2020), in recent years, the growth rate of China's cold chain logistics industry market size has reached 14%, but there are still a lot of product losses, and the daily loss of agricultural products can meet the daily needs of 200 million people. The reason for this is that the imperfect standard construction of the cold chain has caused serious losses of agricultural products, including loading and unloading standards, operating procedures, inspection mechanism, transportation, warehousing, distribution, and sales. At the same time, the report also shows that the annual consumption of perishable products in China is about one billion tons, and the goods that need to be circulated through cold chain logistics account for about 50%. In recent years, the use of refrigerated food has increased by about 10% every year, but the refrigerated transportation volume is not large, reaching about 15% of the whole process. Compared with 80%–90% in some developed countries, the gap is far [3].

In the whole cold chain mode of agricultural products, the intermittent shutdown of refrigeration equipment and the improper handling of goods in the transfer station lead to the interruption of the cold chain from time to time, which cannot maintain the whole refrigeration process and seriously damage the quality of agricultural products [4]. As the cold chain breaks down from time to time, we can collect a lot of useful information data. In order to effectively manage these information data, it is necessary to process these information data with the help of big data technology and guide the next process arrangement.

Today, big data technology is developing rapidly and has a significant impact in many fields. The development of big data is promoting the development process of science and technology. The impact of big data is not only reflected in the Internet field but also in many fields such as finance, education, and health care. In the field of artificial intelligence research and development, big data also plays an important role, especially in machine learning, computer vision, and natural language processing. Big data is becoming the foundation of intelligent society. Its advanced technical concept, as well as the efficient storage and computing capacity of massive data, can improve the productivity of the industry, help people understand the objective laws of the development of affairs, help scientific decision-making, and make related fields have better development prospects. At present, the scale of cold chain logistics of agricultural products in China is expanding rapidly [5]. Because agricultural products are perishable, the cold chain logistics of agricultural products have strict requirements for the whole

process of temperature and humidity, oxygen content, transportation time control, etc. In the era of information and data, the most effective way to improve the distribution efficiency of cold chain logistics of agricultural products, reduce its transportation costs, and improve customer satisfaction is to store, process, and analyze relevant cold chain logistics data scientifically and reasonably, so as to improve the intelligence of cold chain logistics of agricultural products. Therefore, it is necessary to study how to realize the application of big data technology in the field of agricultural cold chain logistics.

The agricultural product cold chain logistics management system based on big data is a complex system engineering, and each link will generate a large amount of data. Therefore, it is very important to develop a suitable small file integration model for the cold chain logistics of agricultural products under the big data model. Analyzing the performance of various solution algorithms requires extensive literature research. Selecting an efficient algorithm to simulate and solve the model requires certain theoretical and code capabilities, and the implementation process is difficult. The main content of this study is to explore the development trend of big data technology in the cold chain logistics industry and introduce the importance of spectral clustering algorithm in detail, so as to provide a reference for subsequent research.

2. Related Work

2.1. Current Situation of Cold Chain Logistics of Agricultural Products. Nowadays, people's quality of life is getting higher and higher, and the demand for cold chain logistics of agricultural products is also rising, but the development across the country is uneven [6]. At this stage, the main characteristics of China's agricultural cold chain logistics are as follows:

- (1) The construction cost of agricultural cold chain logistics is high and the scale is large.
- (2) In order to ensure the agility and high profitability of the system and reduce the loss of agricultural products in the process of production and circulation, there are high requirements for the ability of network information technology of the system. Because of the high requirements of cold chain logistics and the characteristics of agricultural products, cold chain logistics must use information technology to monitor the whole process of agricultural products in multiple directions [7].
- (3) Each link of cold chain logistics needs to have better organization and coordination, because the cold chain of agricultural products has strict timeliness, and there are many links of cold chain logistics, and the main functions of each link are also different.
- (4) *High Energy Consumption.* Because cold chain logistics requires products to be refrigerated and transported in a low-temperature environment, it requires a lot of power and other energy.

(5) *High Risk*. Cold chain logistics of agricultural products are vulnerable to market supply and demand, season and weather, transportation, and other factors, while the income is damaged, with high risk.

At present, China's economy has entered the new normal, and all walks of life are developing in depth. Among them, as the focus of the national logistics industry, the development of cold chain logistics has achieved preliminary results. The no. 1 central document of 2011 attaches great importance to the development of cold chain logistics and the construction of cold chain logistics system [8]. In 2014, the state issued the guiding opinions on further promoting the healthy development of cold chain transportation logistics enterprises, pointing out that it is necessary to strengthen the application of advanced information technology in the field of cold chain logistics [9]. In 2018, cold chain logistics, express logistics, and e-commerce logistics are the three key development areas, and the transformation and development of modern logistics will usher in a war. In 2019, the construction of northern railway logistics nodes is taken as an opportunity to promote the connectivity of "the Belt and Road" facilities. Harbin Europe container export cold chain train will be launched in 2022 [10]. According to the China cold chain logistics development report (2020), the demand of China's cold chain transportation industry is growing at an annual rate of 30%. At present, China's cold chain logistics system is not perfect, the transportation scale is small and scattered, and the infrastructure construction lags behind, which is prone to product loss. Therefore, we should begin to explore the construction path of agricultural cold chain logistics system to help the rapid development of agricultural cold chain logistics economy [11].

As shown in Figure 1, the cold chain transportation rate of primary agricultural products in China has always been lower than that in developed countries [12]. It can be seen from Figure 1 that at present, the cold chain transportation rate of agricultural products in China is only 15%, and that of aquatic products is only 40%, which is far lower than the cold chain transportation rate of nearly 100% abroad. Even with such a low transportation rate, there are also 10% and 17% chain breaks, while the chain break rate in developed countries is only 5%. The cold chain circulation rate of agricultural products and aquatic products is only 5% and 23% respectively, which is not popular in China, while the cold chain circulation rate of developed countries has reached 95%, indicating that there are still many difficulties to be overcome in China's cold chain transportation [13].

At present, the quantity and quality of refrigerated warehouses and refrigerated trucks in China cannot meet the needs of the rapid development of cold chain logistics. In addition, the regional imbalance of cold storage distribution in China is also one of the factors that lead to the "broken chain" of a large number of products. Perishable products will produce a lot of waste without sufficient circulation in the cold chain environment. The annual loss of fruits and vegetables is as high as 15% (the loss in developed countries is only about 5%), and the loss value is more than 50 billion yuan. In addition, the consumption of frozen food per capita

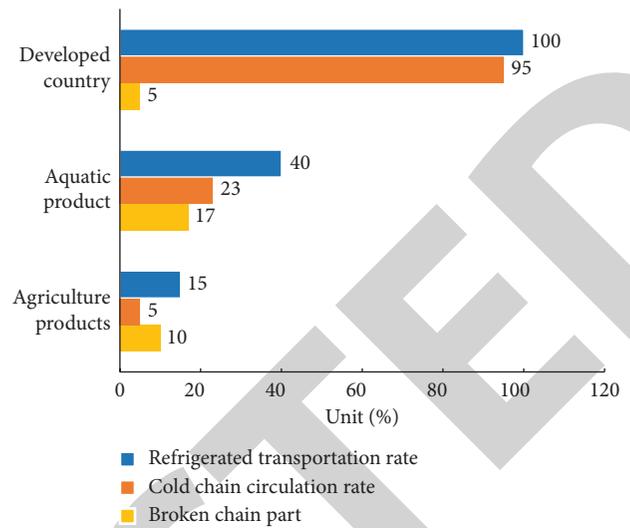


FIGURE 1: Transportation and circulation of cold chain logistics at home and abroad. Source: public data sorting.

in China is only 10 kilograms per year, half that of Japan and one sixth that of the United States. There are two reasons. One is the lack of industrialization and small scale of the cold chain food supply side; second, Chinese people do not have a strong dependence on cold chain food.

2.2. Problems in Cold Chain Logistics of Agricultural Products. During the transportation and storage, many agricultural products may be damaged, affecting the profitability and freshness of the cold chain [14]. At present, the problems existing in the cold chain logistics of agricultural products in China mainly include the following aspects:

2.2.1. Low Level of Agricultural Cold Chain Logistics Equipment and Technology [15]. In recent years, China's economic development model has gradually transformed, and the scale of cold chain logistics industry of various types of products, including agricultural products and medicine, has also expanded rapidly. However, due to weak infrastructure and imperfect technology, the efficiency of cold chain logistics in China is low. The daily operation of cold chain logistics of agricultural products requires a lot of manpower. In addition, the maintenance of cold storage, vehicles, and other facilities and the improvement of technology require a lot of capital output. However, due to seasonal characteristics, enterprises are prone to capital risks. At the same time, due to the perishable nature of agricultural products, if the cold chain link fails, it may cause a lot of losses, and there is a great risk for enterprise operators. The construction of enterprise cold chain logistics system should be paid in batches, and the funds are too large to meet the needs of some small and medium-sized logistics enterprises. The investment and construction force is relatively scarce, resulting in the relatively insufficient construction of China's agricultural cold chain logistics system.

2.2.2. The Supervision System of Product Quality Is Not Perfect. While most cold chain logistics companies attach importance to economic benefits, they do not pay enough attention to the quality of agricultural products and transportation. In practical application, advanced storage technology, temperature control, supervision, and management technology cannot be reasonably and effectively used. In order to save energy consumption, some companies cannot ensure that agricultural products are in a scientific and standardized low-temperature environment in the whole transportation process of cold chain logistics because of the use or incorrect operation of staff, which may reduce the quality and safety of agricultural products.

2.2.3. The Application of Cold Chain Logistics Technology Is Low, and the Development of Digital Information Lags behind. From the current application of cold chain logistics technology, the application degree of digital technology and information technology is relatively low. First, the data and information management of product circulation is missing. Second, the control of intelligent temperature and humidity and other environmental factors in agricultural product storage is not perfect. Digital information technology is very important to the construction of cold chain logistics system. At present, the application of cold chain logistics technology is low, which greatly hinders the construction of agricultural cold chain logistics system. The relevant operators in all links of agricultural cold chain logistics have not received standardized training, and they do not understand the standardized operation process of agricultural cold chain logistics industry and the characteristics of agricultural products, which affects work efficiency [16]. The technology inheritance of cold chain logistics enterprises mainly depends on the old and new models. This model relying on the experience of senior employees can no longer keep up with the rapid development of cold chain logistics. If the experience is unscientific and imperfect, then the development can only lag behind. In the process of cold chain transportation of agricultural products, good operation technology is needed to ensure the quality of agricultural products, but there is a lack of relevant cold chain logistics personnel training methods and systems. Such systems and standards need to be jointly formulated and cultivated by education units and work units.

3. Development Trend of Agricultural Cold Chain Logistics Based on Big Data

Under the pressure of market competition, the efficiency of agricultural cold chain logistics enterprise management needs to be greatly improved. Through the efficient use of big data technology, we can optimize and innovate the cold chain logistics enterprise management system, realize the modernization of enterprise logistics management, and effectively improve the efficiency of agricultural cold chain logistics management. The introduction of big data technology in the field of agricultural cold chain logistics can not only improve the efficiency of information docking and the

stickiness of enterprise customers but also speed up the value-added of cold chain logistics data. Based on the background of big data, enterprises can make full use of big data technology to intelligently count and analyze the relevant data information of enterprise logistics, which can fully ensure that the value-added of enterprise cold chain logistics data tends to be standardized and reasonable; the characteristics of big data technology are combined scientifically and reasonably to collect and analyze valuable data information, so as to promote the effective value-added and reasonable promotion of cold chain logistics data. In order to promote the development of enterprises, we should provide a large amount of data information support [17].

Foreign scholars have conducted in-depth research on how to apply modern information technology in cold chain logistics and achieved good results. Many developed countries use information technology to help develop cold chain logistics, so as to control the whole process of related products from raw materials, production, storage, and transportation. The United States has applied a variety of advanced modern information technologies to the production and operation of cold chain logistics, especially after it quotes the relevant technologies of big data, which has greatly improved the efficiency of cold chain logistics and obtained huge profits. Canada has established the world's largest agricultural production and transportation network covering nine countries and regions and has established a contact network with relevant personnel and organizations to share information about the whole process of cold chain transportation [18].

Saxena suggested using real-time sensor data to support supply chain decisions and described a model for measuring and improving the availability of real-time sensor data. By analyzing the data reported by wireless sensor networks, it is helpful to predict the shelf life of perishable food and take measures to prevent its deterioration. Analyzing sensor data to make decisions, rather than relying on intuition, will make decisions more scientific. These studies will encourage cold chain enterprises operating in the United States to explore value-added innovation opportunities through modern technologies such as the Internet of things and improve the relevant links of the supply chain through the experience of warehouse workers and carriage division, so as to improve the competitiveness of enterprises [19]. Reeves studied several cases from the perspective of resource dependence theory to explore the reverse logistics strategy used by some American supply chain enterprises, which controls costs by reducing risks. The data were collected through face-to-face, semistructured interviews and reviews of relevant company documents. Thematic analysis of data is carried out in five steps: compilation, decomposition, reassembly, interpretation, and conclusion. Three key themes emerging from data analysis are communication strategies, inspection strategies, and cost allocation strategies. The significance of this study includes that supply chain leaders may reduce the cost of consumers' food and beverage products by effectively implementing reverse logistics and avoid or reduce the flow of damaged and deteriorated food and beverage products into the consumer market [20].

Compared with developed countries, the development time of cold chain logistics in China is relatively short, and the relevant equipment is still insufficient. Technology research results are lacking and often only use information technology to complete the real-time monitoring function. As an important part of future scientific and technological development, big data has always been one of the strategic cores of the country. The fifth plenary session of the fifth CPC Central Committee, which ended in November 2015, clearly proposed “big data” as a national strategy. At the same time, the application of “big data” technology in the logistics industry is also gradually heating up.

Yan et al. analyzed the development form of cold chain logistics of agricultural products from the perspective of production companies and sales companies of agricultural products and gave some suggestions for improvement of the form of cold chain logistics of agricultural products combined with big data technology [21]. Wang and Shen briefly analyzed the development of China’s fruit cold chain logistics, and combined with the characteristics of big data technology, gave optimization measures in system design, facility configuration, data collection, and other aspects. It provides a reference for China’s fruit cold chain logistics, fresh fruit supply, and China’s fruit processing [22]. Li et al. analyzed the ability of information sharing to reduce the inventory of expected suppliers and retailers by studying the effect of information sharing strategies on inventory and checked the impact of information sharing on the accuracy and stability of inventory demand forecasting, so as to improve the accuracy of market demand forecasting and reduce inventory [23]. The results show that the combination of information sharing and buffer inventory can better improve the performance of fresh product supply chain. Take a project as an example. Hua analyzed and summarized the plight of the cold chain logistics company and the deficiencies of its information system. A framework suitable for processing cold chain logistics big data projects is given. A large number of Chinese scholars have begun to conduct in-depth research in this area, and with the gradual development of China’s big data model, the development of the cold chain logistics industry will usher in significant and considerable progress [24].

4. Feasibility Analysis of Spectral Clustering Algorithm in the Agricultural Big Data Cold Chain Transportation Mode

In the whole process of agricultural cold chain logistics, there will be a large number of agricultural products related information, cold storage and carriage environment information, order details information, path information, etc., which constitute a complex high-dimensional agricultural cold chain logistics information data. High-dimensional information data can more truly reflect the status of agricultural cold chain logistics and can provide guidance for the further optimization of cold chain logistics. Ordinary data processing technology cannot deal with high-dimensional cold chain logistics data efficiently, so it is particularly important to find a suitable data analysis method for the construction of agricultural cold chain logistics management system. Collecting valuable data

information and using appropriate cluster analysis methods to reduce the dimension of high-dimensional cold chain logistics data, intelligently statistics and analyze relevant data, and mining the laws and knowledge hidden in the cold chain logistics information can promote the sustainable, efficient, and coordinated development of China’s cold chain logistics industry.

4.1. Clustering Algorithm for Logistics. Big data technology has a very strong data processing ability, which can effectively process massive data information, deeply mine large amounts of data and various types of data, obtain specific value information, and actively collect all kinds of information [25]. There are many data mining methods applied in the field of agricultural products cold chain logistics, and clustering mining algorithm is one of the more commonly used methods at present. The algorithm classifies the data according to the similarity and difference between the data pairs and puts the highly universal data into the same category, so as to explore the relationship between things. The clustering mining process is shown in Figure 2. The process of clustering mining mainly includes feature selection and transformation of clustering object data source, selection and design of relevant clustering algorithm, evaluation and physical analysis of its results, and finally knowledge archiving [16].

The clustering mining algorithm in many real scenes can be more efficient data mining. For cold chain logistics data having no prior knowledge, when selecting the mining algorithm, the clustering algorithm becomes the first choice. Several common clustering algorithms are shown in Figure 3.

K-means, EM, and other algorithms will be selected in many related studies that need clustering algorithms, because their dataset dimension is not high, and the distribution space is mostly convex. However, the data space of cold chain logistics big data is often not convex and has the characteristics of high dimension, cumbersome data, periodicity, and so on. Using the above commonly used algorithms is inefficient and is likely to fall into local optimal solutions. This study presents an improved spectral clustering algorithm for data analysis, which can solve the above shortcomings to a certain extent.

4.2. Research on Spectral Clustering Algorithm. Spectral clustering is a typical algorithm based on graph theory. Unlike most classical clustering algorithms, which can only deal with convex datasets, spectral clustering algorithms do not specify the spatial style of datasets. At the same time, it can overcome the shortcoming that several classical clustering algorithms are easy to converge to local optimization. High-dimensional datasets can be mapped to low dimensions using the spectral clustering algorithm, and then the classical clustering algorithm (such as k-means and K-means) can be used to cluster the data in low dimensional space. In recent years, spectral clustering algorithm has been widely used, which has the following advantages:

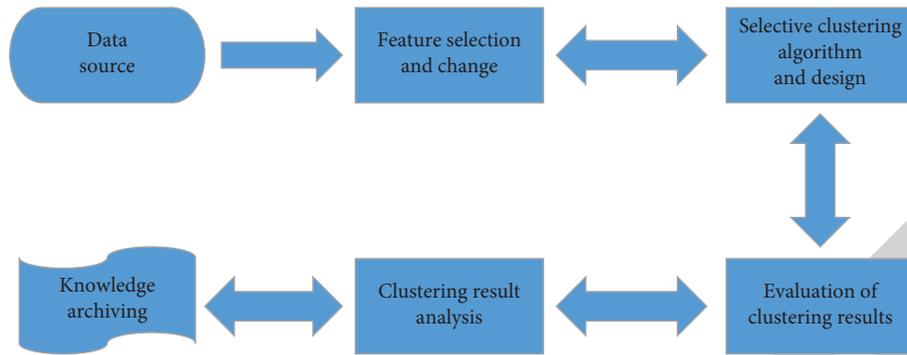


FIGURE 2: General clustering flowchart.

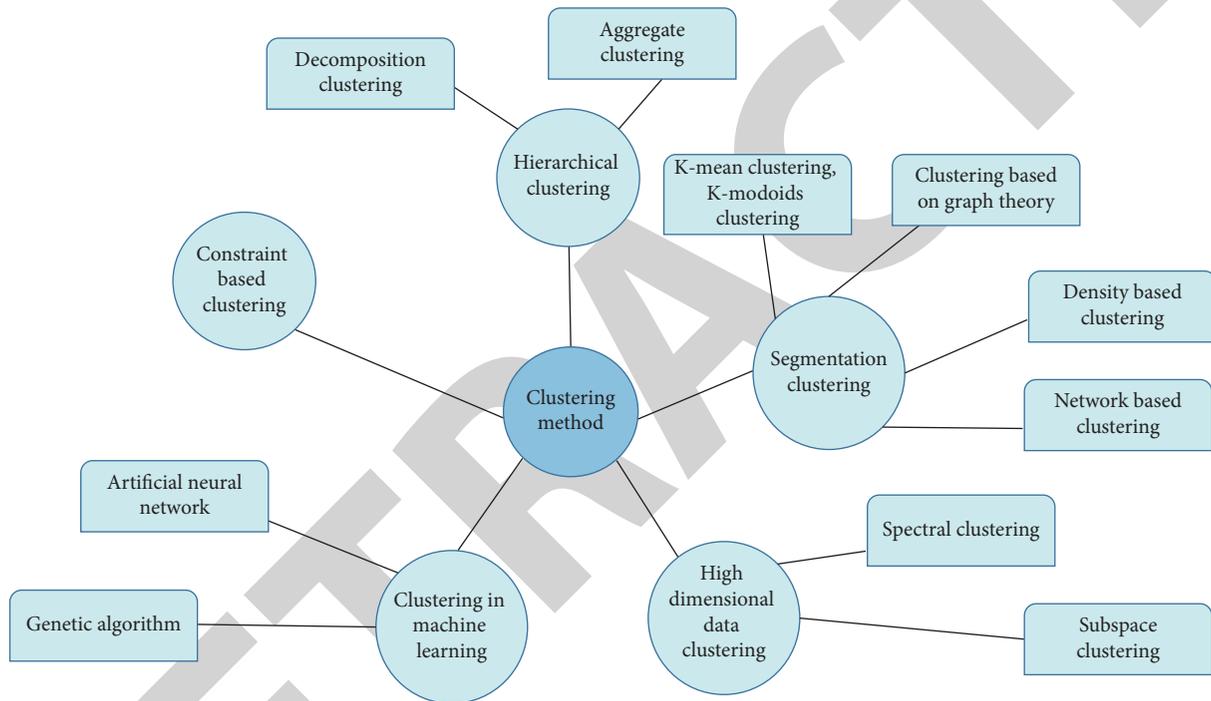


FIGURE 3: Clustering algorithm in data mining.

- (1) When clustering, it is not necessary to consider which spatial type the distribution of data points presents, that is, the spectral clustering algorithm is not limited to specific spatial characteristics.
- (2) Clustering is only accomplished by solving the eigenvector of the general matrix.
- (3) It can compress high-dimensional data, reduce the dimension to lower dimensional data, and then calculate it to reduce the workload.
- (4) The clustering process of spectral clustering algorithm is not easy to fall into the local optimal solution.

The basic idea of spectral clustering algorithm is as follows: first, the similarity matrix is obtained according to the original dataset, then the corresponding degree matrix and Laplacian matrix are obtained, then the Laplacian matrix is used to obtain the eigenvectors of the first k eigenvalues of the Laplacian matrix, and then these eigenvectors form an $n * k$

matrix U , and each row of u is used as a new dataset. These newly generated datasets are clustered into k classes by K-means or other classical clustering algorithms, and finally, the clustering results are output.

In the spectral clustering algorithm, the clustered data points are regarded as node V in the graph, and the undirected edge E is used to connect the nodes. Its weight represents the similarity between the data points, which is the undirected weighted graph, $G = (V, E)$.

The Gaussian kernel function is used to calculate the similarity of spectral clustering algorithm. The similarity of sample points is calculated as follows:

$$S_G(x_i, x_j) = \exp\left(-\frac{d(x_i, x_j)^2}{2\sigma^2}\right), i, j = 1, 2, \dots, n, \quad (1)$$

where S is the similarity between point pairs, $d(x_i, x_j)$ is the Euclidean distance of data points, and σ is the scale parameter.

The degree of vertex d_i is shown in the following equation:

$$d_i = \sum_{j=1}^n W_{ij}, \quad (2)$$

where W is the adjacency matrix, which is also the similarity matrix between data points.

Laplace matrix L is composed of the difference between degree matrix D and weight connection matrix W . The non-normalized Laplace matrix is shown in the following equation:

$$L = D - W, \quad (3)$$

where L is the Laplacian matrix and D is the degree matrix.

Normalized Laplace matrix is shown in the following equation:

$$L = D^{-1/2} L D^{-1/2}. \quad (4)$$

Finally, the algorithm calculates the normalized L matrix constructed according to the above formula to obtain K eigenvalues and eigenvectors, form an $n \times k$ eigenmatrix with K eigenvectors, and record it as Q . K-means is used to cluster Q , and the n -dimensional vector C is obtained as the result. The steps are described as follows:

The set of input n data points is shown in the following equation:

$$x = |X_1, X_2, \dots, X_n|. \quad (5)$$

The output K clustering results are shown in the following equation:

$$C = \{C_1, C_2, \dots, C_k\}, \quad (6)$$

where C is the output result of K clustering, namely, the partition result of sample points.

The original spectral clustering algorithm process can be summarized into the following four steps:

- (1) The source dataset is processed to obtain a similarity map, which is calculated from the sample space similarity. Then, calculate the connection matrix W and degree matrix D .
- (2) The Laplace matrix L matrix is obtained and normalized.
- (3) For the eigenvalue decomposition of Laplace matrix, take the eigenvectors mapped by the first k eigenvalues and combine them into $n \times k$, a eigenvector matrix.
- (4) Use the k-means clustering algorithm to cluster the eigenvector matrix obtained in step (3), and finally get the classification of each sample.

4.3. Improvement of Existing Spectral Clustering Algorithms at Home and Abroad. The key of spectral clustering algorithm is to measure the similarity between data points and divide nodes in the similarity graph and cluster the data, so as to get the results. Therefore, it is particularly important to choose

an appropriate method to calculate the similarity between data points. Appropriate similarity distance measurement can well show the correlation between data points. Constructing an appropriate similarity matrix is the key to evaluate the advantages and disadvantages of spectral clustering algorithm, which can improve the accuracy of clustering results.

Many scholars at home and abroad have studied the spectral clustering algorithm. Jia gave WLSN-NJW algorithm, which is an improved spectral clustering algorithm based on weighted local standard deviation and can make the clustering results more stable [26]. Liang et al. used variance optimization initial center to improve the K-medoids clustering algorithm and then applied it to the final step of spectral clustering algorithm to improve the clustering quality [27]. Zhang proposed a self-coding spectral clustering algorithm that combines metric fusion and landmark representation [28]. Sapkota et al. introduced a new algorithm combining spectral clustering and k-means. This algorithm replaces the initialization method of clustering centroid in the classical k-means algorithm and solves some limitations of K-means algorithm [29]. Xiang et al. derived the coefficient matrix of the object by combining the "concept of reachability similarity" of the object with the given distance-based similarity, thus solving the spectral clustering problem on multiscale data. A CAST algorithm using a lasso regularization coefficient matrix is proposed. It is proved that the coefficient matrix has a "grouping effect" and shows "sparsity." These two characteristics mean very effective spectral clustering [30]. Pourkamali-Anaraki proved that the previously popular spectral clustering method based on Nyström has serious limitations, which ignores the key information because it prematurely reduces the rank of the similarity matrix related to the sampling points [31]. In addition, the current understanding of how to use Nyström approximation to affect spectral clustering embedding is limited. In order to solve these limitations, a principled spectral clustering algorithm is proposed, which uses the spectral characteristics of the similarity matrix related to the sampling points to adjust the tradeoff between accuracy and efficiency. Chong et al. proposed a robust model fitting method, namely, spectral clustering to eliminate outliers (ORSC) and estimate multiple inner chain structures in the presence of a large number of outliers [32]. Its basic idea is to project each data point into the concept space, in which the distance distribution between the inner point and the distance from the point is significantly different from the origin. Therefore, according to the distribution of each subspace point, all points are divided into interior points and outliers by the spectral clustering algorithm. In addition, when dealing with complex multi-structure models with a large proportion of outliers, clustering results can be used to guide subsequent sampling to obtain clearer data points, so as to make assumptions.

4.4. Effect Differences of Different Types of Spectral Clustering Algorithms in the Application of Cold Chain Logistics Big Data. The transportation mode of "big data + logistics" is the standard mode of the logistics industry in the future. Big

data will play an important role in both normal temperature transportation and cold chain transportation. Therefore, as the basis of big data operation, the algorithm also has important research value, and the continuous improvement and progress of the algorithm will also be fed back to the transportation mode of “big data + logistics.”

Efficient processing and analysis of massive and multi-dimensional cold chain logistics data have become the key to solve the problems related to cold chain logistics. Cold chain logistics-related data have high-dimensional characteristics, and most of the datasets show nonconvex space. Based on this, Zuo [33] proposed an improved adaptive spectral clustering algorithm based on local standard deviation and optimized initial center (DCSC-NJW), and analyzed the data in the field of agricultural cold chain logistics. The SEEDS dataset is selected from UCI data as experimental data. The UCI database is a standard test database, which is widely used in the field of data mining and machine learning. The dataset used has a clear classification in the UCI database, which enables us to analyze and compare the accuracy of the final clustering results more intuitively.

It is very important for the spectral clustering algorithm to select appropriate similarity measurement methods. To calculate the similarity, we should first analyze the attributes of its data, such as ordinal, nominal, numerical, and other attributes. The data in this study are numerical attributes, so we use distance to measure the similarity. There are many distance calculation methods, such as Euclidean and weighted Euclidean distance, Markov distance, Hamming distance, and similarity coefficient. According to the spatial distribution characteristics of cold chain logistics data, this study uses Euclidean distance as the similarity measure like the traditional spectral clustering algorithm.

Figure 4 describes a group of cold chain logistics data. From this group of data, it can be seen that in data points a , b , and c , b and c should be classified into one class, so data points c and b should have high similarity. Its similarity should be much greater than that of a and b , so as to ensure that the calculation results classify point c and point b into the same category. When Euclidean distance is selected to measure the similarity of the above group of data, the Euclidean distance between b and a will be much greater than that between c and b . The probability that points b and a are divided into one class is far less than that of c and b . Judging from the Euclidean distance, most of the data in the above group can be classified reasonably. Through the simple analysis of the example, it is easy to see that using Euclidean distance as the method of calculating the similarity matrix is more responsive. According to the global consistency, Euclidean distance can achieve better clustering effect.

The Hadoop version used in this experiment is 2.7.2, which runs in a fully distributed mode. The operating system is centos-6.8, and the number of copies of data blocks is 3. The computer processor used is Intel (R) core (TM) i7-3540M@3.0GHz, 8 GB memory. Using 9000 files as the experimental dataset, it can be divided into three file types. The first is the image type file, with an average file size of 2.8 MB, accounting for 30.2%. The second type is text files, with an average file size of 135 kB, accounting for 63.22%.

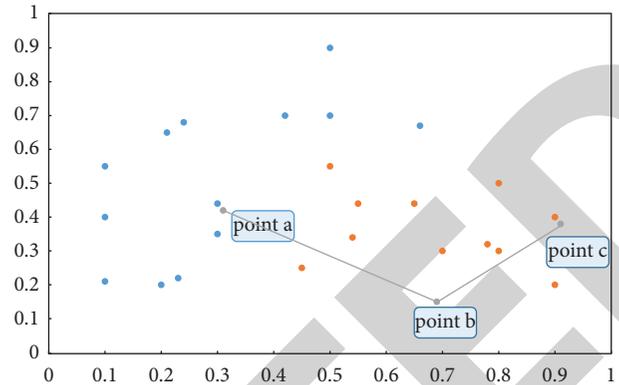


FIGURE 4: Euclidean distance can roughly reflect the global consistency of cold chain logistics data.

The third type is video files, with an average file size of 68 MB, accounting for 6.58%. The 9000 experimental documents were randomly disrupted and divided into 6 groups for testing. The number of documents was 1500, 3000, 4500, 6000, 7500, and 9000, respectively.

Next, the writing speed, the memory occupation of the namenode, and the file reading speed are tested.

4.4.1. Write Speed Test. In order to verify that this method improves the file upload rate, six groups of files are uploaded to HDFS according to the original HDFS method, HAR method, and the method proposed by Zuo, and their upload time is recorded. Experiment is repeated three times and the average value is calculated. The experimental results are shown in Figure 5.

Experimental results show that the proposed method takes much less time than the original HDFS and C. The reason is that the initial HDFS client will send a write request to the namenode every time it uploads a small file, and the time spent sending the request is much longer than that spent writing the file, so the upload speed is the slowest. The har method needs to run mapreduce to merge before uploading. The method proposed by Zuo does not need to run mapreduce to directly write the merged files into HDFS before uploading, and its more refined grouping and merging method makes its upload speed faster.

4.4.2. Namenode Memory Usage Test. Six groups of files are uploaded to HDFS through the original HDFS method, har method, and the method proposed by Zuo to test the difference between this method and other methods in the memory occupation of namenode. After uploading each group of files, the growth of the space occupied by the namenode editing log edit in progress is recorded. The experimental results are shown in Figure 6.

It can be seen from the experimental results that the method proposed by Zuo has good performance in the memory occupation of the namenode. The main reason is that this method can classify and merge small files into large files more efficiently to reduce the amount of metadata uploaded to HDFS, which significantly reduces the memory

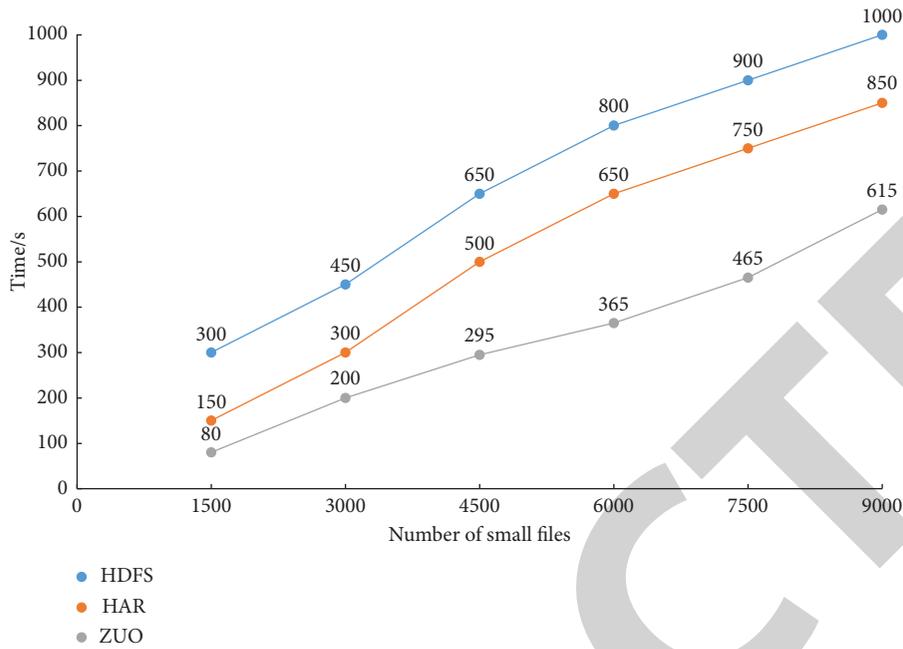


FIGURE 5: Small file write time comparison.

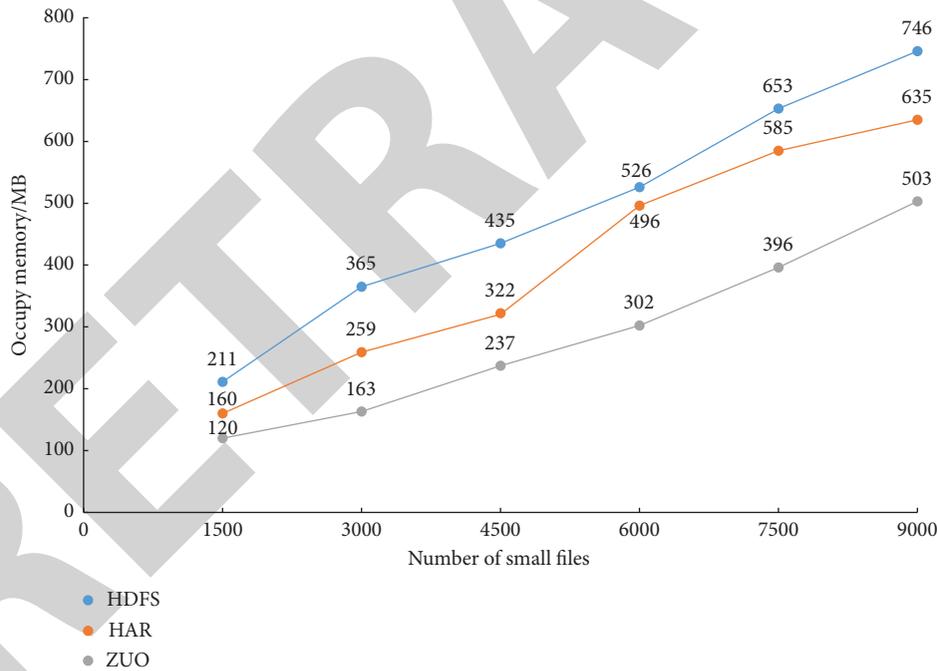


FIGURE 6: Namenode memory comparison.

consumption of namenode, and the memory occupation of namenode is significantly reduced, which also shows that the number of data blocks occupied by datanode is also greatly reduced.

4.4.3. *Read Speed Test.* In order to verify the difference between the method proposed by Zuo and the original HDFS and har methods in the speed of reading files, the time spent reading each group of files was recorded, and each

group of experiments was carried out five times and the average value of the data was taken. At the same time, setting the threshold of cache capacity n to 1000 can verify the impact of the cache mechanism on reading files. The experimental results are shown in Figure 7.

Experimental results show that Zuo's method is significantly better than the original HDFS and har methods. In the process of repeated reading, because of its cache mechanism, the reading time can be reduced. The reason is that the performance of HDFS metadata retrieval decreases

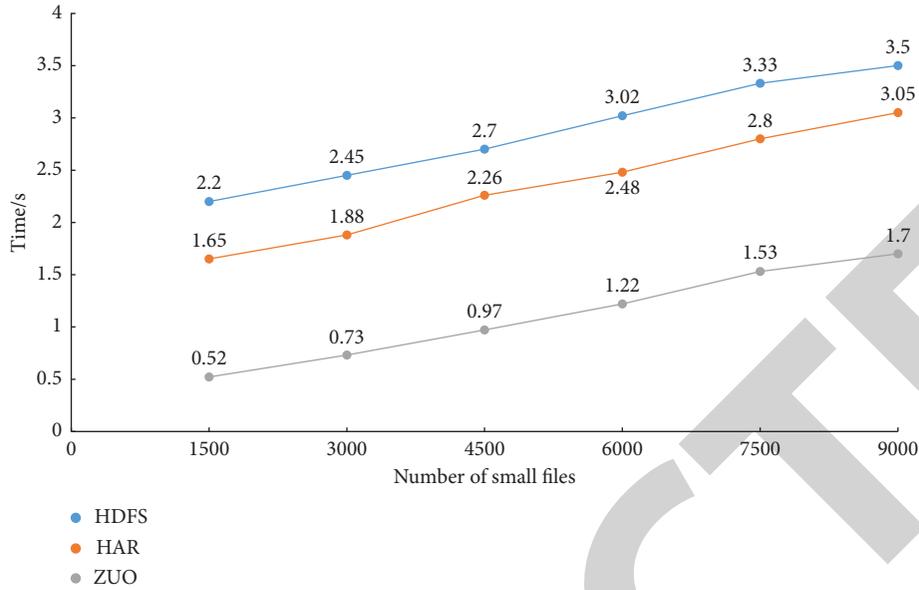


FIGURE 7: Comparison of small file reading time.

TABLE 1: Clustering results of the SEEDS dataset (number of wrong clusters per class).

Average number of error clusters in 10 experiments				
Dataset	Category	Spectral clustering	Xie's improved spectral clustering	Zuo's improved spectral clustering
SEEDS	1	13	8	7
	2	7	9	5
	3	9	6	4

TABLE 2: Clustering time of the SEEDS dataset.

Average clustering time of 10 experiments				
Dataset	Time category	Spectral clustering	Xie's improved spectral clustering	Zuo's improved spectral clustering
SEEDS	Evd_time	0.335	0.346	0.301
	Kmeans_time	0.102	0.081	0.071
	Total_time	0.501	0.478	0.389

Note: Evd_time: running time of feature decomposition; Kmeans_time: running time of K-means; Total_time: total running time.

with the increase of small files stored in HDFS. However, the method proposed by Zuo uses hash mapping to save the index information and directly obtains the index information when searching the index information according to the file name, with a time complexity of $O(1)$. When reading duplicate files, since the index information has been added to the hash mapping object in memory at the first reading, it is more efficient to query the index information directly from the hash map in memory at the second reading.

In this section, three spectral clustering algorithms are analyzed experimentally, including traditional spectral clustering, Xie et al.'s [34] improved spectral clustering algorithm, and Zuo's improved spectral clustering algorithm. SEEDS dataset is selected from UCI data as experimental data. UCI database is a standard test database, which is widely used in the field of data mining and machine learning. The dataset used has clear classification in UCI

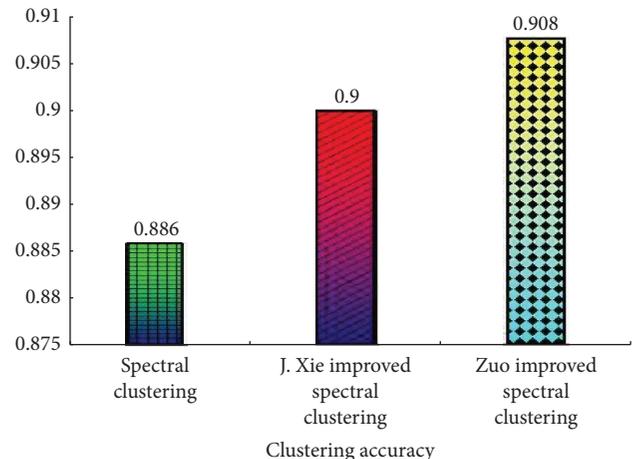


FIGURE 8: Line graph of average clustering accuracy of three algorithms in 10 experiments.

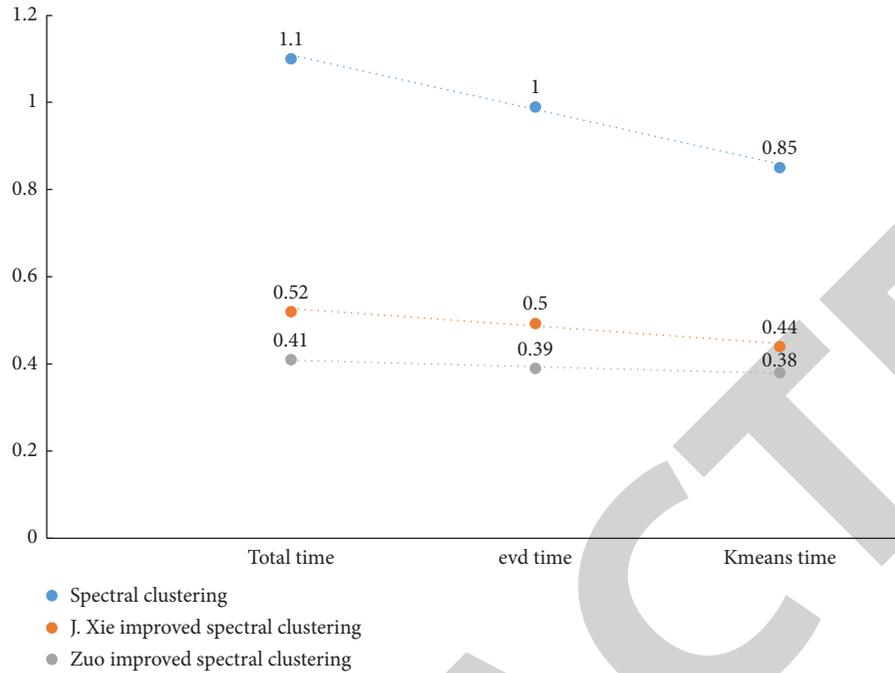


FIGURE 9: The average computing time of 10 experimental clusters of the three algorithms.

database, which enables us to analyze and compare the accuracy of the final clustering results more intuitively.

The experiment was completed in MATLAB software environment, the computer processor used is Intel (R) core (TM)i7-3540M@3.0 GHz, 8 GB memory, and the mechanical hard disk has a memory of 60 GB. Experiments are carried out on SEEDS dataset, and the algorithm is run 10 times to take its average clustering accuracy, the p value of the algorithm, and σ . According to the results of literature [35], the values are 1.3 and 9, respectively. The calculation process of the algorithm records the number of wrong clusters and classifications and the calculation time of the algorithm, as listed in Tables 1 and 2.

Through the simulation experiments on classical datasets, the data in Tables 1 and 2 show that the clustering accuracy and time of Zuo's improved spectral clustering algorithm are improved compared with the original spectral clustering algorithm and the improved spectral clustering algorithm proposed by Xie, and the error rate of clustering results is significantly reduced.

This study compares three kinds of spectral clustering algorithms on SEEDS dataset and completes the clustering according to the steps of spectral clustering algorithm. The line chart of the results is analyzed and the accuracy of the algorithm is compared. The test results are shown in Figure 8; at the same time, the operation speed of the algorithm is compared, and the results are shown in Figure 9.

By comparing and analyzing the clustering accuracy of spectral clustering, Xie's improved spectral clustering algorithm and Zuo's improved spectral clustering algorithm, it is found that Zuo's improved spectral clustering algorithm has higher clustering accuracy, which fully proves the feasibility and superiority of an improved adaptive spectral clustering algorithm based on local standard deviation and

optimizing the initial center. The simulation results are true and effective and provide an algorithm basis for the subsequent mining work of the system. According to the comparative analysis of Figure 9, the total time and step-by-step time of the three algorithms are decreasing. The time of Zuo's improved spectral clustering algorithm is reduced compared with the traditional spectral clustering algorithm and the improved spectral clustering algorithm proposed by Xie. Experiments show that the improved spectral clustering algorithm proposed by Zuo has the advantage of fast processing speed and can meet the requirements of fast processing speed of cold chain logistics big data.

From the above experimental results, it can be seen that different kinds of algorithms have certain differences in the process of cold chain logistics transportation management, and an algorithm does not have universal applicability. China's cold chain transportation mode is developing again, and the algorithm should also be constantly upgraded and improved. It should make coordinated progress with the cold chain transportation mode, combine and confirm each other, and form a set of reasonable big data management mode.

5. Conclusions

In recent years, China's agricultural cold chain logistics has developed rapidly, with huge market demand and development prospects. Big data technology is a new driving force for the development of this era. Effective use of big data technology can optimize and innovate the cold chain logistics enterprise management system, realize the modernization of enterprise logistics management, and effectively improve the efficiency of agricultural cold chain logistics management. Introducing big data technology into

the field of agricultural cold chain logistics can improve the efficiency of information docking, improve the stickiness of enterprise customers, and speed up the value-added of cold chain logistics data. At the same time, the foundation of big data supporting cold chain transportation is all kinds of mathematical algorithms. While paying attention to the appearance, we need to take into account the internal development. Constantly upgrading and improving the algorithm mode will also speed up the promotion and popularization of big data cold chain transportation, because, in the future logistics industry, the rapid transmission and exchange of information will be the basis to ensure the normal operation of the whole system.

Data Availability

The labeled dataset used to support the findings of this study is available from the author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

- [1] M. Zhou, B. Dan, S. Ma, and Z. Xumei, "Supply chain coordination with information sharing: the informational advantage of GPOs," *European Journal of Operational Research*, vol. 256, no. 3, pp. 785–802, 2016.
- [2] M. Pilkington, *Blockchain Technology: Principles and Applications*, p. 145. Social Science Electronic Publishing, Tirana, Albania, 2015.
- [3] J. Ylihuomo, D. Ko, S. Choi, S. Park, and K. Smolander, "Where is current research on Blockchain Technology?—a Systematic review," *PLoS One*, vol. 11, no. 10, Article ID e0163477, 2016.
- [4] J. Li, L. Fan, and D. Dong, "A new route optimization approach of cold chain logistics distribution based on fresh agricultural products," in *Proceedings of the 2018 Chinese Control And Decision Conference (CCDC)*, pp. 6652–6657, IEEE, Shenyang, June 2018.
- [5] Y. Zhang, "Research on optimizing transportation of fresh agricultural products by Internet of things," in *Proceedings of the .2020 International Conference on E-Commerce and Internet Technology (ECIT)*, pp. 299–301, IEEE, Zhangjiajie, China, April 2020.
- [6] D. Wang, Q. Yuan, and F. Xiang, "Analysis of current situation of joint distribution in cold chain logistics," in *Proceedings of the 2020 International Conference on Urban Engineering and Management Science (ICUEMS)*, pp. 216–220, IEEE, Zhuhai, China, April 2020.
- [7] A. Juneja and N. N. Das, "Big data quality framework: Pre-processing data in weather monitoring application," in *Proceedings of the 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-ITCon)*, pp. 559–563, IEEE, Faridabad, India, February 2019.
- [8] T. T. Htay and S. Phyu, "Improving the performance of Hadoop MapReduce applications via optimization of concurrent containers per node," in *Proceedings of the 2020 IEEE Conference on Computer Applications (ICCA)*, pp. 1–5, IEEE, Yangon, Myanmar, March 2020.
- [9] U. R. Muhammad and X. Zhao, "A Comprehensive Overview of BIG DATA technologies: a Survey," in *Proceedings of the 2020 5th International Conference on Big Data and Computing (ICBDC 2020)*, pp. 23–31, Association for Computing Machinery, New York, NY, USA, July 2020.
- [10] D. Li, Y. Gong, G. Tang, and Q. Huang, "Research and design of Mineral resource management system based on big data and GIS technology," in *Proceedings of the 2020 5th IEEE International Conference on Big Data Analytics (ICBDA)*, pp. 52–56, IEEE, Xiamen, China, May 2020.
- [11] Bo. Yu and Y. Chen, "Strengthen the establishment of enterprise logistics management system by means of big data," in *Proceedings of the 2020 4th High Performance Computing and Cluster Technologies Conference & 2020 3rd International Conference on Big Data and Artificial Intelligence (HPCCT & BDAI 2020)*, pp. 209–211, Association for Computing Machinery, New York, NY, USA, July 2020.
- [12] Y. Pang, J. Xie, F. Nie, and X. Li, "Spectral clustering by joint spectral embedding and spectral Rotation," *IEEE Transactions on Cybernetics*, vol. 50, no. 1, pp. 247–258, Jan 2020.
- [13] C. Zhang, H. Chen, C. Wu, and M. Chen, "Multiscale fast spectral clustering based on k-d tree," in *Proceedings of the 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, pp. 1607–1611, IEEE, Chengdu, China, June 2019.
- [14] F. A. Sebayang, M. S. Lydia, and B. B. Nasution, "Optimization on Purity K-means using variant distance measure," in *Proceedings of the 2020 3rd International Conference on Mechanical, Electronics, Computer, and Industrial Technology (MECnIT)*, pp. 143–147, IEEE, Medan, Indonesia, August 2020.
- [15] Y. Matsuyama, "Divergence Family Contribution to data evaluation in Blockchain via Alpha-EM and log-EM algorithms," *IEEE Access*, vol. 9, pp. 24546–24559, 2021.
- [16] M. Yin, S. Xie, Z. Wu, Y. Zhang, and J. Gao, "Subspace clustering via learning an adaptive low-rank graph," *IEEE Transactions on Image Processing*, vol. 27, no. 8, pp. 3716–3728, 2018.
- [17] A. Tang, L. Mhamdi, D. McLernon, S. A. R. Zaidi, and M. Ghogho, "Deep learning approach for network Intrusion Detection in software Defined networking," in *Proceedings of the International Conference on Wireless Networks*, pp. 258–263, IEEE, Fez, Morocco, October 2016.
- [18] P. Xiao, "Enhanced Monte Carlo Cycle framework and data mining of cold chain logistics transmission efficiency," in *Proceedings of the 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, pp. 999–1002, IEEE, Coimbatore, India, December 2020.
- [19] A. Saxena, *Usability of Real Time Data for Cold Chain Monitoring Systems*, ProQuest Dissertations and Theses Full-text Search Platform, West Lafayette, IND, USA, 2016.
- [20] L. A. Reeves III., *Supply Chain Managers' Reverse Logistics Strategies to Control Cost through Risk Mitigation*, ProQuest Dissertations and Theses Full-text Search Platform, Minneapolis, MIN, USA, 2019.
- [21] D. Yan, W. Yang, and P. Lin, "Analysis on the development mode of modern cold chain logistics of agricultural products based on big data," *China logistics and procurement*, vol. 2020, no. 16, pp. 80–81, 2020.
- [22] X. Wang and S. Shen, "Application of big data in fruit cold chain supply chain," *China fruit tree*, vol. 2021, no. 01, pp. 100–103, 2021.
- [23] L. Li, Z. Yang, and C. D. Tan, "Buffer inventory + information sharing strategy for retailers in two-Level fresh supply chain," in *Proceedings of the 2019 IEEE International Conference on*

- Industrial Engineering and Engineering Management (IEEM)*, pp. 901–905, IEEE, Macao, Macao, December 2019.
- [24] Z. Hua, *Design and Construction of Cold Chain Logistics Big Data center*, Zhejiang University of technology, Hangzhou, China, 2019.
- [25] Z. Wang, Y. Zhou, and G. Li, “Anomaly Detection by using Streaming K-means and Batch K-means,” in *Proceedings of the 2020 5th IEEE International Conference on Big Data Analytics (ICBDA)*, pp. 11–17, IEEE, Xiamen, China, May 2020.
- [26] R. Jia, *Research on Improved Spectral Clustering Algorithm and its Application in Logistics Distribution*, Shandong University of science and technology, Qingdao, China, 2018.
- [27] Z. Liang, C. Yuan, and Q. Xiao, “An algorithm for mining hot spots based on variance optimized spectral clustering,” *Guangxi Branch Student*, vol. 27, no. 6, pp. 1–6, 2021.
- [28] M. Zhang, “Zhou Zhiping A self coding spectral clustering algorithm combining metric fusion and landmark representation,” *Journal of Intelligent Systems*, vol. 15, no. 04, pp. 687–696, 2020.
- [29] N. Sapkota, A. Alsadoon, P. W. C. Prasad, A. Elchouemi, and A. K. Singh, “Data Summarization using clustering and classification: spectral clustering combined with k-means using NFPH,” in *Proceedings of the 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, pp. 146–151, IEEE, Faridabad, India, February 2019.
- [30] Li Xiang, B. Kao, C. Shan, D. Yin, and M. Ester, “CAST: a correlation-based adaptive spectral clustering algorithm on multi-scale data,” in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 439–449, Association for Computing Machinery, New York, NY, USA, June 2020.
- [31] F. Pourkamali-Anaraki, “Scalable spectral clustering with Nyström approximation: practical and Theoretical aspects,” *IEEE Open Journal of Signal Processing*, vol. 1, pp. 242–256, 2020.
- [32] M. Chong, Q. Li, T. Lai, X. Lan, X. Wang, and J. Li, “Outliers Removed via spectral clustering for robust model fitting,” in *Proceedings of the 2018 11th International Symposium on Computational Intelligence and Design (ISCID)*, pp. 229–233, IEEE, Hangzhou, China, December 2018.
- [33] K. Zuo, *Agricultural Products Cold Chain Logistics Management System Based on Big Data*, Nanjing University of Posts and telecommunications, Nanjing, China, 2021.
- [34] J. Xie, Y. Zhou, and L. Ding, “Local standard deviation spectral clustering,” in *Proceedings of the 2018 IEEE International Conference on Big Data and Smart Computing (BigComp)*, pp. 242–250, IEEE, Shanghai, January 2018.
- [35] Y. Xu, *Implementation of Agricultural Big Data Platform and Data Analysis Algorithm*, Northeast Agricultural University, Harbin, China, 2017.