

## Research Article

# Real-Time Controllable Optimization Algorithm for Correlated Big Data in Cloud Computing Environment

Rutao Li<sup>1</sup> and Zaiyi Pu<sup>2</sup> 

<sup>1</sup>School of Electronic Information Engineering, Ningbo Polytechnic, Ningbo 315800, Zhejiang, China

<sup>2</sup>Education and Information Technology Center, China West Normal University, Nanchong 637009, Sichuan, China

Correspondence should be addressed to Zaiyi Pu; [zaiyipu@cwnu.edu.cn](mailto:zaiyipu@cwnu.edu.cn)

Received 10 February 2022; Accepted 28 April 2022; Published 31 May 2022

Academic Editor: Jong M. Choi

Copyright © 2022 Rutao Li and Zaiyi Pu. 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.

The era of big data (BD) has arrived. How to train models to find correlations in data and help people make decisions has become a major research topic and direction. As an elastic and scalable distributed computing mode (refers to a whole consisting of multiple interconnected computers that cooperate to perform a common or different task in a set of system software environments, with minimal reliance on centralized control processes, data, and hardware), cloud computing can provide powerful computing and storage capabilities and has been widely used in BD query and difficult processing. This paper aims to study the algorithm in the environment of cloud computing. Different from the traditional research algorithms, the relational BD algorithm is controllable in real time. Moreover, it has been optimized and upgraded on the previous real-time controllable algorithm. Also, it performs serial and parallel simulation tests on the algorithm. When the optimal situation of the parallel algorithm is obtained, the test results show that the relevant mining time of the optimized algorithm is significantly shorter than the traditional data mining time under the same dataset. The traditional mining time is about 3.5 times the data mining time of this paper, and the running power consumption of the optimization algorithm is reduced to 20 W.

## 1. Introduction

Computers and the Internet have accelerated the change and dissemination of information. The advent of the era of information explosion has also promoted the transformation of storage technology from optical disks, chips, and card storage to disk arrays and even today's large-scale network disk array storage. The rapid development of data storage and data processing methods has accelerated the arrival of BD. Generally speaking, BD are massive data with complex structure, huge quantity, diverse types, and low value density. Their scale is too large to be understood. From a computer science perspective, BD refer to a collection of various structured and semistructured data. From a popular point of view, BD refer to a massive collection of data. Unlike traditional data, which tend to discover why things happen, BD have the advantage of predicting what will happen. Finding causal relationships is not helpful for users to make

decisions, and finding correlations between things is more valuable. For example, people place more importance on when the airfare is cheaper than why the price of the plane changes. The relationship implied by BD is more worthy of social exploration.

BD's relational data topic mining refers to mining the implicit relationship of each data item in the database. Mining frequent itemsets (the most basic pattern is an itemset, which refers to a collection of several items; frequent patterns refer to itemsets, sequences, or substructures that frequently appear in a dataset; frequent itemsets refer to sets with support greater than or equal to the minimum support) of transaction sets in database is an important part and main goal of association rules. The relationship of each data item in the database can be clearly expressed by mining association rules, which is easy for people to understand. Using cloud computing to mine the relevance of BD is an efficient and low-cost method. Static algorithms can only process

some conventional data. For some places that need emergency processing and high risks, such as emergency fire monitoring, cabin gas monitoring, and so on, static algorithms are obviously not enough to meet the requirements. Although the current dynamic algorithms can control the data in real time, these algorithms need to be further optimized and upgraded in order to calculate the data more accurately and even with the smallest error.

The innovation of this paper lies in the use of dynamic real-time controllable data algorithms. The algorithm has stronger applicability and wider range of use, and the data can be seen at any time. In addition, this paper optimizes on the basis of the previous real-time controllable algorithm, which speeds up the data detection speed, makes the data update faster, and makes the calculation results more accurate.

## 2. Related Work

With the research and development of BD such as the Internet of Things, the Internet, and medical care, more and more scholars have entered the study of BD. Among them, Puig et al. applied data analysis to aid in fraud detection and maintenance demand forecasting. Puig et al. introduced new algorithms and methods [1]. Qian et al. collected real-time observations of hourly mean wave height, temperature, and pressure at the Maida station in Qingdao, China. Using eight quality control methods, they explored data quality and identified the most effective methods for the Wheat Road station. After using eight quality control methods, the average wave height, temperature, and pressure data passed the test in percentages of 89.6%, 88.3%, and 98.6%, respectively [2]. Puthal et al. called this an online security verification problem. To solve this problem, Puthal et al. proposed a dynamic key length based security framework (DLSeF). The theoretical analysis and experimental results of the DLSeF framework show that the efficiency of processing streaming data can be significantly improved [3]. To address these challenges, Ren et al. proposed a workshop material delivery framework. They studied the key technologies of the proposed framework. To demonstrate the implementation of the proposed framework, proof-of-concept scenarios were designed to demonstrate the implementation of the framework [4]. Bao et al. proposed a novel gamma control method to tune the vertical growth rate, a method for estimating plasma vertical instability. The experimental results show that the time evolution of the real-time vertical growth rate conforms to the target value, and the real-time vertical growth rate can be adjusted by gamma control [5]. Janzen and Mann introduced a feedback control method that automatically adjusts multiple exposure settings for compositing to increase the dynamic range of sensory processes. It is synthesized to capture an extremely high dynamic range with minimal uncertainty [6]. The research of the above scholars can promote the development of big data in certain aspects, but the research on the correlation and real-time control of big data is not in-depth and perfect and needs to be further optimized.

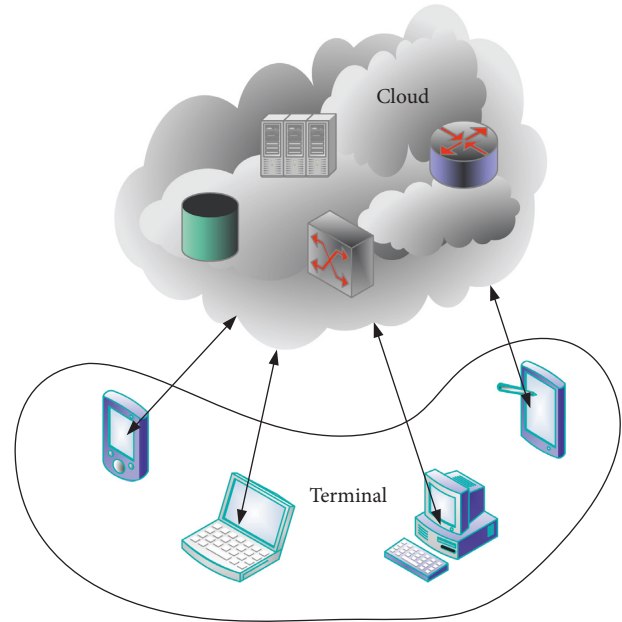


FIGURE 1: The principle of cloud computing.

## 3. Data Controllable Algorithm Based on Cloud Computing

### 3.1. Cloud Computing

- (1) *Origin.* In August 2006, Google CEO first proposed the concept of “cloud computing” at the SES San Jose 2006 conference. After this concept was formally proposed in the history of cloud computing, it caused a great sensation [7].
- (2) *Definition.* The definition of cloud computing is not yet unified. According to NIST’s improvement, cloud computing is a model that can obtain the required resources from a shared pool of configurable computing resources anytime, anywhere, conveniently, and on demand, and resources can be quickly provided and released [8, 9]. It minimizes the effort of managing resources or interaction with service providers [10, 11]. The principle of cloud computing is shown in Figure 1.
- (3) *Features.* The main features of cloud computing are as follows. First, without manpower, users can utilize the existing computing resources of any service provider. It includes processing power, storage space, or applications without human interaction. Second, it is convenient and fast, and the resources in the cloud computing system can be accessed and used anytime, anywhere. The third is information sharing. The computing resources of service providers can be combined to provide services, and these combined resources may be distributed in multiple data centers around the world. At the same time, the provider’s computing resources can be shared and used by multiple users. The fourth is resource payment. Users can use these computing resources flexibly. A user can

TABLE 1: Types of cloud computing services.

Genre	Concept	The proportion of applications (%)
IaaS	It provides virtualized computing resources to individuals or organizations of cloud computing providers	6.03
PaaS	Provides a platform for developers to build applications and services over the global Internet.	25.57
SaaS	Providing pay-as-you-go software applications over the Internet; cloud computing providers host and manage software applications and allow their users to connect to and access applications over the global Internet.	67.4

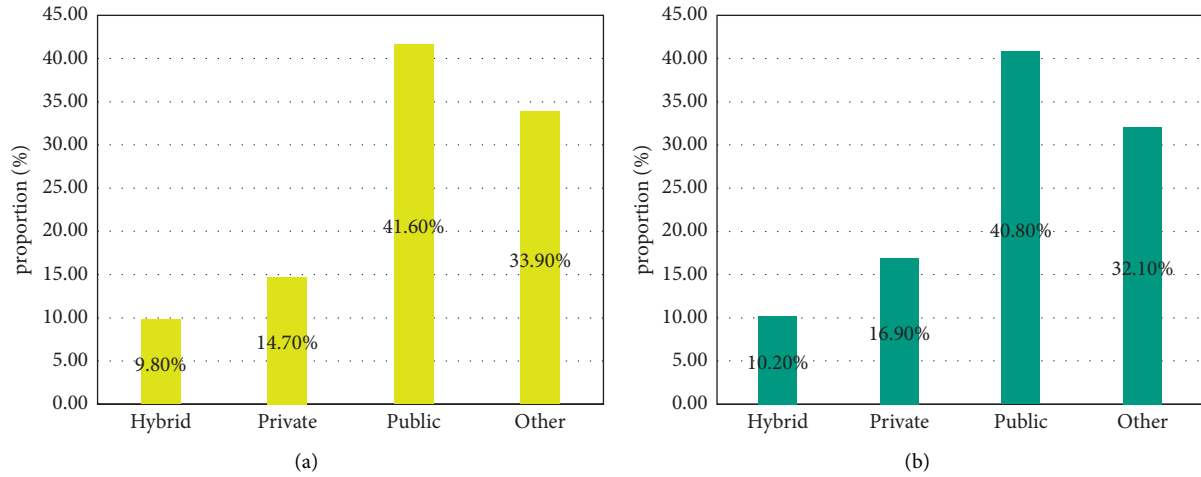


FIGURE 2: Domestic cloud service usage in the past two years. (a) 2019. (b) 2020.

apply for more resources when they are needed and can release them when they are no longer needed. From the user’s point of view, these resources are unlimited, and users only pay for the resources they use. The fifth is strong adaptability. The cloud computing system is an adaptive system, which can automatically balance the load and optimize the utilization of resources and can update resources adaptively with the changes of the data age. Users can also monitor resource usage [12–14].

- (4) *Classification.* According to the type, there are software (SaaS), platform (PaaS), and infrastructure (IaaS). There are private type, public type, mixed type, and community type according to the method. The specific content of the three service types is shown in Table 1, and the application proportion of the four service methods is shown in Figure 2 [15].
- (5) *Core Technology.* The core technologies of cloud computing mainly include programming technology and information security technology. Among them, the first six technologies are the most important [16].
- (6) *Application Areas.* Cloud computing technology is commonly used in today’s Internet services. The most common are online mailboxes (such as Google and Baidu) and online search engines. Users can search for the resources they need at any time on their mobile devices and share data resources through the network cloud. The same goes for online

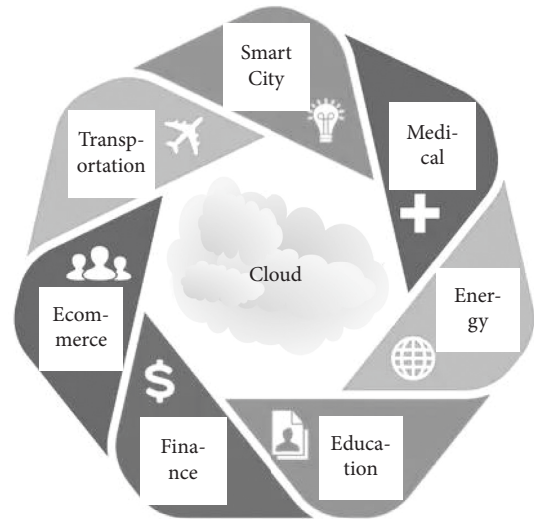


FIGURE 3: Application areas of cloud computing.

mailboxes, where sending and receiving e-mails used to be a tedious and time-consuming process. With the popularization of cloud computing and network technology, e-mail has become a part of social life. As long as it is in the network environment, real-time delivery of e-mails can be realized. At present, the main application areas can be divided into storage cloud (such as Microsoft, Google, and other large networks), financial cloud (such as Alibaba’s Alibaba



FIGURE 4: BD usage.

Cloud, Tencent Financial Cloud, and so on), education cloud, medical cloud, and so on. The details are shown in Figure 3.

- (7) *Problems Faced*. At present, several problems faced by cloud computing mainly include serious information leakage, users have no right to access, data systems are not comprehensive, and there is no sound legal protection. To solve these problems, we must first improve the legal system and publicize the awareness of legal security. Secondly, the access rights are used correctly, and the data system is comprehensively improved [17].

### 3.2. The Correlation of BD

- (1) *The Definition of BD*. Generally speaking, under different requirements, the required time processing range is different [18]. Figure 4 shows some important uses of BD.
- (2) BD are characterized by complex structure, huge quantity, diverse types, and low value density. At present, BD are changing from calculating some data to analyzing all data, from calculating microscopic results to discovering macro trends, and from exploring cause and effect to exploring information correlation.
- (3) *The Core Technology of BD*. The core technologies of BD include BD collection technology, storage technology, mining and analysis technology, and visualization technology. The details are shown in Figure 5.

- (4) The main application fields of BD include e-commerce industry, financial industry, biotechnology, smart government, education industry, transportation industry, medical industry, etc. The usage ratios of these main application areas are shown in Table 2. The data on the proportion of applications here are compiled from public information.

- (5) *The Association Rules of BD*. When this paper studies the correlation of BD, it first analyzes the attributes of BD. BD refer to a massive collection of unfiltered data. They can rationally express objective things in many aspects, such as words, symbols, letters, shapes, and so on, which are all data expressions in different forms. Data exist in data values and data structures, and large amounts of data form a complex network. In the network, the data value is its information entity, and the data structure can be regarded as the relationship between entities. The association between data can include time and space association, entity and virtual association, network level association, and so on [19–22].

The relationship between time and space refers to the attribute of describing data using space and time, which is beneficial to people's data mining. Time correlation can usually be divided into time points and time periods, and the expression of development and change of things needs to be represented by time attributes. Spatial correlation is mostly used in the visualization of geographic data, and it is more helpful for people to understand when the background is selected as a familiar map in the visualization based on

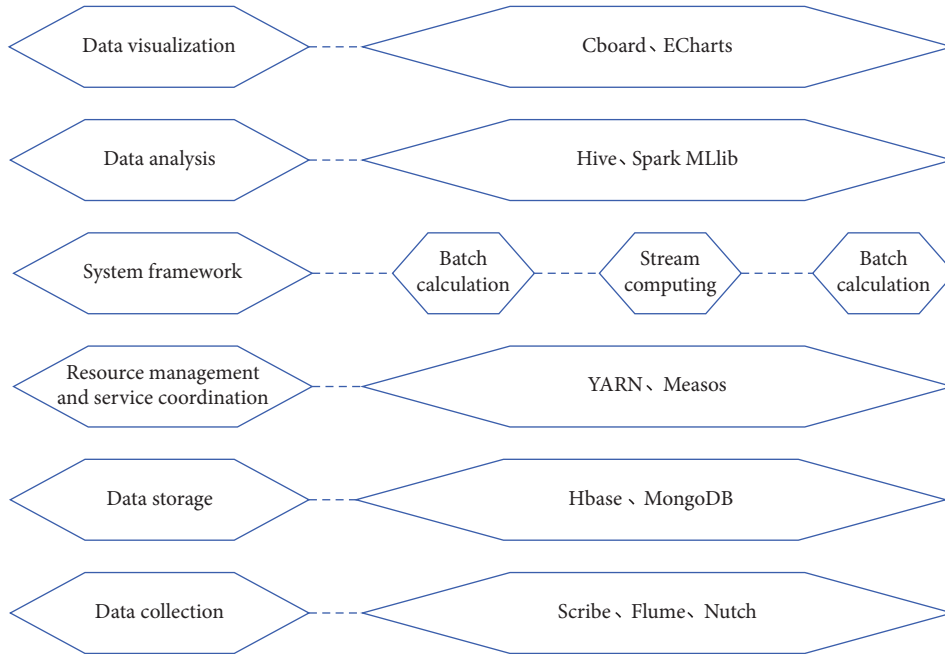


FIGURE 5: Core technologies of BD.

TABLE 2: The proportion of BD applications.

Type of enterprise	Proportion (%)
Government affairs	13
Internetwork	8
Medical and health	14
Education	8
Traffic	6
Finance	30
Electronic commerce	4
Others	17

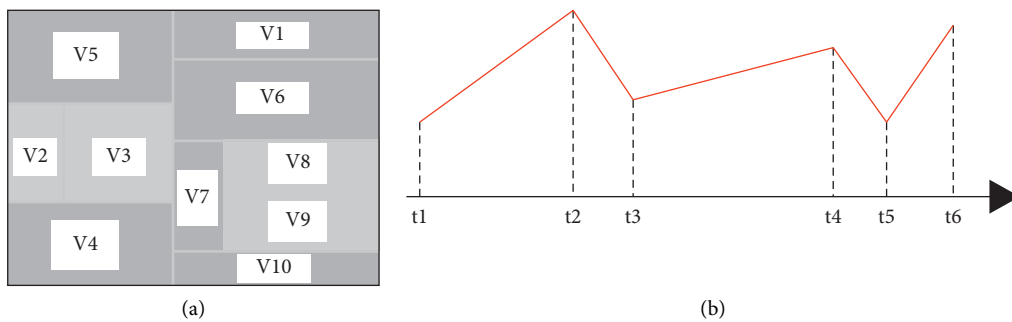


FIGURE 6: Spatiotemporal correlation of data. (a) Spatial correlation. (b) Time correlation.

geographic information. As shown in Figure 6, there is no BD correlation in time and space [23, 24].

Entity associations are often used in visualization to represent entities with different visual representations. Entity attributes in BD can usually be changes and combinations of three types of entity attributes: category attributes, interval attributes, and numerical attributes. The combination of virtual and real BD can more conveniently detect and control entities, as shown in Figure 7.

3.3. *BD Real-Time Controllable Algorithm.* The so-called real-time controllability of BD is to monitor and control these data anytime and anywhere at the terminal.

Taking the K-means algorithm as an example and assuming that the center of the data cluster is  $W$ , then

$$W_i = \left( \frac{1}{n_i} \sum i_{i1}, \frac{1}{n_i} \sum i_{i2}, \frac{1}{n_i} \sum i_{i2}, \dots, \frac{1}{n_i} \sum i_{ip} \right), \quad (1)$$



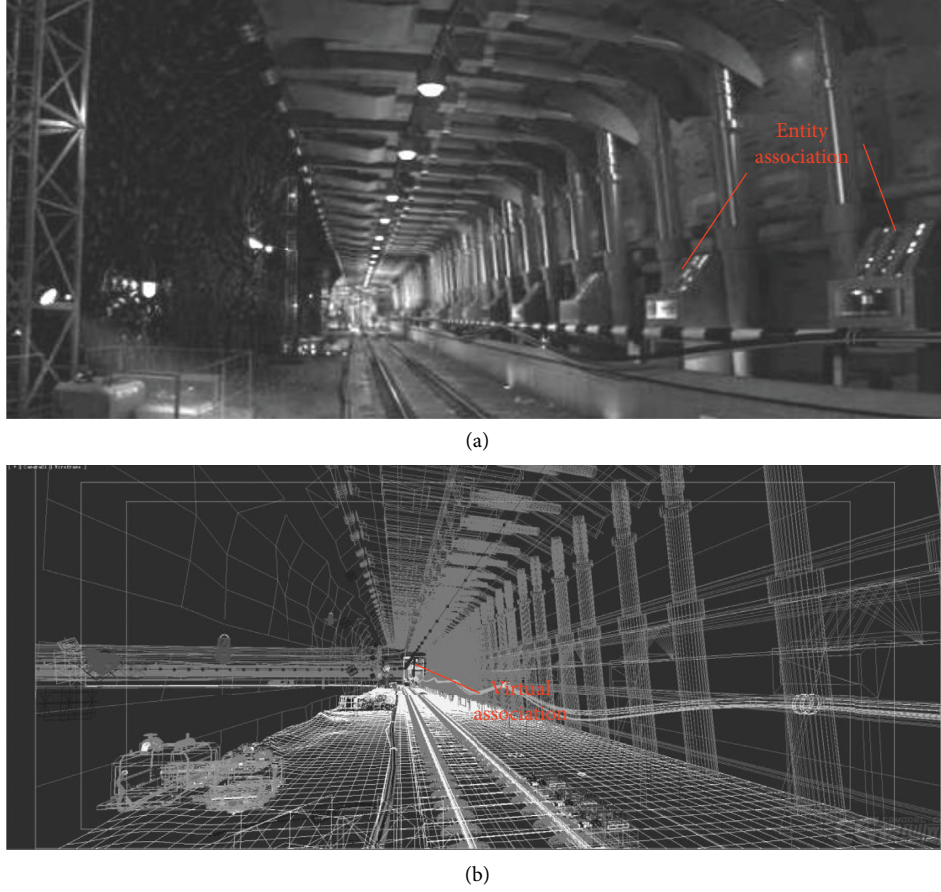


FIGURE 7: Virtual-real association of data. (a) Entity association. (b) Virtual association.

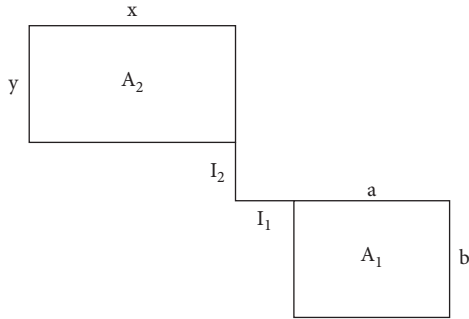


FIGURE 8: Side length and distance relationship of two minimum circumscribed matrices.

where  $n_i$  is the number of data in clusters and  $i_{ip}$  is the object of  $p$ -dimensional data.

The Euclidean distance of the two data is

$$d(I_i, I_j) = \sqrt{|i_{i1} - i_{j1}|^2 + |i_{i2} - i_{j2}|^2 + \dots + |i_{ip} - i_{jp}|^2}, \quad (2)$$

where  $i_{jp}$  is another object of  $P$ -dimensional data.

The average distance from all data points in the cluster to the center point is called intracluster similarity, which can be expressed as

TABLE 3: Serial-parallel algorithm simulation parameters.

Dataset size (kB)	Number of association rules
1220	14975
2484	14810
4960	14804
9960	14831
19921	14546
39866	14523

$$\text{inner} = \frac{1}{n} \sum_{I \in C_i} \sqrt{|i_1 - W_{i1}|^2 + |i_2 - W_{i2}|^2 + \dots + |i_p - W_{ip}|^2}, \quad (3)$$

where  $C_i$  represents the cluster, and the smaller the inner, the higher the similarity.

The minimum distance between clusters and cluster centers is called intercluster similarity, which can be expressed as

$$\begin{aligned} \text{ext} &= \min W_i - W_j \\ &= \min \sqrt{|W_{i1} - W_{j1}|^2 + |W_{i2} - W_{j2}|^2 + \dots + |W_{ip} - W_{jp}|^2}. \end{aligned} \quad (4)$$

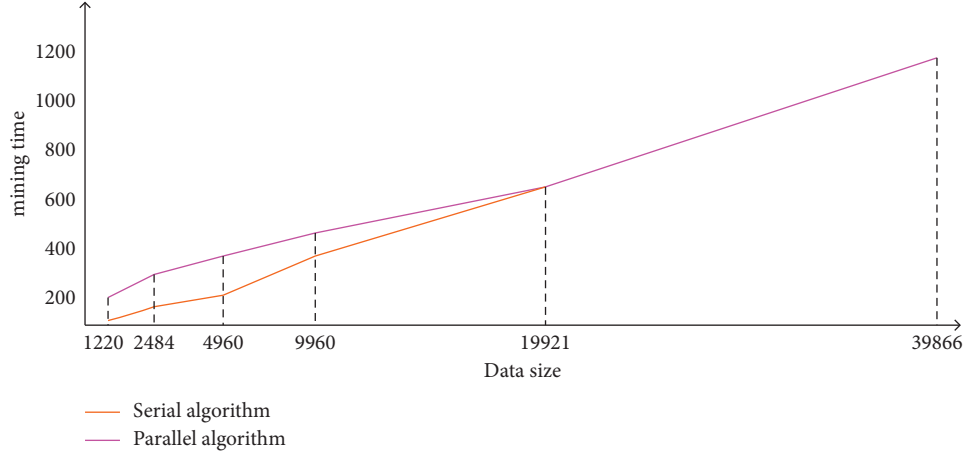


FIGURE 9: Comparison of serial and parallel algorithms.

The smaller the ext, the higher the similarity. On the contrary, it is smaller.

Taking the nearest neighbor algorithm as an example, suppose a dataset is  $U$ ,  $x_i$  represents the data in it, and the  $m$  attributes of the data  $x$  are  $x_{im}$ .

Then, there are

$$U = \{x_{ij}\}_{n \times m}. \quad (5)$$

Normalize to get

$$R = \{r_{ij}\}_{n \times m}. \quad (6)$$

According to the information entropy theory (information entropy is a measure of the uncertainty of information; the greater the uncertainty of the information, the greater the information entropy and the greater the value of the information), the expression of each attribute weight is

$$E_j = \frac{1}{\ln n} \sum_{i=1}^n r_{ij}^X \ln r_{ij}^X, \quad (7)$$

$$r_{ij}^X = \frac{r_{ij}}{\sum_{i=1}^n r_{ij}},$$

where  $r_{ij}^X$  represents the matrix weight; then, the weight of the  $j$ th attribute is

$$w_j = \frac{1 - E_j}{\sum_{i=1}^m E_i}. \quad (8)$$

The weighted distance for any two data in the dataset can be expressed as

$$D(x_i, x_j) = \sqrt{\sum_{p=1}^m w_p (x_{ip} - x_{jp})^2}. \quad (9)$$

Then, the sum of the distances of all the data in the set is

$$D_u = \sum_{i=1}^{n-1} \sum_{j=i+1}^n D(x_i, x_j). \quad (10)$$

TABLE 4: Simulation parameter settings.

Dataset size (kB)	Number of blocks
3800	50
12500	145
25000	300
33000	405
40000	536

Set the distance standard of adjacent data to  $\varepsilon$ ; then,

$$\varepsilon = D_u \times \frac{n(n-1)}{2}. \quad (11)$$

Set the neighborhood of the data to  $\delta_i$ ; then,

$$\delta_i = \{x \in U(x_i, x_j) < \varepsilon\}. \quad (12)$$

In order to determine whether the data in a dataset or function are in the neighborhood content, define

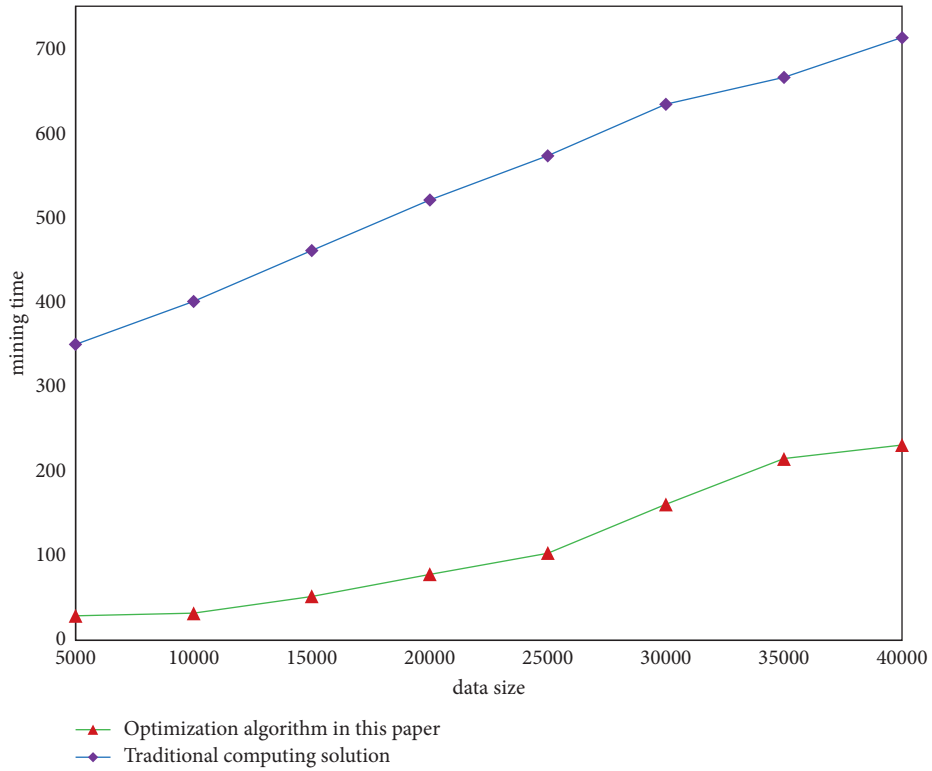
$$IS(x_i, x_j) = \begin{cases} 0, & x_i \notin \delta(x_j), \\ 1, & x_i \in \delta(x_j). \end{cases} \quad (13)$$

Assuming that the probability of data  $x_j$  appearing in the neighborhood space of the rest of the data is  $P_C$ , then

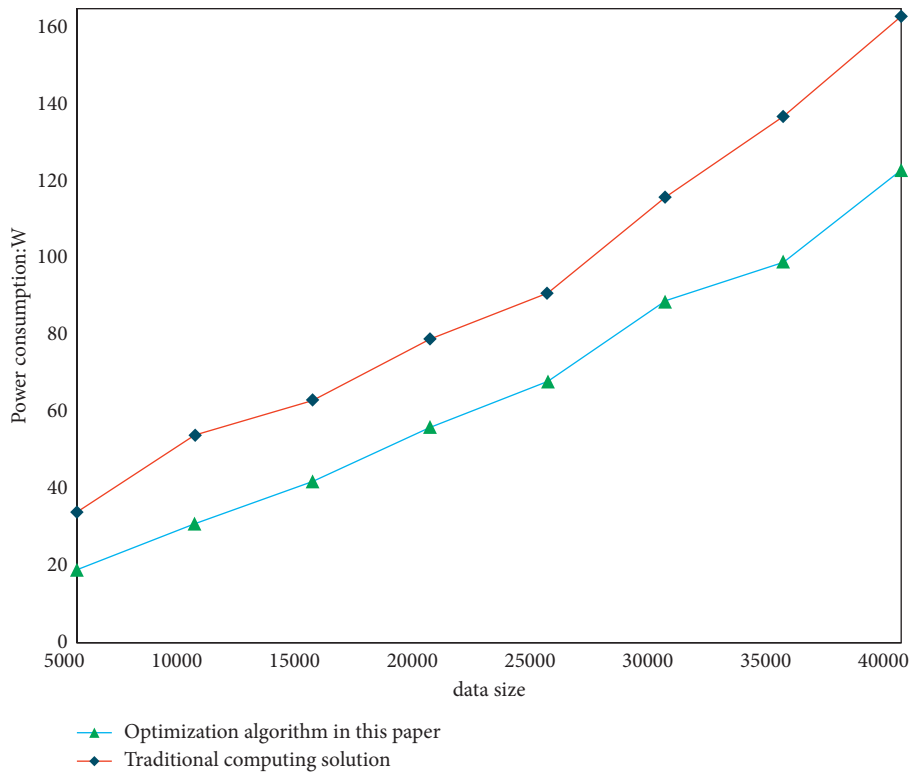
$$P_C = \frac{\sum_{j=1}^n IS(x_i, x_j)}{n}. \quad (14)$$

#### 4. Optimization Algorithm Test of Dynamic $K$ Value

**4.1. Real-Time Controllable Optimization Algorithm for Correlated BD.** This paper combines the proximity algorithm to determine the dynamic  $K$  value of the  $K$ -means algorithm, so as to construct the network model of the correlation data. The construction steps are as follows: first, the entire dataset is to be clustered, and the algorithm is applied to cluster it. The obtained clustering results are adjusted. Then, using each of the clusters as a dataset,



(a)



(b)

FIGURE 10: Comparison between optimization algorithms and traditional cloud computing. (a) Comparison of mining time. (b) Power consumption comparison.



perform corresponding clustering and adjustment. This step is performed iteratively on each resulting subcluster until the data network is constructed [25–27].

Assuming that the network space composed of one or more data has the two minimum circumscribed matrices  $A_1$  and  $A_2$ , then the relationship between the side lengths and distances between the two can be shown in Figure 8.

Then, the area of the new space minimum circumscribed matrix obtained by the two is

$$S = (a + l_2 + x) \times (b + l_1 + y). \quad (15)$$

When the center points of  $A_1$  and  $A_2$  remain unchanged, the formula can be expanded to get

$$S = xy + (b + l_1)x + (a + l_2)y + (a + l_2)(b + l_1). \quad (16)$$

It can be seen from this that the area of the newly constructed minimum circumscribed rectangle is related to the area and perimeter of the smallest circumscribed rectangle that composes it.

According to the previous proximity algorithm and K-means algorithm, the distance between the two circumscribed matrices can be obtained:

$$D(A_1, A_2) = e^{(S/1+S)(C/1+C)} \sqrt{\sum_{j=1}^M W_j (c_{1j} - c_{2j})^2}, \quad (17)$$

where  $S$  is the area and  $C$  is the perimeter.

In this paper, the parameter  $\lambda$  is introduced on the basis of formula (11) for data optimization, namely:

$$\varepsilon = \lambda * D_u \times \frac{n(n-1)}{2}. \quad (18)$$

In order to optimize the clustering effect between each data, the value of the number of clusters of any child node data of the intermediate node is

$$k' = \left\lfloor \frac{\sum_{i=1}^n k_i}{n} \right\rfloor. \quad (19)$$

**4.2. Optimization Algorithm Simulation Test.** Firstly, the serial algorithm is simulated, and the introduced improved adjacent K-means algorithm is run under Win7, and the specific performance of the Oracle database (Oracle database system is a popular relational database management system in the world; the system has good portability, convenient use and powerful functions, and is suitable for all kinds of large, medium, and small computer environments; it is an efficient, reliable, and high-throughput database solution) for mining data is recorded. The serial and parallel algorithms mine the same dataset and keep increasing the size of the dataset, and both algorithms set the same threshold, different datasets must be tested, a total of 6 times. The simulation parameters are shown in Table 3, and the experimental results are shown in Figure 9.

From Figure 9, the experimental results show that as the scale of data processed by the serial-parallel algorithm gradually

increases, the memory consumed by the serial algorithm gradually increases. When the data size reaches about 39 M, the serial algorithm will report insufficient memory and cannot complete association rule mining, while the improved parallel algorithm can complete the task. But since the Hadoop platform runs in pseudo-parallel mode, the performance of a single node is the same as that of a serial algorithm. Also, the parallel algorithm requires interaction between tasks, so when the data size is small, the parallel algorithm takes much more time than the serial algorithm. As the data size increases, this gap gradually decreases. It can be seen that parallel algorithms are more advantageous when dealing with large datasets.

Then, the optimization algorithm and the traditional algorithm are simulated and tested in the fully distributed mode by using the parallel method. The same thresholds are also set for the two algorithms:  $\delta = 0.2$ ,  $\delta = 0.6$ , and  $\delta = 0.4$ , and different datasets are tested for 6 times. The simulation parameters are shown in Table 4, and the results are shown in Figure 10.

It can be seen from Figure 10 that compared with the traditional cloud computing, the optimization algorithm in this paper has significantly lower mining time for data correlation. When the amount of data reaches 40 MB, traditional cloud computing takes about 700 s, while the optimized algorithm in this paper only takes about 200 s, and the traditional cloud computing time is about 3.5 times that of this paper. In addition, in terms of power consumption, the power consumption of the algorithm optimized in this paper is always lower than that of traditional cloud computing, and the difference between the two is about 20 W.

## 5. Discussion

The work done in this paper is purely aimed at the algorithm, but the research of the algorithm is to maximize its effect. If only the dynamic algorithm is used to replace the static algorithm, with the emergence of high-speed and multi-core processors, this research may not be of great significance. Therefore, if the research in this paper can be applied to the protocol, it may have greater practicality.

When the paper improves the algorithm simulation, because the dataset is still small, it is expected that the dataset will be enlarged in the follow-up work and then simulated to show the advantages of the improved algorithm. At present, most of the data operations in the system rely on manual arrangement of data information, which cannot really be better used by the majority of users. It cannot design and implement fully automated operations, provide users with a good interface, and cannot directly obtain data mining results.

With the development of society, BD has increasingly penetrated into human's daily work, life, and various safety-critical application environments. In particular, the rapid development of mobile electronic products and mobile Internet applications has brought new research topics for energy saving and reliability optimization of real-time systems. To this end, continuous and in-depth research and practical work are required to adapt and optimize real-time controllable algorithms.

## 6. Conclusion

In the research on the real-time controllable optimization algorithm of BD, this paper firstly explains the background meaning of the era of BD in the abstract part. Then, this paper explains the purpose of this research and the theoretical algorithm used. Then, this paper explains the research value of BD background and cloud computing in the reference part. This article exemplifies the related research of many scholars in the real-time controllable optimization technology of BD and analyzes their research results and the shortcomings of their research.

In the theoretical research part, this paper first introduces cloud computing. It includes its source, definition, characteristics, classification, core technology, application fields, and the challenges cloud computing is facing. Then, this paper introduces the concept, characteristics, core technology, main sources, and association rules of BD and explains with the help of charts.

Finally, in the algorithm design of optimization and upgrading, this paper proposes a combination of proximity algorithm and K-means algorithm. It is optimized and upgraded through parallel computing. After several simulation tests of performance parameters, the results obtained are compared with traditional cloud computing, and the advantages of the scheme proposed in this paper are obtained.

## Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## Conflicts of Interest

The authors declare that they have no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Acknowledgments

This study was supported by the Zhejiang Provincial Natural Science Foundation of China under grant no. LGG20F020013 (Rutao Li). This study was also supported by the Excellent Talent Foundation of China West Normal University (no. 17YC497) (Zaiyi Pu).

## References

- [1] V. Puig, C. Ocampo-Martínez, R. Pérez, G. Cembrano, J. Quevedo, and T. Escobet, "Advances in industrial control] real-time monitoring and operational control of drinking-water systems," *BD Analytics and Knowledge Discovery Applied to Automatic Meter Readers*, Springer, Berlin, Germany, 2017.
- [2] C. Qian, A. Liu, R. Huang et al., "Quality control of marine BD—a case study of real-time observation station data in Qingdao," *Journal of Oceanology and Limnology*, vol. 37, no. 06, pp. 195–205, 2019.
- [3] D. Puthal, S. Nepal, R. Ranjan, and J. Chen, "DLSeF: a dynamic key-length-based efficient real-time security verification model for bd stream," *ACM Transactions on Embedded Computing Systems*, vol. 16, no. 2, pp. 1–24, 2017.
- [4] S. Ren, X. Zhao, B. Huang, Z. Wang, and X. Song, "A framework for shopfloor material delivery based on real-time manufacturing big data," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 3, pp. 1093–1108, 2019.
- [5] N. N. Bao, Y. Huang, B. J. Xiao et al., "Design of real-time feedback control of vertical growth rate on EAST," *Nuclear Science and Techniques*, vol. 32, no. 7, pp. 1–9, 2021.
- [6] R. Janzen and S. Mann, "Extreme-dynamic-range sensing: real-time adaptation to extreme signals," *IEEE multimedia*, vol. 24, no. 2, pp. 30–42, 2017.
- [7] Y. Wang, J. Li, and H. H. Wang, "Cluster and cloud computing framework for scientific metrology in flow control," *Cluster Computing*, vol. 22, no. 1, pp. 1–10, 2017.
- [8] L. Ogiela, M. R. Ogiela, and H. Ko, "Intelligent data management and security in cloud computing," *Sensors*, vol. 20, no. 12, p. 3458, 2020.
- [9] S. Namasudra and P. Roy, "Ppbac: popularity based access control model for cloud computing," *Journal of Organizational and End User Computing*, vol. 30, no. 4, pp. 14–31, 2018.
- [10] Z. Cao, L. Jin, C. Wan, Y. Song, Y. Zhang, and X. Wang, "Optimal cloud computing resource allocation for demand side management," *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1943–1955, 2017.
- [11] J. Du, L. Zhao, J. Feng, and X. Chu, "Computation offloading and resource allocation in mixed fog/cloud computing systems with min-max fairness guarantee," *IEEE Transactions on Communications*, vol. 66, no. 4, pp. 1594–1608, 2018.
- [12] T. Hirai, H. Masuyama, S. Kasahara, and Y. Takahashi, "Performance analysis of large-scale parallel-distributed processing with backup tasks for cloud computing," *Journal of Industrial and Management Optimization*, vol. 10, no. 1, pp. 113–129, 2017.
- [13] X. Liu, X. Zhang, W. Li, and X. Zhang, "Swarm optimization algorithms applied to multi-resource fair allocation in heterogeneous cloud computing systems," *Computing*, vol. 99, no. 12, pp. 1–25, 2017.
- [14] Y. Zhang, X. Xiao, L. X. Yang, Y. Xiang, and S. Zhong, "Secure and efficient outsourcing of PCA-based face recognition," *IEEE Transactions on Information Forensics and Security*, vol. 15, 2019.
- [15] W. Wei, X. Fan, H. Song, X. Fan, and J. Yang, "Imperfect information dynamic stackelberg game based resource allocation using hidden markov for cloud computing," *IEEE Transactions on Services Computing*, vol. 11, no. 99, pp. 78–89, 2018.
- [16] S. Wang, K. Liang, J. K. Liu, J. Chen, J. Yu, and W. Xie, "Attribute-based data sharing scheme revisited in cloud computing," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 8, pp. 1661–1673, 2017.
- [17] A. Barsoum and M. A. Hasan, "Provable multicopy dynamic data possession in cloud computing systems," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 3, pp. 485–497, 2017.
- [18] J. L. Tsai and N. W. Lo, "A privacy-aware authentication scheme for distributed mobile cloud computing services," *IEEE Systems Journal*, vol. 9, no. 3, pp. 805–815, 2017.
- [19] L. Ping, J. Li, Z. Huang et al., "Multi-key privacy-preserving deep learning in cloud computing," *Future Generation Computer Systems*, vol. 74, pp. 76–85, 2017.
- [20] W. Tang and W. Feng, "Parallel map projection of vector-based big spatial data: c," *Computers, Environment and Urban Systems*, vol. 61, no. Part B, pp. 187–197, 2017.

- [21] Z. Lv, S. Lv, and J. Li, "Clinical characteristics and analysis of risk factors for disease progression of COVID-19: a retrospective Cohort study," *International Journal of Biological Sciences*, vol. 17, no. 1, pp. 1–7, 2021.
- [22] M. Lee, L. Mesicek, and K. Bae, "AI advisor platform for disaster response based on big data," *Concurrency and Computation: Practice and Experience*, vol. 12, Article ID e6215, 2021.
- [23] T. G. Rodrigues, K. Suto, H. Nishiyama, and N. Kato, "Hybrid method for minimizing service delay in edge cloud computing through VM migration and transmission power control," *IEEE Transactions on Computers*, vol. 66, no. 5, pp. 810–819, 2017.
- [24] M. Abdel-Basset, M. Mohamed, and V. Chang, "NMCDA: a framework for evaluating cloud computing services," *Future Generation Computer Systems*, vol. 86, pp. 12–29, 2018.
- [25] H. Cai, B. Xu, L. Jiang, and A. V. Vasilakos, "IoT-based big data storage systems in cloud computing: perspectives and challenges," *IEEE Internet of Things Journal*, vol. 4, no. 1, pp. 75–87, 2017.
- [26] Z. Lv, "The security of Internet of drones," *Computer Communications*, vol. 148, pp. 208–214, 2019.
- [27] Y. Guo, H. Mi, and M. Habibi, "Electromechanical energy absorption, resonance frequency, and low-velocity impact analysis of the piezoelectric doubly curved system," *Mechanical Systems and Signal Processing*, vol. 157, Article ID 107723, 2021.