Research Article

Agricultural Information Management Data Model Construction Based on Cloud Computing and Semantic Technology

Chunhang Wang and Zhe An

Tangshan Vocational and Technical College, Hebei, Tangshan, China

Correspondence should be addressed to Zhe An; 1400440135@xs.hnit.edu.cn

Received 16 December 2021; Revised 28 February 2022; Accepted 10 March 2022; Published 11 April 2022

Academic Editor: Qiangyi Li

Copyright © 2022 Chunhang Wang and Zhe An. 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.

In order to improve the data retrieval and mining ability of agricultural information management system, an agricultural information management data model based on cloud computing and semantic technology was proposed. Fuzzy C-means algorithm is used for adaptive fusion and clustering of semantic association features of distributed data in large-scale information management systems; feature compressors are used to reduce the storage space dimension of large-scale information management system to improve the target data mining capability and adaptive scheduling capability of information management system. The registration rate improves than the traditional method by 12.46% and computational cost by 23.76%. This method has higher accuracy in storage data mining and retrieval for large information management systems and superior performance than the traditional method. The introduction of the concept and technical methods of related data is the best practice to realize the fine disclosure, standardized description, semantic organization, and in-depth integration of massive agricultural science and technology information resources, which will play an important role in improving the visibility, visibility, and accessibility of agricultural information resources. It is innovative in the construction of the multidimensional semantic-related model and field knowledge service system driven by related data.

1. Introduction

As the basic industry of China’s economy, agriculture’s development level has a profound impact on the quality of life of Chinese people [1]. At present, China’s agriculture is in the stage of developing to informationized agriculture, and the foundation of agricultural informationization is agricultural spatial information [2]. The construction of cloud portal for agricultural spatial information management is studied, and the research background of this study is expounded from the perspectives of theory, technology, and application. Since the 21st century, with the geometric growth of data, cloud computing technology has become one of the hotspots in the field of Internet high-tech research [3]. Cloud computing belongs to distributed computing, while cloud GIS, based on the basic theory and technology of cloud computing, realizes the storage, management, processing, and analysis of geo-spatial information, thus changing the traditional GIS structure and application mode [4]. The cloud GIS technology can realize the storage and processing of massive agricultural spatial data and provide the platform with more efficient agricultural spatial data computing and agricultural space application analysis capabilities; it greatly improves the management and sharing efficiency of agricultural spatial information and is a powerful technical support for the construction of agricultural information management platform in the future. With the continuous development of modern society, human beings have entered an era of information explosion. With the increasing number of application users, increasing demand for computing power, and increasing security requirements, enterprises have to increase their investment in hardware equipment to meet the growing needs. At the same time, in order to maintain the safety and reliability of the hardware equipment, the system operation and maintenance requirements are also increasing day by day. More
importantly, this growth model is conducted exponentially. SMEs and individuals are not only discouraged from the cost of software development but also difficult for large companies. Therefore, cloud computing technology, which is easy to use and easy to obtain resources, has gradually entered people’s sight.

In the network information environment, the era of “BigData” has come, and the “fourth paradigm” of scientific research characterized by data-intensive computing is rising. Existing information organization and service means are difficult to adapt to the development and change of agricultural scientific research environment and scientific research methods; due to the lack of orderly integration, organization, standard control, mining, and utilization of massive agricultural scientific and technological information resources, agricultural researchers are in the “information ocean;” however, it is often faced with the dilemma of “excessive information and lack of knowledge and medical treatment,” making it increasingly difficult for people to find and obtain information accurately, quickly, and comprehensively [4]. The large information management system is built in the database setting based on the calculation and information retrieval algorithm and combined with the information scheduling technology, conducts the overall design framework and database development of the information management system, improves the information retrieval and integrated information processing ability, and studies the data mining technology of the large information management system, which is of great significance to improve the distributed retrieval and scheduling ability of the information management system. In large-scale information management systems and cloud database media, effective data mining methods are needed to access and schedule distributed data in large-scale information management systems to improve data utilization efficiency and resource-sharing degree; in traditional methods, semantic feature mining of large-scale information management system is based on statistical analysis and feature extraction of distributed data bit information flow in large-scale information management system; typical data mining methods include association rule mining, semantic ontology model mining, spectrum feature extraction, and fuzzy K-means mining; the fuzzy clustering center of big data feature distribution in large-scale information management system is constructed, combined with spectral feature extraction and semantic analysis method, data mining is carried out, and some research results have been obtained [5]. The construction of agricultural information management data model can significantly improve the visibility, visibility, accessibility and scientific value of agricultural science and technology information resources will play an important role and has an important research value and practical significance for promoting agricultural scientific research and innovation.

According to the above problems, this study proposes a data mining technique for a large information management system based on semantic correlation feature extraction. This section first briefly analyzes some new trends and characteristics of the environment of agricultural scientific research innovation and information needs of scientific research users under the background of data-intensive computing and the rise of the “fourth paradigm” of scientific data and the arrival of the era of “big data.” It is proposed to introduce new service concepts and technical means and other related data to innovate service content and functions and improve the role and contribution of knowledge service to scientific and technological innovation [3, 6].

2. Research Methods

2.1. Analysis of Data Storage Structure and Characteristics of Large-Scale Information Management System

2.1.1. Data Storage Structure Model. In the cloud computing and semantic technology environment, especially from the traditional file network (Web of Document) to the structured and semantic-rich data network (Web of Data) and traditional knowledge organization systems (such as narrative lists, topic vocabulary, and taxonomy) have regained their attention and attention and also play an increasingly important role in organizational management, mining, and analysis and development and utilization of massive information resources. The database storage basic block of large-scale information management system is \( m_{i,j} (i \leq n, j \leq k) \); combined with the closed frequent item set fusion method, the target data semantic training set of large-scale information management satisfies \( t_i, G_i \subseteq G_j \subseteq Y_i \iff Y_j \); let \( A = \{a_1, a_2, \ldots, a_n\} \) be the distributed amplitude of the data structure; the probability of occurrence of target number in large-scale information management system is represented by \( p(i, j), p(i, j) = \lim_{\eta \rightarrow 0} p(a_i = i, b_j = j) \); the interval unit \( i \in [0,n], j \in [0,n] \) of big data sampling is obtained. The time window of data mining is expressed as:

\[
Y_N = X_n + \eta, \tag{1}
\]

where \( Y_N \) represents the characteristic component of data association rules, while \( \eta \) represents the mutual information, the data storage element \( \alpha(i, j) \) is input into the storage link layer of the large-scale information management system, the steady-state characteristic quantity of information scheduling is \( \delta_n \), and the transfer function of distributed storage of large-scale information management system can be obtained as follows:

\[
\alpha(i, j) = \begin{cases} 
0, & i = 0 \text{ or } j = 0, \\
1, & n - j < i, i \geq j, \\
1, & n - j < i, j \geq i, \\
\frac{1 - n_j B_j}{B_i}, & n \geq j, i \geq j, \\
\frac{1 - n_i B_i}{B_j}, & n \geq j, j \geq i.
\end{cases} \tag{2}
\]

The model of associated rule distribution for large information management systems (Figure 1).
2.1.2. Feature Analysis of Semantic Correlation Dimension. The information management system is optimized and reorganized by combining the feature recombination method of big data information flow; the feature quantity of semantic association dimension of information management distributed data is extracted from the topology structure of the reorganized information management system. The decision statistics of association rule information fusion are as follows:

\[ Q_s = \frac{\langle (x_{n-d} - \bar{x}) (x_{n-d} - \bar{x}) \rangle}{\langle x(n) \rangle^2} \]  

(3)

where \( x_n \) represents the big data sampling sequence in the large-scale information management system, \( d \) represents the data sampling interval, \( D = 2d \), \( \bar{x} \) is the mean, and \( \langle x(n) \rangle \) represents the weight of the evaluation index, which is

\[ \langle x(n) \rangle = \frac{1}{N} \sum_{i=1}^{N} x(n). \]  

(4)

Combining with the binary semantic weight analysis method, the binary semantic decision matrix is obtained as follows:

\[ R = \begin{bmatrix}
  r(V_1, V_1) & \cdots & r(V_1, V_{k-1}) \\
  \vdots & \ddots & \vdots \\
  r(V_{k-1}, V_1) & \cdots & r(V_{k-1}, V_{k-1})
\end{bmatrix}. \]  

(5)

By constructing the evaluation decision matrix and combining with the adaptive information fusion method, the decision criteria of data mining of large-scale information management system are obtained as follows:

\[ p(Q_s) = \frac{1}{\sqrt{2\pi} \sigma_s} \exp \left[ -\frac{(Q_s - \langle Q_s \rangle)^2}{2\sigma_s^2} \right], \]  

\[ \int_{-\infty}^{\infty} p(Q_s) dQ_s = 1. \]  

(6)

The ontology clustering center was built according to each evaluation object and the optimal scheme, and the absolute error of \( Q_8 \) and \( \langle Q_s \rangle \) was calculated; the decision threshold value of data mining of large-scale information management system is \( Q_{cr} \), when

\[ p(\{Q_{cr} < Q_s\} \leq 0.05. \]  

(7)

At the confidence level of 95%, the accurate distribution probability of semantic association feature mining can be obtained as follows:

\[ 0.025 = \int_{-\infty}^{z_1} p(Q_s) dQ_s = 1 - \int_{-\infty}^{z_1} p(Q_s) dQ_s. \]  

(8)

Type: \( z_1 = -z_{1.025} \), when \( S \geq 2.00 \), the target data mining of the information management system is aggregated with a probability of 95% in the semantic ontology model, and the semantic association feature extraction results are

\[ \max U = u_1 + u_2 + \cdots + u_n, \]  

\[ u_i = p_i, \]  

\[ \sum_i p_i = 1, \quad 0 < p_i < 1, \]  

\[ \frac{p_i (1 - p_i)}{w_i} = \frac{p_i (1 - p_i)}{w_n} = \frac{1}{K}. \]  

(9)

2.2. Optimization of Data Mining Algorithm

2.2.1. Semantic Relevance Feature Extraction. The integrated scheduling and data mining of the information management system are carried out with the semantic association feature quantity as the training sample set; fuzzy C-means algorithm is used for adaptive fusion and clustering of semantic association features of distributed data in large-scale information management systems; distribution data distribution sample attribute set is \( i \in S_p \), and its equivalent similarity transitive mapping relationship satisfies

\[ a_i^T Q a_i = \sum_{j=1}^{n} a_i a_j Q_{ij} \geq 0. \]  

(10)

The semantic mapping relationship between ontology is established, for the target data containing \( n \) samples, where the correlation characteristic distribution vector of sample \( x_i, i = 1, 2, \ldots, n \), is

\[ s(t) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} a_{mn} g_{mn}(t) + n(t), \]  

(11)

where \( a_{mn} \) refers to the envelope amplitude of potentially useful information of distributed data in the information management system to be large scale, \( g_{mn}(t) \) is the statistical average value of data, and \( n(t) \) is an interference term, and each concept in the ontology model can also be called a class. Semantically, when \( \Phi: M \rightarrow R^{d+1} \), an embedded high-dimensional phase space exists for \( \Phi(z) = (h(z), h(\phi_1(z)), \ldots, h(\phi_d(z))^2); \) two groups of very similar ontology fragment data sample sequence \( \{x(t_0 + i\Delta t)\}, i = 0, 1, \ldots, N - 1 \), are obtained; thus,
The semantic mapping relationship between ontology of information management system is obtained:

\[ x_n = a_0 + \sum_{i=1}^{M_{AR}} a_i x_{n-1} + \sum_{j=0}^{M_{LA}} b_j \eta_{n-1}, \quad (12) \]

where \( a_0 \) represents the information flow vector of distributed data input to the large-scale information management system, \( a_i \) is the embedding dimension of phase space, \( M_{AR} \) is the correlation order of semantic information, and \( \eta_{n-j} \) is the data sampling interval. In the topology structure of reorganized information management system, the feature quantity of semantic association dimension of information management big data samples is calculated, and information scheduling and association rules mining are carried out by combining data clustering algorithm [7, 8].

### 2.2.2 Information Management System Target Data Mining

Considering two groups of ontology fragments with the similar structure, the objective function of semantic association feature fusion of distributed data in large-scale information management system is constructed:

\[
X_p(u) = \begin{cases} 
\|x_k - V\|^2, & x(t) = 0 \quad \text{for} \quad t \geq 0 \\
\|\int_{-\infty}^{\infty} x(t)e^{j2\pi \cot \alpha/2} \sin \alpha \frac{2/2 \cot \alpha}{\alpha} \ dt, & x(t) \neq 0 \quad \text{for} \quad t < 0 
\end{cases}
\]

where \( x(t) \) is the ontology segment in the distributed data information of the massive large-scale information management system that contains the target attribute set of mining. \( P \) is the measure distance, expressed as equation (14), by Euclidean distance:

\[ p = \|x_k - V\|^2. \quad (14) \]

According to the semantic mapping SM-Context, the output information management target sample set meets the following requirements:

\[ \sum_{j=1}^{n} \mu_{jk} = 1, \quad k = 1, 2, \ldots, n. \quad (15) \]

Combined with the context feature optimization method of the concept, the target data scheduling of the information management system is carried out, and the optimal output solution is obtained:

\[ F_j = \frac{n}{k=1} X_{kj}, \quad (16) \]

\[ Q_j = \frac{n}{k=1} \left(X_{kj}\right)^2. \]

The semantic components contained in the concept are clarified, and then, the feature quantity of semantic association dimension of distributed data in large-scale information management system is extracted, and the iterative formula of feature extraction is obtained:

\[ x_i^{(k+1)} = (1 - \omega) x_i^{(k)} + \left(b_j - \sum_{j=1}^{n} a_{ij} x_j^{(k)} - \sum_{j=1}^{n} a_{ij} x_j^{(k)} \right), \quad i = 1, 2, \ldots, n; \quad k = 1, 2, \ldots, n. \quad (17) \]

where \( \omega \) is the initial weight of semantic mapping, \( a_{ij} \) is the joint characteristic quantity of the disjunctive relation, and \( x_i^{(k)} \) represents the difference between the semantics of a concept and the elements around it. The semantic association feature input is used into the fuzzy \( C \)-means classifier, and the feature compressor is used to improve the information mining ability and the adaptive scheduling ability of the information management system.
2.3. Cloud Computing Technology

2.3.1. Features of Cloud Computing. Since 2007, cloud computing has gradually appeared in people's field of vision. With the rapid development of computer technology, cloud computing has also begun to shift from academic theoretical research to practical applications and has brought convenience to people's lives. Compared with other computing modes such as traditional grid, cloud computing, as a new computing mode, has the following characteristics:

1. Virtualization: cloud computing uses virtualization technology to encapsulate hardware and software resources into services, and users can access and use these services on demand through the Internet. These services execute somewhere deep in the "cloud," where the user does not know and need not know where they are running. Users only need a cloud terminal device, such as a laptop, to easily obtain services in the cloud, even including powerful computing tasks.

2. It can be dynamically expanded and has the ability to provide functionality quickly and elastically to scale and to quickly add or release resources to scale down. For cloud users, the resources available for rent are virtually endless, and any number of purchases can be made at any moment.

3. Universality: the application of cloud computing has nothing to do with the physical platform. All operations are unified on the virtual platform built by the cloud, and various applications can be run on its platform.

4. On-demand deployment: because the "cloud" uses virtualization technology, the deployment of its platform is very flexible. Cloud computing can change capacity according to user needs and deploy applications into virtual computing pools. When a user submits a task, virtual machine resources are deployed according to the actual number of user tasks.

5. Pay on demand: charges for features use pay-per-use services or use an advertising-based charging model to increase cloud computing resource usage. For example, storage, computing, and bandwidth resources are measured and charged monthly based on the number of users used. An enterprise's private cloud may increase the cost of the enterprise, and the internal personnel of the enterprise may use physical currency, or the private cloud may be used for free.

6. High reliability: the user's services and computing will not be affected by the failure of a certain machine in the cloud. Data redundancy technology, virtual machine dynamic migration technology, etc., are used to ensure the smooth completion of user tasks.

7. High cost performance: cloud data centers are composed of a large number of cheap server clusters. The configuration of a single server may be low, but after they are unified to form a huge computing resource pool, its computing power will far exceed that of a mainframe.

2.3.2. Classification of Cloud Computing. At present, there are many classifications of cloud computing in the industry, but there are two generally recognized classification methods, namely, by service type.

Classification and Classification by Service:

1. Classification by Service Type. Cloud computing can provide users with a variety of services. By using these services, users can obtain corresponding resources. According to the type of service, it can be roughly divided into three categories from the bottom up, namely, Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Application Software as a Service (SaaS).

2. Classification by Service Mode. As an emerging computing model, although cloud computing has many advantages that existing computing models do not have, there are still some problems. The first is the issue of security. For companies with high security requirements, such as banks, operators, and the military, once security is compromised, the consequences will be disastrous. The second is the issue of supervision. Some enterprises have very strict internal management and operation and maintenance systems and do not want to be grasped and interfered with by related industries outside the company. Although cloud computing can provide protection for enterprises and users through security isolation measures, it still cannot meet the needs of all users. Therefore, in response to these problems, cloud computing is divided into public cloud, private cloud, community cloud, and hybrid cloud according to the relationship between service providers and users.

3. Result Analysis

The construction of the large-scale information management system is based on the database design and information retrieval algorithm, combined with the information feature extraction and optimization technology, which carries out the overall design framework and database development of the information management system and improves the information retrieval and integrated information processing ability. The design of a platform is to design platform objectives, specify design rules, and conduct actual platform construction according to the design rules. Only in the process of building the platform, strictly observing the specified design rules and improving the functions according to the design rules can the platform that meets the needs and purpose be made. This study designs the following rules for the built agricultural space information management cloud portal. The distributed data in the information management system includes 32 concept sets and 80 attribute sets and contains 100 semantic information instance sets [9]. The initial position of semantic mapping (0, 0, 25), the correlation coefficient is 0.23, the inertia weight acceleration factor of information management system data retrieval is 0.56, and the initial sampling frequency of sample data is $f_1 = 1.46Hz$, terminating the sampling frequency $f_2 = 2.12Hz$, the simulation duration is 200 ms, the management data of the
The information management system contains 32 concepts, 62 attributes, and 100 instances, and the mapping of the semantic relationship is shown in Table 1.

According to the semantic ontology association mapping in Table 1, the semantic mapping results between the reference ontology and the first 10 ontologies are given, as shown in Table 2.

The data mining of the information mining system obtains the time-domain and frequency-domain waveforms of the original data samples collected in the large information management system. Taking sample data as the test object set, semantic correlation feature extraction and big data mining can obtain feature extraction results. The analysis results show that the data mining conducted by this method has strong anti-interference feature extraction and good data amassation. In order to compare the performance of different mining algorithms, the present method and the traditional method conduct the target data mining and retrieval of the information management system under the same conditions. The check rate comparison results of the test data mining are shown in Table 3, and the data mining accuracy comparison results are shown in Figure 2.

According to the analysis of Table 3 and Figure 2, the greater the number of iterations, the higher the registration rate of this method is increased by 12.46%, and the calculation cost is decreased by 23.76%. The accuracy of storage data mining and retrieval of large information management systems is higher, and the performance is superior to the traditional method. However, the method designed in this study has big data and low accuracy of data scheduling, so the next step needs to be improved in this respect [10]. The construction of a large information management system is based on the database design and information retrieval algorithm, combined with information feature extraction and optimization, to carry out the overall design framework and database development of the information management system and improve the information retrieval and integrated information processing ability.

4. Conclusions

Through the above research work, we have mastered the core technical principles of associated data, the construction method of semantic association model, and the principles, processes, and key technical methods of associated data from construction, publication to consumption, and application. An agricultural information management data model based on cloud computing and semantic technology is proposed.
The cloud storage model is constructed for big data distributed storage design in large information management system, extracts the semantic correlation dimension features of information management distribution data in reorganized information management system, uses fuzzy C mean algorithm for the adaptive integration and clustering of distributed data in large information management system, uses the storage space dimension reduction processing of large information management system, and improves the target data mining ability and adaptive scheduling ability of information management system. Significantly improving the availability, visibility, availability, and scientific value of agricultural science and technology information resources will play an important role, which has important research value and practical significance for promoting the innovation of agricultural scientific research. Methods: the data mining of large-scale information management system has good accuracy and strong anti-interference ability.

Although the massive data contains a large amount of valuable information, the vast majority of them are semi-structured or unstructured and isolated data, lacking normative control, deep disclosure, and semantic correlation, making it difficult for computers to process these data automatically. To this end, people carried out the construction of semantic network research and launched the associated open data movement; more and more data providers and application system will release their own data in the Internet in the form of the associated data, implanted more standard semantic information in the data and interconnection with other data sources, formed a huge data network rich in semantics, and laid the foundation for computer accurate understanding and intelligent processing of the data.

**Data Availability**

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

**Conflicts of Interest**

The authors declare no conflicts of interest.

**References**


