Scientific Research Publishing

Open Journal of Business and Management, 2022, 10, 3132-3141
https://www.scirp.org/journal/ojbm
ISSN Online: 2329-3292
ISSN Print: 2329-3284

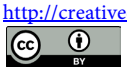
Research on the Construction of Enterprise Accounting Data Analysis Platform Based on Cloud Computing

Weihua Zhang

Zhejiang Yuexiu University, Shaoxing, China
Email: zwh_601@126.com

How to cite this paper: Zhang, W. H. (2022). Research on the Construction of Enterprise Accounting Data Analysis Platform Based on Cloud Computing. *Open Journal of Business and Management*, *10*, 3132-3141.
https://doi.org/10.4236/ojbm.2022.106156

Received: September 15, 2022
Accepted: November 8, 2022
Published: November 11, 2022



Abstract

With the emergence of massive data inside and outside the enterprise, paying much attention to the processing and analysis of accounting big data can bring huge value-added value to enterprises, and adapt to complex and changing economic environment. Based on this, by analyzing the related theories of accounting big data, accounting information and cloud computing, we build a cloud computing storage module of big data analysis platform, and apply Apriori data mining algorithm based on association rules to deal with massive data of accounting. A comparative prediction of the financial status of a group shows that the maximum error is less than 8% compared with the actual results, thus verifying the reliability and superiority of the established accounting big data analysis platform based on cloud computing.

Keywords

Cloud Computing, Enterprise Accounting, Data Mining Algorithm

1. Introduction

In the era of big data, data, information and knowledge are also important resources for enterprises (Hashem, Yaqoob, Anuar et al., 2019). Now many developed countries are striving to become knowledge-based enterprises, constantly using advanced information technology to excavate knowledge from accounting big data and enhance their core competitiveness. China has also clearly defined the "upgrading of information technology" into one of the goals of building a moderately prosperous society in an all-round way, so as to ensure sustained economic development and enhance comprehensive national strength. How to adapt to the complex economic environment and realize the value added of the

enterprise is also the focus of the enterprise information construction (Zhang, 2021). The core of enterprise informatization is accounting information. Because of the high cost, low efficiency, long construction cycle, limited technology and other factors, the existing accounting information system is difficult to get large amounts of accounting data from inside and outside enterprises, and find knowledge from them, so as to provide a scientific basis for business managers to make timely decisions (Xu, Huang, Chen et al., 2018). Therefore, how to acquire and excavate the valuable knowledge hidden behind the big data and promote the sustainable development of enterprises is a difficult problem for academic and business circles to tackle together.

With the development of the technology of the Internet of things, cloud computing is also widely used. It has the advantages of low cost, large storage space, fast processing speed and so on. All the information related research in the world has started to rely on it (Zhang & Chen, 2016). In the construction of the accounting information system, the cloud computing technology is used to obtain, cluster and analyze large accounting data, which not only overcomes the problem of large cost of traditional accounting information mode, but also greatly improves the efficiency of analyzing massive accounting data, and gradually exerts the value of accounting big data analysis. It provides a new idea for further developing and utilizing the big data of accounting, improving the relevance of accounting information, helping managers to make decisions, and achieving low cost accounting control in enterprises (Yang, Huang, Li et al., 2019).

For example, Alibaba Cloud migrated the cloud platform for Zheshang Securities. In order to ensure that customers across the country experience similar network access speeds, Zheshang Securities has set up several clusters in the computer rooms of different operators across the country to ensure a good user experience. However, deploying clusters in multiple regions of the country not only has a long construction period and high cost, but also faces great challenges in system expansion and management. At present, Zheshang Securities has deployed part of the market quotation and entrustment system on Alibaba Cloud. With the support of Alibaba Cloud's powerful network platform, multi-line BGP access has been realized. No matter which operator the end customer accesses through, whether in the north or the south, they can obtain good network access speed and quality, reduce the difficulty and cost of management, and facilitate flexible expansion, thus ensuring the end customer good experience.

2. State of the Art

As information technology is widely used in the field of accounting information, the research on accounting information has entered a new stage in the academic circle. Foreign information technology has been applied to accounting since 1950s. Under the impetus of information technology, foreign scholars study the process reengineering and construction of management information system (Sookhak,

Gani, Khan et al., 2015). To the existing research, scholars point out that there are a variety of research methods in the framework of accounting information research, which can only promote the construction of accounting information (Park, Ki, Jeong et al., 2016).

The domestic scholars put forward that accounting informationization relies on the existing information technology to integrate information flow, capital flow, business flow and logistics, which realizes the accounting information system with the combination of accounting and information technology, digital, dynamic, diversified and real-time (Huang, Lu, & Zhang, 2020). According to the investigation of the scholars, the research on the existing accounting information theory is not deep enough, and the theoretical framework of the comprehensive accounting information system is not formed (Liu, 2016). The scholars also think that the main reason is that scholars pay too much attention to the application of accounting information technology and its influence on accounting theory, ignoring its influence on the essence of accounting informatization (Yang, Huang, Li et al., 2019). Scholars believe that the intelligent accounting information system is the theoretical knowledge of complex system. We analyze the accounting data of all transformed into knowledge, and knowledge expression and complete intelligent storage, learning and memory, push, update and other functions, to achieve man-machine combination, intelligent decision accounting software. In addition, the scholar pointed out the characteristics of accounting information in the era of big data, and proposed that we should strengthen the construction of cloud computing and set up an accounting big data analysis platform to excavate the value behind big data. Because the cloud computing technology has many advantages, the use of cloud computing in the accounting information system can greatly improve the value of the accounting data (Ranjan, Georgakopoulos, & Wang, 2018).

3. Modeling of Cloud Computing and Big Data

3.1. The Theoretical Basis of Accounting Large Data Analysis

The financial department is one of the most closely related departments in the enterprise. The arrival of the big data era provides more information resources for accounting department managers' financial analysis. Big data is a large and complex data set, which can provide huge value for acquisition, storage, management, sharing and analysis within a reasonable time. The large data has four characteristics: the rapid and continuous increase of data, the fast speed of data input/output, the diversity of data types and sources, and the low value of the data.

Accounting information is a comprehensive application of modern information technology, such as computer, network communication, and so on, to obtain, process, output and apply the accounting data resources. It provides adequate, real-time and all directional accounting information for enterprise management, control decision and economic operation and is beneficial to the man-

agers' information decision making, thus improving the core competitiveness of the enterprise. Accounting information has the characteristics of universality, integration, dynamic and gradual. Its essence is the process of accounting data processing, and the goal is to establish modern accounting information system by means of information technology, and its purpose is to provide sufficient, real-time and omnibearing accounting information. Figure 1 is the application of information technology in the large data processing of accounting.

Using data mining technology can extract useful knowledge from large amounts of data, which can be applied to all aspects of enterprise financial early warning, product sales budget, customer value analysis and so on, so all walks of life pay close attention to data mining. The research of data mining is also a hot spot in today's application development. With the production of massive data and the sharing of information resources, the original information technology can not deal with mass data, and cloud computing technology is produced. Cloud computing is a business computing model. It concentrates all computing resources and uses hardware virtual technology to provide cloud computing users with powerful computing power, storage space and bandwidth. It assigns computing tasks to a large number of computer integrated resource pools, enabling different application systems to get corresponding storage space, information service and equal or more computing power with traditional large servers according to their needs.

3.2. Architecture of Large Data Storage Mechanism in Cloud Computing Environment

Cloud computing is through the Internet to provide dynamic and easy to expand large data storage space and structure model. In order to realize the clustering and classification of large data storage in cloud computing environment, it is

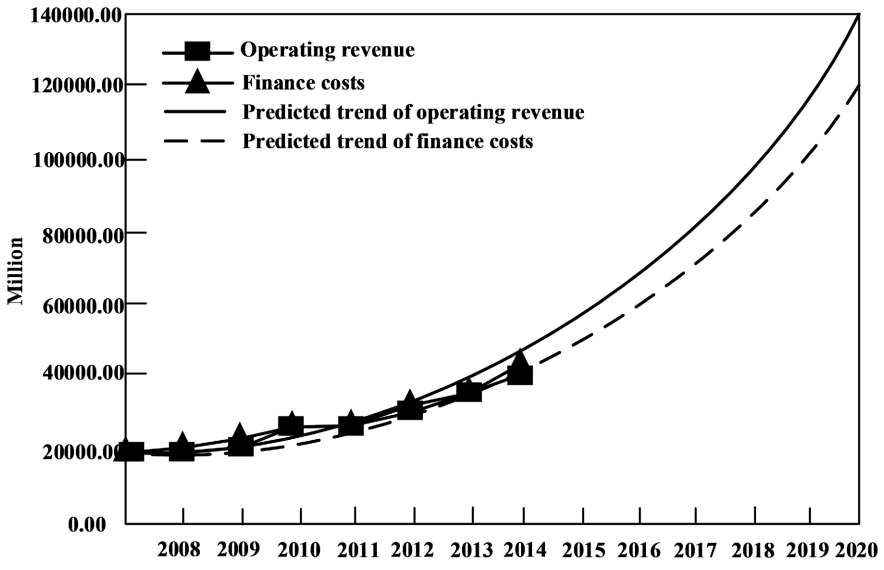


Figure 1. The predict results of operating revenue and finance costs.

necessary to build a large data storage system architecture in the cloud computing environment. In the cloud computing environment, large data storage uses a virtualized storage pool structure, and cloud computing deployment depends on the computer cluster. From top to bottom, they are the I/O virtual computer, the USB interface sequence and the disk layer respectively. The enterprise data center acquires application services through various terminals, so that computing is distributed on a large number of distributed computers.

When all the cloud computing virtual machines are allocated to the physical machine, the utilization (1) can be used to calculate the global optimal solution in this cluster. It can also distribute the large data feature clustering center V_{M_i} in all cloud computing to the physical machine P_{M_i} according to the optimal solution.

$$N = \frac{1}{n}\sum_{j=1}^{n}\left|U_{ti_j^{CPU}} - U_{ti_{avg}^{CPU}}\right| + \frac{1}{n}\sum_{j=1}^{n}\left|U_{ti_j^{Mem}} - U_{ti_{avg}^{Mem}}\right| + \frac{1}{n}\sum_{j=1}^{n}\left|U_{ti_j^{bw}} - U_{ti_{avg}^{bw}}\right| \tag{1}$$

The sample is analyzed and collected to judge whether the sample is a typical sample, and the sample is the data, and a large database letter and data flow sample $S = \overline{X}_1, \overline{X}_2, \cdots, \overline{X}_k$ is set up. Data information sampling is carried out in time period $T_1, T_2, \cdots, T_K$ respectively. Now we divide the large data set X into the c class in the cloud computing environment, of which $1 < C < n$. The segmentation of data is transformed into space segmentation, and the central vector of large data storage structure is as follows:

$$V = \left\{v_{ij} \mid i = 1, 2, \cdots, c,\ j = 1, 2, \cdots, s\right\} \tag{2}$$

Among them, V_i is the i vector of the target clustering feature (the i cluster center vector). The fuzzy partition matrix is expressed as:

$$U = \left\{\mu_{ik} \mid i = 1, 2, \cdots, c, k = 1, 2, \cdots, n\right\} \tag{3}$$

Single data source is processed by redundant data reduction. In the process of virtual machine clustering mining for multi-channel QoS requirements, the input part (the set of virtual machines and physical machines) and the related parameters are $V_{MS} = \left\{V_{M_1}, V_{M_2}, \cdots, V_{M_m}\right\}$, $P_{MS} = \left\{P_{M_1}, P_{M_2}, \cdots, P_{M_n}\right\}$. The heuristic factor is α, and the expected value of the heuristic factor is β, and the maximum number of mining times is $I_{\max}$. Thus, the data blocks uploaded by the client provide a fixed size data block to achieve cloud clustering. Through the analysis of the large data storage system architecture in the cloud computing environment, it provides an accurate data base for large data analysis.

3.3. Modeling of Large Data Mining Algorithm Based on Association rules

Because of the huge amount of data in the accounting work, the Apriori data mining algorithm based on association rules is used to deal with large amount of accounting data. Apriori data mining algorithm belongs to a kind of association algorithm, which has the basic. It describes the underlying association between

data items, which are classified as single, single, and Boolean association rules. The algorithm uses a sequential search cycle by layer to dig frequent sets.

The Apriori algorithm has the following properties: any infrequent $(k-1)$ set can not be a subset of frequent k item sets. This is because if the percentage of transactions that contain $(k-1)$ item sets is not greater than the minimum support threshold, then the percentage of transactions that contain the k item set and another K item set can not be greater than or equal to the threshold of minimum support. If we use the concept to explain, the connotation of the concept of transaction composition containing k item set is increased than that of the concept consisting of $(k-1)$ item set, so its extension will inevitably decrease, and the number of transactions involved will also decrease.

Therefore, it is possible to delete a set of infrequent items in a set of $(k-1)$ before the K item set is generated by this property. The frequent itemsets are obtained by deleting the infrequent itemsets in the candidate $(k-1)$ item set $(k-1)$. The basic process of the algorithm is as follows: 1) irst, calculate all the C_1; 2) scanning database, deleting infrequent subsets and generating L_1; 3) connect L_1 with its own to generate C_2; 4) Scan the database, delete infrequent subsets in C_2, and generate L_2; 5) By analogy, the C_k is generated through the L_{k-1} connection with itself, and then the database is scanned to generate L_k until no more frequent itemsets are generated.

Frequent item sets are divided into two steps: connection and pruning.

Join step: we use the recursive connection method to find L_k, and use L_{k-1} to connect the set of candidate k item sets with its own connection. We record the set of candidate k item sets as C_k. Setting l_1 and l_2 is an item set of L_{k-1}, and the entries in l_1 are lexicographic and the mark $l_1[j]$ represents the l_1 item of the j. If l_1 and l_2 's previous k-2 corresponding items are equal, then l_1 and l_2 can be connected:

$$\left(l_1[1]=l_2[1]\right)\wedge\cdots\wedge\left(l_1[k-2]=l_2[k-2]\right)\wedge\left(l_1[k-2]<l_2[k-2]\right) \tag{4}$$

l_1 and l_2 produce new item sets:

$$l_1[1]l_1[2]\cdots l_1[k-2]l_1[k-1]l_2[k-1] \tag{5}$$

This item is added to the set until a new item set is not produced.

Pruning step: a set C_k is generated by a connection. It contains some k item sets that are not frequent k entry sets but all the frequent k items are included in the collection C_k. If the k item set is compared to the transaction in the database, the support of the set in the transaction can be obtained. If the K item set that does not meet the minimum support degree is deleted, the frequent K itemsets can be obtained.

3.4. Construction Feasibility Analysis of Large Data Analysis Platform Based on Cloud Computing

Y can be regarded as the integral of the independent variable x for the large accounting data of the enterprise:

$$Y = \int \rho(x) \mathrm{d}x \tag{6}$$

The $\rho(x)$ is the density of the large accounting data. Get all the objective information through the upper form. On this basis, it is confirmed that the useful accounting data V is the value amendment of the large accounting data Y:

$$V = Y^r \tag{7}$$

The value coefficient $r \in [0,1]$, when $r = 1$, $V = Y$, all the big accounting data are valuable; When $r = 0$, $V = 1$, there is an accounting data value. The knowledge K is the integral of the useful accounting data V

$$K = \int iV \mathrm{d}V \tag{8}$$

Among it, i is the knowledge conversion coefficient of useful information. According to the above theory, the accounting information system is to classify, summarize and excavate the objective information Y, and automatically provide the decision information K. The accounting big data analysis platform in this paper is to expand the scope of accounting data. Based on the above theories, we use information technology to do mining and analysis of accounting big data and provide decision information, so the establishment of platform is theoretically feasible.

4. Analysis of Examples

4.1. Forecast of Operating Income and Operating Cost

Analysis of the performance of a large data analysis platform is based on a group of cases. The group is a large enterprise with pharmaceutical retail as the core. It has the characteristics of large scale, rapid development and many institutions. With the continuous development and expansion of the group, the problem of management mode is becoming more and more obvious. The group adopts centralized management, and the group headquarters centrally manages the budgets and financial decisions of subsidiaries, factories and departments. In order to test the predictability of the platform, we import all the financial data of the group from 2008 to 2013, and predict the financial status of operating income and operating cost from 2015 to 2020. The accuracy of the platform prediction is verified by comparing the predicted number with the actual number.

According to the operating income and operating cost from 2008 to 2013, the platform automatically draws the corresponding trend line. The user can choose to predict the next few years, and the forecast line will extend automatically, as shown in **Figure 2**. It is predicted that the operating income of 2008 is 41,829,787,656 yuan, the actual amount is 38,721,656,259 yuan, and the error rate is 8%. It is predicted that the operating income of 2009 is 54,795,373,842 yuan, the actual amount is 68,078,217,820 yuan, and the error rate is 0.05%. It is predicted that the operating income of 2010 is 67,497,346,641 yuan, and the actual rate is 68,078,217,820 yuan, and the error rate is 1%. It is predicted that the operating income of 2011 is 78,313,511,241 yuan, and the actual operating

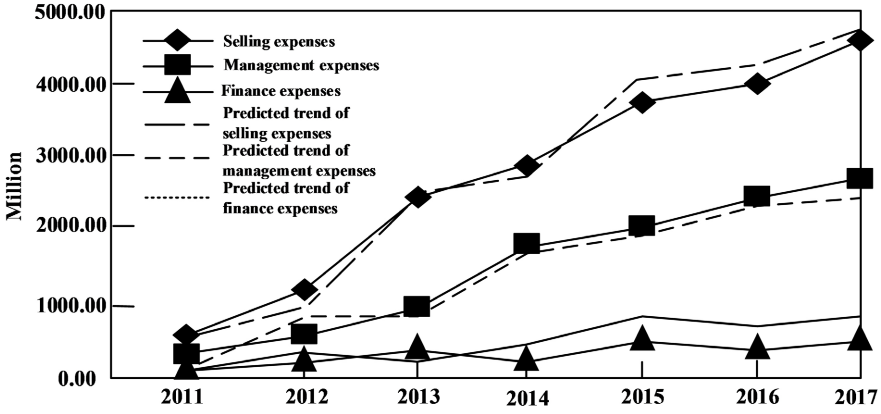


Figure 2. The predict results of three charges.

income is 78,232,818,357 yuan, and the error rate is 0.1%. Except the prediction error of 2008, the other years are within the range of prediction error. The forecast of operating costs is affected by operating income. Except to the forecast error of 8% in 2008, other forecasting errors are also allowed. In 2011, the main reason for the sales volume less than the forecast is that the introduction of the new health care reform has brought a new round of structural adjustment and market expansion. The macroeconomic environment is uncertain, and the changes in policies and regulations have an impact on the industry, resulting in a decline in sales. The group then adopted measures to integrate other companies to restore sales in 2011.

4.2. Three Items of Cost Forecast

It forecasts the financial situation of the group's sales expenses, management expenses and three expenses of financial expenses, and now there are 2011-2017 years of actual financial situation. In 2011 years, the forecast sales cost is 2847458501.50 yuan, the actual amount is 3071521304.81 yuan and the error is 3%. In 2012, the forecast sales cost was 3272341449.25 yuan, the actual sales cost was 3288786883.64 yuan, and the forecast error was 0.5%. In 2013, the sales cost was 3857956.28 yuan, and the actual sales cost was 4407325498.15 yuan, and the error was 3%. In 2014, the sales cost was 4,394,103,832 yuan, and the actual sales cost was 4417235498.08 yuan, and the error was 0.3%. The sales costs of 2015, 2016 and 2017, and the forecast results for the management and financial costs of each year are shown in Figure 2.

There is a big error between the cost of management and the cost of sales in 2011 and 2013. The main reason is the integration of many enterprises in 2011 and 2013, with a large number of sales and management costs in the consolidated statement. The increase in financial costs in 2012 was caused by a large amount of remittance losses, and 2014 returned to normal. Moreover, internal financing is adopted to reduce the financial cost, but there is a real difference between the predicted intra platform exchange rate and the real exchange rate,

resulting in the prediction error in 2011 and 2013. The above situation is a sudden situation, the platform can not be foreseen, but also the shortcomings of the platform.

5. Conclusion

With the development of economy, the improvement of computer level, network and intelligence, modern society has already stepped into the information stage. Especially in the field of accounting, with the deep application of cloud computing and other technologies, the construction of enterprise accounting information has been accelerated. The current situation of the traditional accounting information system is analyzed, and its shortcomings such as single accounting data collection, poor integration between the system and other systems, insufficient accounting information and so on are analyzed. It is difficult to meet the new needs of the managers and the need for the value creation of accounting data. Therefore, it is necessary to apply advanced technology to solve the above problems. Using cloud computing technology to build platform has the advantages of high efficiency of data processing and low construction cost. Based on the theory of accounting information, we use cloud computing technology to build accounting information system based on accounting big data analysis, and mainly build an accounting big data analysis platform. The cloud computing storage module of the big data analysis platform is modeled in detail, and the Apriori data mining algorithm based on association rules is applied to deal with the massive data of accounting. Finally, based on the analysis and prediction of the past few years and future financial situation of a group, by comparing the actual data in the past few years, it is proved that the prediction data of the platform is reliable. The large data analysis platform based on cloud computing can help enterprises to develop rapidly and enhance the core competitiveness of enterprises.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

Hashem, I. A. T., Yaqoob, I., Anuar, N. B. et al. (2019). The Rise of "Big Data" on Cloud Computing: Review and Open Research Issues. *Information Systems, 47,* 98-115. https://doi.org/10.1016/j.is.2014.07.006

Huang, M. L., Lu, L. F., & Zhang, X. (2020). Using Arced Axes in Parallel Coordinates Geometry for High Dimensional Big Data Visual Analytics in Cloud Computing. *Computing, 97,* 425-437. https://doi.org/10.1007/s00607-014-0383-z

Liu, D. Y. (2016). Analysis of Problems and Countermeasures in the Application of Accounting Information Systems. *Finance, No. 3,* 21-22.

Park, J. S., Kim, H., Jeong, Y. et al. (2016). Two-Phase Grouping-Based Resource Man-

agement for Big Data Processing in Mobile Cloud Computing. *International Journal of Communication Systems, 27,* 839-851. https://doi.org/10.1002/dac.2627

Ranjan, R., Georgakopoulos, D., & Wang, L. (2018). A Note on Software Tools and Technologies for Delivering Smart Media-Optimized Big Data Applications in the Cloud. *Computing, 98,* 1-5. https://doi.org/10.1007/s00607-015-0471-8

Sookhak, M., Gani, A., Khan, M. K. et al. (2015). Dynamic Remote Data Auditing for Securing Big Data Storage in Cloud Computing. *Information Sciences, 380,* 101-116. https://doi.org/10.1016/j.ins.2015.09.004

Xu, J., Huang, E., Chen, C.-H. et al. (2018). Simulation Optimization: A Review and Exploration in the New Era of Cloud Computing and Big Data. *Asia-Pacific Journal of Operational Research, 32,* Article ID: 1550019. https://doi.org/10.1142/S0217595915500190

Yang, C., Huang, Q., Li, Z. et al. (2019). Big Data and Cloud Computing: Innovation Opportunities and Challenges. *International Journal of Digital Earth, 10,* 13-53.

Zhang, L. J. (2021). Editorial: Big Services Era: Global Trends of Cloud Computing and Big Data. *IEEE Transactions on Services Computing, 5,* 467-468. https://doi.org/10.1109/TSC.2012.36

Zhang, Q., & Chen, Z. (2016). A Weighted Kernel Possibilistic c-Means Algorithm Based on Cloud Computing for Clustering Big Data. *International Journal of Communication Systems, 27,* 1378-1391. https://doi.org/10.1002/dac.2844