International Journal of Multiphysics
Volume 18, No. 2, 2024
ISSN: 1750-9548

Research on Engineering Cost Data Management System based on Decision Tree Algorithm

Zhaogang Wang

China tobacco corporation Guizhou province company audit Chu, Guiyang city, Guizhou province, 550003, China

Abstract

The informatization level of the engineering cost industry in China is still relatively low and the development degree is unbalanced. After China acceded to the WTO, some foreign cost consulting and related enterprises began to enter China's engineering cost market, and some large domestic enterprises also began to actively explore the international market. Under such a situation, we need to take the research on big data as an opportunity to drive the industry through data and establish and develop a highly autonomous big data industry chain through the government's organization and enterprises' participation, aiming to achieve leapfrog development.

Keywords: Engineering cost; Data management; Decision tree algorithm; Informatization

1. Introduction

With the rapid development of modern society, human society has gradually entered the era of big data. The research on big data is the development and extension of informatization construction. Meanwhile, the analysis and mining of big data can better discover the potential knowledge in the data, provide better feedback for the industry construction, improve the industry management level, save social resources, and improve the social investment efficiency. Facing the impact of big data and the increasingly open construction market in China after joining the WTO, China's engineering cost industry should take measures in advance, accelerate the pace of industry development, narrow the gap with the management level of the construction engineering industry in developed countries, and enhance its competitiveness in the domestic and international markets.

In recent years, the research on big data has been paid much attention to by governments around the world. In March 2012, the Obama Administration of the United States announced the "Big Data Research and Development Initiative", elevating big data research to the highest national policy. This year was called the enlightenment year of big data, and online attention to big data also rose sharply this year. Countries around the world are also competing to follow suit. For example, of the eight types of high-tech invested 600 million pounds by the United Kingdom, 189 million pounds were invested in big data research. Besides, Japan also regards big data research as a national strategy.

2. Development Situation of Informatization Construction of Project Cost Management in China

During the process of construction engineering activities, project cost management runs through the entire process and often accumulates a large amount of data, which guides the construction of future projects.

After the reform and opening up, China's construction industry has developed rapidly. Meanwhile, the continuous progress of computer and network technologies makes it possible for the project cost industry to carry out informatization construction. In this paper, it is believed that the informatization of project cost is the process of developing and utilizing project cost information using modern information technology, and then serving the construction engineering industry and improving the industry efficiency.

China's major construction departments attach great importance to the construction of informatization on project costs. In the first half of 2002, the Quota Division of the Ministry of Housing and Urban-Rural Development (formerly the Ministry of Construction) established the China Construction Project Cost Information Network, which uniformly releasedrelevant information through the network, and the construction of the administrative affairs information database related to project cost management was initially completed. Meanwhile, the networking methods of local project cost information networks and the methods to achieve a unified national cost database were stipulated. In 2011, the Quota Division of the Ministry of Housing and Urban-Rural Development issued *Several Opinions on Doing a Good Job in the Information Management of Construction Project Cost*, which once again clarified the goal of information management of project cost, made unified arrangements for the division of labor of local cost management departments at all levels, and required further strengthening the construction of the information platform for project cost.

Relevant enterprises are actively promoting the industry's informatization construction. Since the 1990s, project cost management software independently developed by Chinese enterprises has been available. After more than ten years of rapid development, it has now developed into a relatively mature industry. Cost software such as Glodon and THSWARE have developed into leaders in the domestic cost software industry. Currently, the penetration rate of computerization in the construction project cost industry in China is already relatively high.

3. Countermeasures for the Project Cost Management Industry in the Era of Big Data

Although China has made great efforts in the informatization constructionprocess and has made some achievements in infrastructure construction and information management, there are still many problems, such as uneven development and the inability of market cost software and government cost information management systems to exchange data. Therefore, utilizing the opportunities brought by the era of big data and actively learning from the relatively mature methods and approaches abroad to establish a management model and supporting measures in line with China's national conditions can achieve twice the result with half the effort.

The general process of big data processing is information collection, data integration, data preprocessing, data mining, pattern evaluation, and knowledge representation. SPSS and SAS companies have proposed their respective commercial data mining standard processes on this basis, namely, CRISP-DM and SEMMA. According to the characteristics and requirements of big data, corresponding measures for key links are proposed in this paper.

3.1 Establish a unified engineering data standard to avoid information isolated islands

The research on big data requires the collected data to be unified and standardized, aiming to facilitate later data analysis and data mining. Although China currently implements a unified list-based pricing standard for engineering measurement, most enterprises still refer to the methods stipulated in the local engineering quotas in the pricing link. However, the local engineering quotas are different in the measurement units, calculation methods and applicability of some sub-projects, resulting in poor universality of the included data and inconvenience for collation and analysis. Breaking this regional difference and establishing a unified metadata standard is an important step in the big data collectionprocess, which will facilitate data exchange between different systems and significantly reduce the time and cost of data processing. Meanwhile, the research on big data requires allowing and regulating the input of semi-structured and unstructured data and standardizing their input rules accordingly. For example, for the description of materials, there can be not only text descriptions but also corresponding pictures. For process flows, relevant videos can be provided, but pictures and videos, etc. must comply with the corresponding standards.

3.2 Establish distributed databases for data storage

At present, the main mode of data storage is the non-relational (NoSQL) database. This type of database stores massive data in different regions and different servers, and obtains the information required by users through rapid information exchange between servers. Recently, the NewSQL-type database, which combines the advantages of both SQL and NoSQL, has been developed.

Due to the large number of contracted projects and relatively rich resources, large construction engineering enterprises can build their own distributed databases by using their own conditions. However, small and medium-sized enterprises cannot independently construct a huge database, requiring the unified construction of the project cost data warehouse by the government's cost management department. China's construction management department has carried out the initial informatization construction, so it has a certain foundation and owns the basic conditions for establishing databases.By establishing a unified database platform through the government's cost department, more enterprises, especially small and medium-sized enterprises are allowed to join in, which not only strengthens data supervision and saves social costs but also may discover more potential value and create huge social and economic benefits.

3.3 Focus on data analysis and data mining

China's project cost industry has accumulated a certain foundation in the informatization construction process, but there is a lack of corresponding analysis and mining tools for the accumulated relevant data of project cost, resulting in "rich data but scarce information". Currently, the algorithms for data mining have been relatively well-developed. The main algorithms include classification algorithms, clustering algorithms, association algorithms, sequence mining, etc. Currently, data mining technology is developing rapidly in some developed countries. Some data analysis companies have developed mature commercial data mining tools. These commercial data analysis companies are also actively exploring the Chinese market, and even some foreign data analysis companies develop markets in the form of free data analysis. However, if excessively relying on foreign analysis software, it will inevitably lead to information leakage. Moreover, project cost data often involves many trade secrets and even state secrets, which is neither conducive to the information security of enterprises and the government nor conducive to the expansion of enterprises in overseas markets. Actively researching and developing data mining tools with independent intellectual property rights and forming the corresponding industrial chain are conducive to the healthy and rapid development of China's project cost industry.

3.4 Pay attention to the visualization research of the mining process and results

Boring data analysis and mining require to be carried out by professionals. However, a large number of project cost personnel may not be able to understand the huge and complex digital results intuitively. With the help of visualization technology, more cost information managers can understand the entire data analysis process and the conclusions drawn from data mining, providing better guidance for production applications.

Currently, some relevant websites at home and abroad attach great importance to the development of data visualization, such as Google Charts, Tableau Software, etc. Domestic related cost software and cost websites are also actively conducting visualization research. For example, Glodon Software can convert its input data into 2D and 3D graphics during its modeling process and associate it with other data, which is convenient for cost personnel to view intuitively during use. Big data also puts forward higher requirements for visualization technology, and it requires relevant visualization software to be built on distributed databases to help managers view and analyze timely. Paying attention to the research on visualization can significantly reduce the threshold of big data management and enable more cost personnel to master the weapon of big data analysis and mining.

3.5 Establish an open, timely and effective information query and release model

Currently, the Chinese government, as the largest owner and publisher of cost information resources, has a series of problems in the release of cost information, such as information lag, insufficient processing of the released data, lack of guiding role, and low sharing rate. However, ifthe information obtained from big data research and mining is not released to society timely, its economic and social benefits will be significantly reduced. Foreign experience shows that establishing an open big data information query service that can provide for society can not only further promote the development of small and medium-sized enterprises but also increase industry transparency, promote exchanges between enterprises, and maintain industry vitality. As the organizer, it is very necessary and feasible for the government to disclose resources to a greater extent. Of course, attention should also be paid to the classification and processing of relevant data, publicizing the general and low-tech socially common data and strengthening the confidentiality of information involving high technology, business secrets of

enterprises and state secrets. This can not only enhance the transparency of project costs and reduce the average construction cost of social construction projects but also encourage the development of technological and innovative enterprises.

4. Conclusion

Overall, the rise and development of research on big data has a huge impact on traditional data analysis and management. The project cost industry has huge data resources. The industry construction is analyzed from the perspective of big data research, and an industry big data industrial chain is built by establishing unified data entry norms, building distributed databases, and strengthening data analysis and mining, aiming to improve the management level and competitiveness of the entire industry.

Refrences

[1] Yuan Xin. Application Research of Data Mining in Construction Cost Data Management System [D]. Wuhan University of Technology [July 30, 2024] DOI:CNKI:CDMD:2.1015.001586.

[2] Chu Yanqiu. Design and Implementation of Engineering Cost Management Information System [D]. Xiamen University, 2014.

[3] Zhang Hong. Design and Implementation of Engineering Cost Management System Based on RCP [D] 2012.DOI: http://dspace.xmu.edu.cn:8080/dspace/handle/2288/44618.

[4] Guo Kui. Application Analysis of Building Material Feature Analysis Model in Engineering Cost Data Management System [J]. Jushe, 2024 (20).

[5] Ma Ning. Design and Implementation of XML based Engineering Cost Management System [J]. Computer Application, 2010. DOI: CNKI: SUN: JSJY.0.2010-S1-026.

[6] Gao Yiquan, Xia Rui, Zhan Liping, et al. Innovation and Practice of CIPS Cost Control System for the Whole Process of Power Transmission and Transformation Projects Based on Massive Data [J]. Engineering Economics, 2023, 33 (1): 68-74.

[7] Shao Biao, Yao Wenjuan, Liu Jian. Research on Cloud Computing based Engineering Cost System [J]. Commodities and Quality: Real Estate Research, 2014 (6): 1.