

Research on Computer Software Engineering Project Automation Management based on Data Mining and Fuzzy Clustering

Li Zhang

Wuhan Qingchuan University, Wuhan, China

zhangly0202@126.com

Abstract. The smooth implementation of automation management of computer software engineering projects has played an important role in promoting the faster development of computer software development and further promoting the development of computer software engineering. In the process of realizing automation management of project, data mining technology and method are introduced, and a computer aided quality function configuration model based on data mining is proposed. The data mining tool was developed, and the model for product user demand and process quality selection decomposition was established. At the same time, the data mining model of fuzzy clustering and grey theory algorithm was established, and related application examples were analyzed. At present, all methods of validity discrimination do not have absolute advantages, and all need to be adapted and selected for specific data sets.

Keywords: Data Mining and Fuzzy Clustering, Computer Software, Automation Management.

1. Introduction

Nowadays, it is the era of computer information technology. With the rapid development of computer technology, software development projects are becoming increasingly hot [1]. With the continuous expansion of the scale of software development teams and the increasing number of software developers, there are more and more problems in software engineering projects. Without standardized management, it will bring adverse effects on the sound development of software engineering projects [2]. At the same time, in the application process of computer technology, people generally attach importance to computer software, but they do not attach enough importance to the automation management of software engineering, neglecting to take effective measures to improve and enhance, affecting the full play of the role of computer software [3]. In this context, we should change our mindset, recognize the important role of automation management, and effectively integrate it into computer software engineering projects [4]. As an important part of the development of computer technology today, computer software development has ushered in a series of new problems while achieving excellent results. Therefore, it is urgent to strengthen the automation management of software engineering projects [5].

Software engineering automation technology mainly uses programmable controllers, and programmers perform repeated erasure code debugging on the flash memory inside the programmable controller [6]. The logic of the controller circuit is controlled by software code to program the hardware chip, so that each circuit logic can play its corresponding role, and the internal series of complex operations are intuitive and simple [7]. By performing the operation process inside the memory, some timing and counting operations are performed to complete the instruction. This kind of memory can realize digital or analog data input and output to operate and control computer operation and machine production plant operation, etc. Its performance has been greatly improved compared with the previous control system [8]. Data mining is one of the most advanced and active research directions in the field of database and information decision-making in the world. The main body of this paper is the fuzzy clustering algorithm based on objective function in data mining method, with emphasis on the analysis of clustering effect. Among them, the purpose of validity analysis is to get the ideal number of clusters, so that the clustering results can best reflect the results of data sets [9].

2. Methodology

Data mining is a process of extracting potentially useful information and knowledge hidden in a large number of random practical application data, which people do not know beforehand. It is mainly used in classification, clustering, Association and prediction. Data mining methods include association method, clustering method, concept description method, etc. The commonly used technologies are decision tree, genetic algorithm, neural network, expert system, fuzzy theory, Bayesian theory, grey theory and rough set, etc. Characteristics of software engineering projects [10]. The characteristics of a computer software engineering project can be summarized as follows: First, there is no visibility. Software can't be touched. To perceive the existence of software, you must use special software engineering methods and corresponding testing methods. Secondly, there is no uniqueness. When the software is designed, there is no standard general software design process, and its design methods are various; finally, it is risky. This technology has a wide range of applications in the current industrial production field, effectively promotes the improvement of engineering development efficiency, and provides reliable automation guarantee for the monitoring and management of every aspect of industrial production.

The proposed data mining algorithm is compared with the fuzzy clustering algorithm. The algorithm defines the degree of difference from the perspective of the set, and the data is scanned once by data compression to obtain the clustering result. The amount of data processing and calculation is small. It can be seen from Table 1 and Figure 1. Compared with the fuzzy clustering algorithm, the data mining algorithm has the same running time, but the clustering quality is improved.

Table 1. Average mean and average running time on boundary clear data sets

Algorithm	Average value	Average running time(s)
Data mining	0.56	3.5
Fuzzy clustering	0.44	4.2

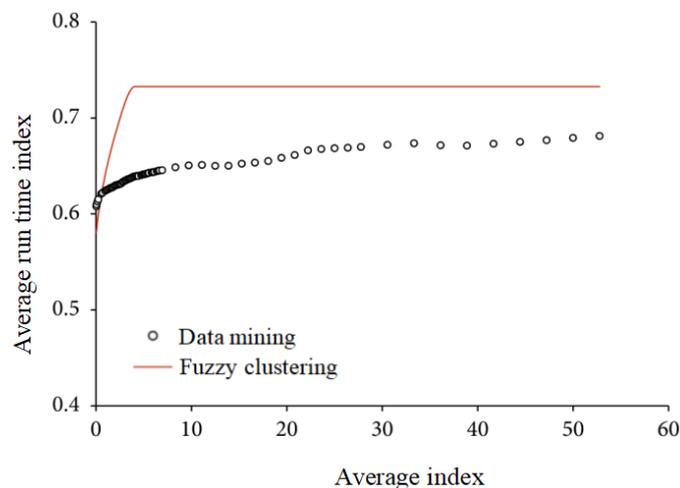


Fig. 1 Average mean and average running time on boundary clear data sets

As the main body of computer software project automation project, staff should be responsible for the implementation of the project, including searching for information, writing software programs, follow-up processing work, etc. In this case, it is necessary to rationally arrange computer software developers in accordance with the scale and nature of the project, and require staff to have a team in addition to their professional knowledge. Team collaboration ability, so as to better complete the complex computer software automation management. Fuzzy clustering analysis can be roughly divided into two kinds, one is based on the fuzzy relationship of the fuzzy clustering method, that is, system clustering analysis, the other is called non-system clustering analysis. After the selection of statistical indicators, the steps of fuzzy cluster analysis mainly include data standardization, calibration and clustering. "Clustering" is the process of distinguishing and classifying things according to certain requirements and laws. A sample set without class markers is divided into several subsets according to certain criteria. According to a certain evaluation method, samples with high

similarity are exhausted. It may be classified into one class, and the samples that are not similar are divided into different classes as much as possible, and $[1, 2]$ is used to indicate the degree of similarity between samples.

In the process of computer software engineering project automation management, it often involves problems such as demand analysis and business processing. Because these two tasks are not considered adequate or low, the software engineering management work cannot meet the expected goals. On the basis of saving manpower, it is possible to browse and use a huge database, and establish a series of complex numbers and manipulation models to calculate and analyze the entire project. As a new field of computer automation development, computer intelligence is increasingly enabling high performance and high analogy. And in the field of industrial production has been effectively realized, automatic management technology has greatly improved the efficiency of industrial production and industrial construction. Automation management technology can also improve the efficiency of project development in the development and application of software engineering projects. Because clustering is unsupervised, the validity of clustering results must be studied, that is, data sets should be divided into several categories to best reflect the structure of data sets. The optimal number of categories is determined by validity indicators, so the study of validity indicators is very important.

3. Result Analysis and Discussion

By calculating the similarity judgment matrix, we get the normalized fuzzy clustering input data set. By setting the appropriate number of clustering categories, we can get the clustering from each sample to the center of the cluster. By setting the appropriate membership threshold, we can get the required classification results. At present, the automation of computer system is widely used in the operation of automation control system. For example, the use of numerical control machine tools to complete an automated manufacturing process can realize the automation of parts production and processing. In addition, the automation of computer software is affected by the development of computer technology, and gradually began to explore the field of computer intelligence. And in the daily work practice to strengthen the accumulation of risk prevention knowledge, so that in the project implementation process, timely and accurate identification of risks, identify potential threats and hidden dangers. All management personnel of computer engineering projects should establish risk awareness and learn risk management knowledge so that they can objectively identify risks and conduct quantitative analysis, and then adopt targeted risk prevention and control methods.

Finally, we give a brief summary of the above experiments. It can be seen from the visualization results of the clustering of the data set shown in Figure 2 under the decision model that the fuzzy clustering integrated decision model proposed in this paper is closer to the true distribution of the data.

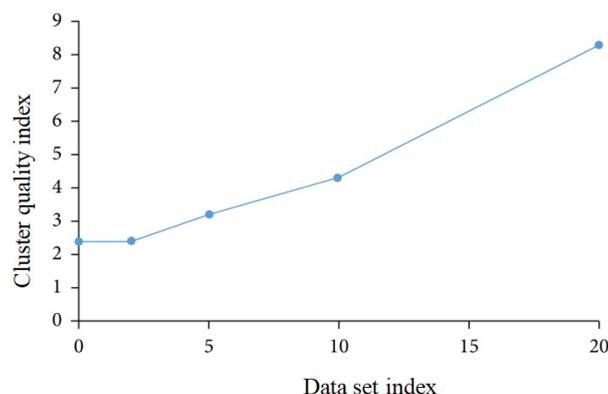


Fig. 2 Clustering visualization results

Because computer software development projects involve a wide range of aspects and aspects, it is necessary to enhance the management of computer software development. The management of computer software development is mainly divided into the following aspects: First, effective

management of actual development and design personnel. The software engineering development and design team usually includes the project's general responsibility, senior management, and junior software programmers. The data mining method based on grey theory can meet the requirements of real-time quality forecasting, and can extract the feature information in the measured data to the maximum extent, thus effectively reducing the influence of measurement system error. Grey relational analysis is a multi-factor statistical analysis method. It is based on historical data of each factor and uses grey relational analysis to mine the strength, size and order of factors. In addition, the inspectors of software engineering need to strictly inspect the engineering automation facilities and software parts to ensure the normal operation and cooperation of software and hardware, thus effectively guaranteeing the quality of the overall development of software engineering.

When using the equipment, we should pay attention to distinguishing and testing the foreign software to determine whether it is suitable for the project and the company's equipment compatibility. At the same time, we should also pay attention to checking and modifying after recording. The inspectors of engineering projects should determine the normal operation of various application software of automation equipment, and communicate with the responsible units and departments in time when they discover functional and equipment problems when inspecting relevant software. We should ensure that the surrounding environment is dry and do a good job of dust removal. However, some staff members did not pay attention to this problem, failed to strengthen the surrounding environment detection, neglect the effective treatment of dust, dust, smoke, magnetic field caused unnecessary interference to computer software engineering, and restricted the automation management level of computer software engineering. When you want to read the data, you should also do it on a computer that is not connected to the external network to prevent interference and information leakage on the network. In addition, you should strengthen the virus monitoring program while strengthening the computer's own security scheme, pay attention to timely update the virus. The database keeps the virus's database up to date with the latest version, minimizing the risk.

4. Conclusion

In summary, with the continuous development of science and technology in computer software engineering projects, the scale of software engineering projects has been continuously expanded, and the number of developers has increased, making it necessary to carry out necessary management of software engineering projects. Automated management plays a very important role in software engineering project management. In addition, the automation of computer software is gradually affected by the development of computer technology, and gradually began to explore the field of computer intelligence. Intelligent operation is improved by simulating human intelligence and operation. On the basis of saving manpower, it can realize the browsing and use of a huge database, and establish a series of complex numbers and manipulation models to calculate and analyze the entire project. In this paper, a tool based on data mining is proposed for the realization of computer software automation system and a framework model is given. Data mining is a novel and very active research field. It is a relatively new attempt to introduce data mining technology into product quality and function allocation, which is of great significance. At the same time, an automatic management method of computer software engineering project based on data mining and fuzzy clustering is proposed. By clustering similar engineering project groups, the classification of automatic management can be more accurately depicted.

References

- [1]. Sangaiah A K, Subramaniam P R, Zheng X. A combined fuzzy DEMATEL and fuzzy TOPSIS approach for evaluating GSD project outcome factors. *Neural Computing and Applications*, Vol. 26 (2015) No. 5, p. 1025-1040.
- [2]. Izakian H, Pedrycz W. Agreement-based fuzzy C-means for clustering data with blocks of features. *Neurocomputing*, Vol. 127 (2014), p. 266-280.

- [3]. Shamshirband S, Goci? M, Petkovi? D, et al. Clustering project management for drought regions determination: A case study in Serbia. *Agricultural and Forest Meteorology*, Vol. 200 (2015), p. 57-65.
- [4]. Kumar S, Kathuria M, Gupta A K, et al. Fuzzy Clustering of Web Documents Using Equivalence Relations and Fuzzy Hierarchical Clustering. *Eprint Arxiv*, (2014), p. 1-5.
- [5]. Karachi A, Dezfuli M G, Haghjoo M S. *Intelligent Information and Database Systems. Lecture Notes in Computer Science*, Vol. 5990 (2017) No. 6, p. 891-901.
- [6]. Srinivas C, Radhakrishna V, Rao C V G. Clustering and Classification of Software Component for Efficient Component Retrieval and Building Component Reuse Libraries. *Procedia Computer Science*, Vol. 31 (2014), p. 1044-1050.
- [7]. Phanikrishnakishore M, Madamsetti A K. Attribute Level Clustering Approach to Quantitative Association Rule Mining. *International Journal of Computer Applications*, Vol. 95 (2014) No.6, p. 17-23.
- [8]. Osborne F, Scavo G, Motta E. A Hybrid Semantic Approach to Building Dynamic Maps of Research Communities. *Lecture Notes in Computer Science*, Vol. 8876 (2014), p. 356-372.
- [9]. Kim J Y. Segmentation of lip region in color images by fuzzy clustering. *International Journal of Control Automation & Systems*, Vol. 12 (2014) No. 3, p. 652-661.
- [10]. Khoshnevisan B, Rafiee S, Omid M, et al. Developing a fuzzy clustering model for better energy use in farm management systems. *Renewable and Sustainable Energy Reviews*, Vol. 48 (2015), p. 27-34.