# Human Resource Intelligent Scheduling Algorithm Based on Decision Tree Algorithm

1st Na Feng
*Yinchuan Institute of Science and Technology*
Yinchuan, China
394348993@qq.com

2022 IEEE 2nd International Conference on Mobile Networks and Wireless Communications (ICMNWC) | 978-1-6654-9111-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICMNWC56175.2022.10031690

***Abstract*—With the development of the global network economy, competition in the global market has become increasingly fierce, the scale and number of commercial projects have continued to grow, and the requirements for human resource planning and management have become higher and higher. Traditional human resource management methods are often based on the personal subjective experience of managers, which is easy to make mistakes. Moreover, with the use of database technology and related computer technology in it, the amount of data in enterprises has increased sharply. Finding the information needed by the enterprise in the data information is related to the success or failure of the operation of the enterprise. At this time, the decision tree application based on data mining can analyze and process the data, and provide a basis for the management of human resources scheduling. The purpose of this article is to study the human resource intelligent scheduling algorithm based on the decision tree algorithm, analyze the decision tree algorithm in detail, find the core problem of the decision tree algorithm, integrate human market resources, and solve the key problems in modern human resource management. It provides a reference for the decision tree algorithm in human resource intelligent scheduling. Experiments show that the human resource scheduling based on the decision tree algorithm significantly shortens the cycle project, up to 27.5%. It greatly saves the time spent on human resource scheduling, which meets the expectations of the experiment.**

***Keywords—decision tree algorithm, human resource management, scheduling algorithm, data mining.***

## I. Introduction

With the rapid development of society, human resource scheduling has gradually increased its status and importance. It has become an important part of business management. At the same time, many emerging technologies have appeared in the field of early human resource management and have been widely used [1]. Nevertheless, there are still some problems that cannot be solved by existing technologies in the field of human resource management. People urgently need a new technology to solve the problem. Decision tree algorithm has been widely used in many fields, and has achieved good financial results, opening up new ideas for intelligent human resource planning [2].

In the research of human resource intelligent scheduling algorithm based on decision tree algorithm. Many scholars have studied it and have drawn many effective conclusions. For example, Sonza RL constructs a human resource scheduling model, and at the same time designs and solves optimization algorithms to realize multi-purpose human resource scheduling. Jeong H Y proposed the concept of data warehouse, which solved the problem of data preparation in the early stage of data mining [3]. The combination of Nurkholis A variable grouping and sample grouping significantly improves the classification effect. These scholars' research on human resource scheduling algorithms provides a lot of theoretical basis for this article[4].

This article first analyzes the decision tree algorithm theoretically, and explains in detail the representation of the decision tree, the core issues and the application of the decision tree method [5]. After the decision tree algorithm is deepened, the random forest algorithm is briefly analyzed, and then the decision tree is simplified for the problem of the decision tree is too large, and finally the decision tree generation experiment is carried out. The result of the experiment is in line with expectations. The application of tree algorithm in human resource intelligent scheduling has certain practical significance [6].

## II. Decision Tree Algorithm for Human Resource Intelligent Scheduling

### A. Decision Tree Algorithm

The decision tree is a tree with multiple crisis nodes. In the field of machine learning, the decision tree is a very classic algorithm, which integrates the idea of divide and conquer well. For complex problems, the decision tree divides the initial set into smaller subsets through the retrograde process of the data set, thereby simplifying the complex problem [7]. The root node is the starting point of the decision tree, which is unique in the world and represents the collection of all training samples. The connection node is an internal node, and the sample is divided into several parts according to a specific separation rule. Leaf nodes, also called terminal nodes, represent a set of data with sorting labels [8]. The construction of decision tree is mainly divided into two parts: decision tree creation and pruning. The creation of a decision tree refers to continuously selecting an attribute from the data set according to a specific separation rule to form a new node until a certain interruption condition is met [9]. This can usually be achieved through the use of retrospective methods. Pruning the decision tree is to remove certain sub-trees from the decision tree created under certain conditions, thereby reducing the risk of over-alignment of the algorithm. According to different strategies, it can be divided into pre-pruning and post-pruning [10]. The whole construction process is essentially the process of learning the separation parameters at each node and obtaining the class probability distribution of each leaf node. Therefore, how to determine the optimal characteristics of each node is particularly critical [11].

### B. Random Forest Algorithm

As it is said that a single tree does not make a forest, the "forest" in the random forest refers to the model that contains many decision trees. The randomness in random forest is mainly reflected in two aspects: random selection of samples and random selection of features. Random selection of samples is random sampling with replacement of the given training samples. Random selection of features means that the

978-1-6654-9111-2/22/$31.00 ©2022 IEEE

features of the sample are not all involved in the training of the decision tree, but several features are randomly selected from all the features for learning. The basic process of the algorithm can be summarized as follows: First, use random sampling to extract with replacement in the original data set to form k sample subsets; then construct k decision trees based on the k different training sets, and generate K classification results; finally, the results of these k models are combined to vote to determine the optimal classification result of the sample [12].

### C. Reasons For Too Large Decision Tree

*1) Representation improper:* Obviously, the size of the decision tree is related to the description language. Some expression languages can express the target concept concisely and clearly, while others cannot. Therefore, choosing the correct description language can significantly reduce the complexity of the decision tree.

*2) The data is noisy:* Data noise includes attribute noise (that is, it contains inappropriate attributes) and attribute value noise (that is, the attribute value of the data is incorrect). Since the decision tree cannot distinguish between right and wrong data, it models right and wrong data. The rules of the correct data itself may be overwhelmed by noise, so the decision tree will follow the noise. Its existence becomes bigger.

## III. Experimental Research on Human Resource Intelligent Scheduling Algorithm Based on Decision Tree Algorithm

### A. Hybrid Scheduling Algorithm

Considering that a single scheduling algorithm always has some unavoidable limitations in the actual application process, but each has its own advantages. For this reason, researchers hope to integrate the advantages of various scheduling algorithms to complement each other, so they propose a hybrid scheduling method. The so-called hybrid scheduling is to organically integrate different single scheduling algorithms according to a certain strategy to form a unified whole for scheduling, rather than just using a certain scheduling algorithm. Nowadays, more and more scholars have joined the research of hybrid scheduling algorithms, and various hybrid scheduling algorithms have emerged one after another. The following mainly introduces some commonly used mixing methods:

*1) Weighted mixing:* Assign certain weights to different scheduling algorithms, and then linearly combine the results of these scheduling algorithms to produce the final scheduling result. This method of mixing is usually relatively simple to implement, and various algorithms can be used in a direct manner, thereby enhancing the scheduling effect of the entire system.

*2) Handover hybrid:* Switch to different scheduling methods according to certain heuristic rules according to the actual application scenarios and requirements of the scheduling system, so as to avoid the problems of a particular algorithm. This hybrid approach is usually too sensitive to the advantages and disadvantages of the included scheduling algorithms. In addition, its conversion rules will increase the number of parameters in the scheduling system, making the scheduling process more complicated.

*3) Cross-hybrid:* Run multiple scheduling algorithms at the same time, and aggregate their scheduling results for unified scheduling to users. Since each item finally scheduled to the user can be associated with the scheduling results of multiple scheduling algorithms at the same time, the overall effect after mixing is not easily affected by a single algorithm.

*4) Waterfall type hybrid:* Connect multiple scheduling algorithms in series, and each scheduling algorithm is regarded as a filter, and the recommended results are filtered layer by layer. That is to say, the scheduling algorithm of the latter level will optimize the scheduling result produced by the algorithm of the previous level. The hybrid method has the characteristics of continuous iteration and refinement, and has a relatively high tolerance for noise.

*5) Feature combination:* Extract different features from different data sources, and combine them to apply to the scheduling algorithm.

*6) Predictive fusion method:* Regard the scheduling algorithm as a kind of "predictive algorithm", use the prediction result of the previous layer of scheduling algorithm as a feature, and input it into the algorithm of the next layer for further training to generate the final scheduling result.

*7) Meta-level hybrid:* The scheduling sub-model generated by one scheduling technology is input into another scheduling algorithm as a whole, which can improve the scheduling performance to a certain extent.

### B. Decision Tree Generation Experiment

The algorithm for creating a decision tree can be described as a retrospective process: first select an attribute of the training sample as a node, and create a branch for each possible attribute value. The steps are as follows:

*1)* All training samples of this node belong to the same category;

*2)* There are no remaining attributes that can be used to further divide the sample;

*3)* There is no sample for this branch. At this point, a complete decision tree is formed.

The information entropy required to classify a given sample is shown in formula (1).

$$\inf o(T) = -\sum_{i=1}^{m} p_i \log_2(p_i) \tag{1}$$

Where Pi is the probability that any sample belongs to Ci. On this basis, to measure the performance accuracy of the generation algorithm, the prediction accuracy can be used to evaluate the quality of the algorithm. The formula is shown in (2).

$$RMSE = \sqrt{\frac{\sum_{u,i \in S} (r_{ui} - f_{ui})^2}{|S|}} \tag{2}$$

Among them, S is a given test set, u represents the user, i is the item, the real score is rui, and the predicted score is fui.

## IV. Experimental Analysis of Human Resource Intelligent Scheduling Algorithm Based on Decision Tree Algorithm

### A. Experimental Analysis of Conventional Human Resource Scheduling

On the basis of the previous discussion, further experiments are carried out on human resource scheduling, and a 50-person enterprise is simulated, which is divided into three project teams: enterprise informatization team, e-government project team and campus informatization project team. Among them, the enterprise information team has 26 people, the e-government team has 18 people, and the campus information team has 6 people. First, the traditional and conventional methods are used to conduct human resource scheduling experiments, and the experimental results are shown in Table I.

TABLE I. SCHEDULING CYCLE COMPARISON TABLE

| project type | Regular scheduling period (days) | Decision tree algorithm scheduling period (days) | Cycle shortening rate (%) |
|---|---|---|---|
| Historical items | 55 | 42 | 23.6 |
| Enterprise Information Project | 46 | 35 | 23.9 |
| E-government project | 48 | 36 | 25.0 |
| Campus Information Project | 40 | 29 | 27.5 |

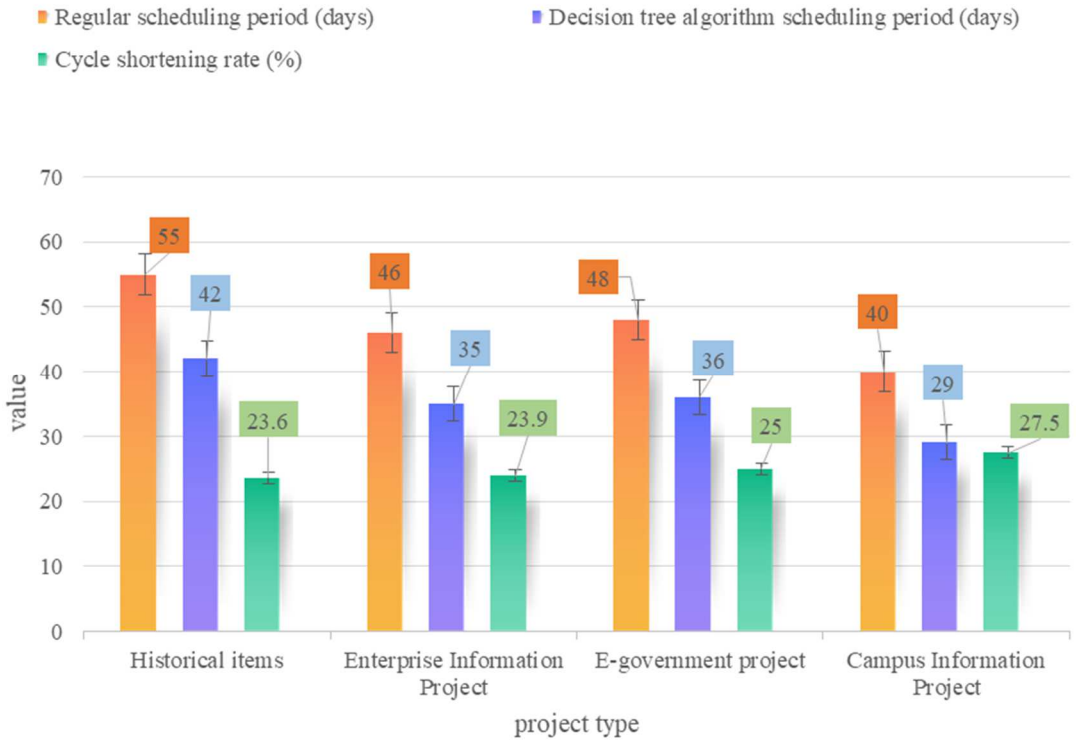


Fig. 1. Comparison of scheduling cycles

It can be seen from Figure 1 that the human resource scheduling based on the decision tree algorithm has significantly shortened the cycle project, the historical project scheduling cycle has been shortened by 23.6%, the enterprise information project scheduling cycle has been shortened by 23.9%, and the e-government project cycle has been shortened by 25%. , The campus informatization cycle has shortened the most time, reaching 27.5%. Greatly save the time spent on human resource scheduling. In order to further explore the impact of the decision tree algorithm on the efficiency of human resource scheduling, on the basis of the cycle experiment, continue to conduct cost experiments to observe the cost impact of the decision tree algorithm on human resource intelligent scheduling. The experimental results are shown in Table II.

TABLE II. COMPARISON TABLE OF SCHEDULING COSTS

| project type | Regular scheduling cost (k) | Decision tree scheduling cost (k) | Cost reduction rate (%) |
|---|---|---|---|
| Historical items | 35 | 24 | 31.4 |
| Enterprise Information Project | 42 | 30 | 28.6 |
| E-government project | 28 | 20 | 28.6 |
| Campus Information Project | 32 | 23 | 28.1 |

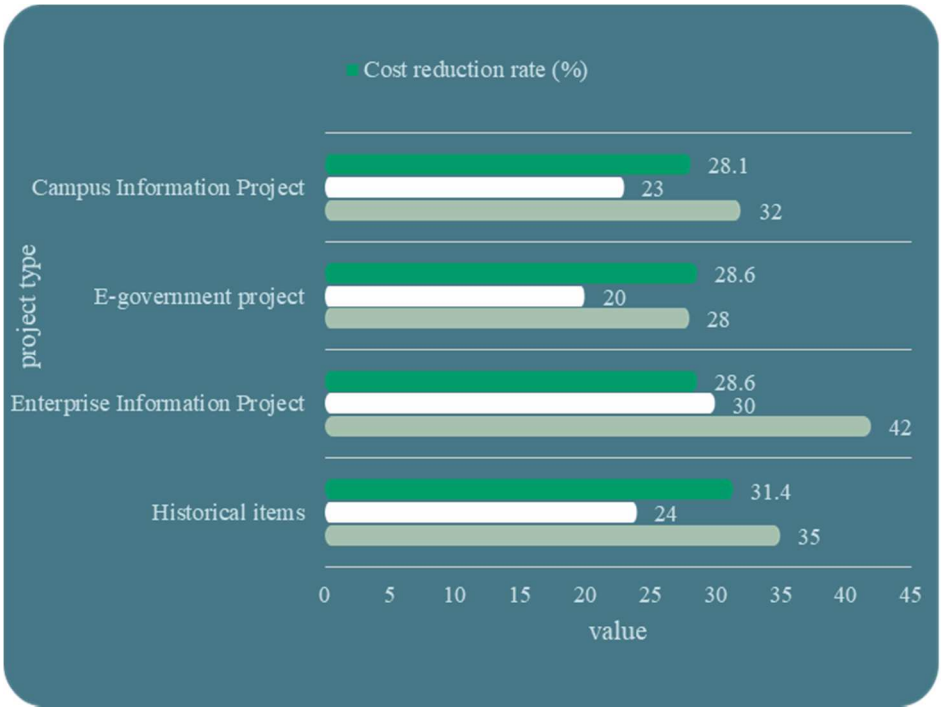


Fig. 2. Comparison table of scheduling costs

Figure 2 shows that the project cost is significantly reduced when the decision tree algorithm is used for human resource scheduling. Among them, the historical project cost reduction rate reached 31.4%, the enterprise information project cost reduction rate was 28.6%, and the e-government project cost reduction rate was 28.6%. , The project cost of campus informatization has been reduced by 28.1%, both at about 30%, which proves that the human resource scheduling of the decision tree algorithm still has a certain practical application value.

## V. CONCLUSIONS

In the era of economic globalization, the importance of human resources has increased rapidly and has become an important resource in enterprises, and its importance has also reached a consensus among various enterprises. A variety of key data mining techniques can assist data analysis and processing, reveal the relationship between data and hidden deep knowledge. The right decision. Therefore, data mining technology can be used to study and explore enterprise management processes, provide valuable guidance for enterprise decision-making, and enable enterprises to obtain greater profits and unique competitive advantages. However, there are still some shortcomings in this article. For example, in the choice of data mining algorithm, this article only chooses the decision tree algorithm. The mining results lack comparison and reference and the amount of data is small, making the algorithm data not accurate enough. Hope it can be improved in follow-up experiments.

## REFERENCES

[1] Sonza R L, Tumibay G M. Decision Tree Algorithm in Identifying Specific Interventions for Gender and Development Issues. Journal of Computer and Communications, 2020, 08(2):17-26.

[2] Jeong H Y, Yoon T S. Analysis of Hepatitis C Virus using Data Mining Algorithm -Apriori, Decision Tree. International Journal on Bioinformatics & Biosciences, 2017, 7(3):01-11.

[3] Nurkholis A, Sitanggang I S. Optimization for prediction model of palm oil land suitability using spatial decision tree algorithm. Jurnal Teknologi dan Sistem Komputer, 2020, 8(3):192-200.

[4] Lalithadevi B. Novel Technique for Price Prediction by Using Logistic, Linear and Decision Tree Algorithm on Deep Belief Network. International Journal of Psychosocial Rehabilitation, 2020, 24(5):1751-1761.

[5] Clarin J A. Academic Analytics: Predicting Success in the Licensure Examination of Graduates using CART Decision Tree Algorithm. Journal of Advanced Research in Dynamical and Control Systems, 2020, 12(1-Special Issue):143-151.

[6] Ayinla, Akinola I B. An Improved Collaborative Pruning Using Ant Colony Optimization and Pessimistic Technique of C5.0 Decision Tree Algorithm. International Journal of Computer Science and Information Security, 2021, 18(12):111-123.

[7] Majumder M. Human Resource Management Practices and Employees' Satisfaction Towards Private Banking Sector in Bangladesh. International Review of Management & Marketing, 2017, 2(1):52-58.

[8] Ones D S, Viswesvaran C, Schmidt F L. Realizing the full potential of psychometric meta-analysis for a cumulative science and practice of human resource management. Human Resource Management Review, 2017, 27(1):201-215.

[9] Cristiani A, Maria Peiro J. Human resource function, unions and varieties of capitalism: Exploring their impact on human resource management practices based on CRANET data. Employee Relations, 2018, 40(6):1072-1098.

[10] Jiang S L, Zheng Z, Liu M. A Preference-Inspired Multi-Objective Soft Scheduling Algorithm for the Practical Steelmaking-Continuous Casting Production. Computers & Industrial Engineering, 2017, 115(JAN.):582-594.

[11] Prati G, Pietrantoni L, Fraboni F. Using data mining techniques to predict the severity of bicycle crashes. Accident Analysis and Prevention, 2017, 101(APR.):44-54.

[12] Buczak A, Guven E. A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection. IEEE Communications Surveys & Tutorials, 2017, 18(2):1153-1176.