A Machine Learning Application for Human Resource Data Mining Problem*

Zhen Xu[1] and Binheng Song[2]

[1] School of Software, Tsinghua University, Beijing, 100084, P.R. China
z-xu03@mails.tsinghua.edu.cn
[2] Dept. Math, Tsinghua University, 100084, P.R. China
bsong@math.tsinghua.edu.cn

Abstract. Apply machine learning methods to data mining domain can be more helpful to extract useful knowledge for problems with changing conditions. Human resource allocation is a kind of problem in data mining domain. It presents machine learning techniques to dissolve it. First, we construct a new model which optimizes the multi-objectives allocation problem by using fuzzy logic strategy. One of the most important problems in the model is how to get the precise individual capability matrixes. Machine learning method by being told is well used to settle the problem in this paper. In the model, appraisal values about employees are saved in knowledge warehouse. Before tasks allocation, machine learning approach provides the capability matrixes based on the existing data sets. Then Task-Arrange or Hungarian Algorithm provides the final solution with our proposed matrixes. After present tasks are finished, machine learning method by being told can update the matrixes according to the suggestions on employees' performance provided by the specialists. Useful knowledge can be well mined in cycles by learning approach. As a numerical example demonstrated, it is helpful to make a realistic decision on human resource allocation under a dynamic environment for organizations.

1 Introduction

As data sets are continuously growing in size and number, there is a need to extract useful information from collections of data. The disciplines of data mining and machine learning are concerned with the application of methods such as clustering, classification, rule induction, and others to potentially large data repositories in order to extract relevant information and eventually convert data and information into knowledge [1, 2].

Machine learning makes the computer be intelligent and deals with the issue of how to build programs that improve their performance at some task through experience. A key research area in machine learning is extracting conceptual description from data samples. Meanwhile, the key purpose of data mining is to search interesting patterns and important rules from large-scale databases. Therefore, many machine

* National 973 Project in China (Grant No. 2004CB719401).

W.K. Ng, M. Kitsuregawa, and J. Li (Eds.): PAKDD 2006, LNAI 3918, pp. 847 – 856, 2006.

learning methods could be used directly into data mining domain. In this paper, we will apply machine learning to human resource allocation, which is one of the problems in data mining domain.

Surveying the past research, there are two types of modeling about Human Resource Allocation. One is linear programming model [3]; the other is goal programming [4]. Linear programming is a single goal optimization technique; this is not the situation for a majority of firms with multiple goals. Goal programming can not deal with the organizational differentiation problems [6]. In addition, neither do the models studied take advantage of the machine learning method.

Currently many resource allocation algorithms have existed. It uses multiprocessors to solve distributed resource allocation in [7], where the processor number is correlated with the task number. The approaches proposed by [8-9] provide algorithms on task scheduling problem. These algorithms do not consider the time constraints among tasks. It limits further application

There are great deals of data about every employee. Before tasks allocation, managers will provide appraisal values aimed to every employee. But the appraisal values each specialist gives are very subjective. Moreover, data is changing according to one's changing experience. Actually, Human resource allocation is a kind of problem in data mining domain. In the process of task allocation, it is crucial to identify and mine useful information from existing data sets. Here we bring forward a machine learning approach to dissolve the allocation problem.

In our paper, a new model is proposed. And we apply fuzzy logic concept into the model. Individual capability matrixes about employees are the data need to be mined. Machine learning methods can offer the matrixes not only based on the previous appraisal data sets the managers provide but also suited to changing conditions. Through the matrixes, we get the fuzzy synthesis appraisal matrixes finally. Then Task-Arrange or Hungarian Algorithm [12] can obtain the best assignment solution based on the matrixes. After the allocation, the specialists feed back the evaluation values which will be saved in the knowledge warehouse. The self-learning process revises the original data in the warehouse. The refreshed matrixes will be offered when next allocation comes. So we can mine useful knowledge based on existing data sets in cycles.

The layout of the paper is as follows: In section 2, we first present our basic model. Section 3 introduces the machine learning method by being told. And then we describe fuzzy concepts and define fuzzy membership sets in Section 4. We give the optimization target in Section 5. Section 6 briefly describes our algorithm which achieves the optimization solution, and a numerical example is also given in section 7 to show how the algorithm works. Section 8 gives the comparison with other algorithms. Finally, we conclude in section 9.

2 The Fuzzy Modeling

The management of human resource includes three aspects: human capital control, tasks control and time control.