

# The Multiple Optimizations for the Physical Fitness Test Data of Campus Based on Repast

Lei Li

Leaders, Yunnan College of Business and Management, Yunnan Kunming, China, 650106

18459423@qq.com

**Keywords:** Physical fitness, Training scheme, Repast, Multiple optimization

**Abstract.** The physical fitness test plays an important role to testify the health level of the campus in our country. These test data is analyzed firstly to guide the training scheme making, which is one of effective approaches to enhance the level of the physical fitness of campus. In this paper, the corresponding data analysis method based on Repast is discussed. The mathematical model for testifying these data is constructed firstly. Then the multiple optimization algorithm is suggested to help the data mining from these data, which aims to form the effective training scheme. The physical fitness data from two universities are used as the experimental data and the corresponding results indicate that the proposed algorithm can obtain the intuitional results to choose the training scheme.

## Introduction

The physical fitness test plays an important role to testify the health level of the campus in our country. These test data is analyzed firstly to guide the training scheme making, which is one of effective approaches to enhance the level of the physical fitness of campus. However, how to deal with these data to obtain better analysis results is one of methods worth to be considered. Meanwhile, the tools to help implementing this analysis is also important to achieve the purpose. During the past two decades, a lot of methods such as regression analysis, no-linear methods, Burg and some other optimization algorithms are used to help analyzing the physical data. Of course, some digital tools as SPSS, Matlab and other statistic software are utilized to implement the ideals of predecessors. Therefore, the trend of this area is to combine the algorithms with software tools to obtain the results rapidly and accuracy.

With the development of the analyzing software, the comprehensive analysis system is spread faster and faster. Such as Repast, by using such system, the complex results can be obtained easily on the multiple dimensions. The structures are constructed firstly and the testifying function is given, when these two factors are already, the system can achieve the results under the given instructions. It means that this system is the latest and easiest software for this multiple analysis tool. In this paper, we use Repast to analyze the physical data to obtain the hided information.

The first step to use this software is to construct the analyzing model and determine the testifying functions. It is the most important work to achieve accurate results. For physical data, the dimension of these data is large which is from the multiple testing contents corresponding to physical testing. In this case, the analyzing model should be suit for the multiple data. During our latest researches, the Multiple Optimization algorithm (MOA) is proposed to resolve the analysis problem for those merged data. Actually, the MOA is similar with the combination of linear optimization and the no-linear optimization. Therefore, the method proposed in this paper contains two aspects of advantages from both linear optimization and no-linear optimization.

In this paper, the data mining method with the MOA for the physical data is proposed. The physical testing items are consisting of the testing vectors firstly. Then the statistic method is employed to obtain the counting vectors describing the distributed features. The MOA is used to analyze the last results. In the implementation of MOA, the Repast is suggested. In our work, we do not analyze the result for the physical education, but only the analysis method is discussed. In section 2, the physical testing items and their corresponding statistic vectors are introduced. In section 3, the description length theory and

MOA are given. In section 4, the proposed algorithm implemented by Repast and the corresponding results are given.

### The Testing Items

In physical testing, some items are the standard contents which are used to testify the physical target of one person roundly. They are 1000m, long jump, shot and shuttle run. Thus, the resulted report for one person is consist of four items, namely, one vector with 4 dimensions is used to describe the data for one person. Actually, for one's physical performance, this vector is sufficient when it is quantized into a standard table. This table is illustrating by Table 1.

Table 1 The standard table for testifying one's physical performance

Testing items	1000m	Long jump	shot	Shuttle run
Standard results	E/P/F	E/P/F	E/P/F	E/P/F

Where symbols E/P/F denote the Excellent, Pass and Fail respectively. This thus the testing vector for one person. If the overall situation will be observed, this vector is actually one row of a matrix with size  $3 \times 4$ . In this case, these three symbols can be considered as the context of the matrix which is used to guide one vector of matrix. Then the testing matrix is obtained. But for amount of persons, its analysis will become more complex. One efficient way to analyze this is to use the statistic method. Namely, by counting the number of E/P/F respectively, the statistic model will be achieved to help the analysis. However, for this statistic model, the trend that these data play is the most important point we pay more attention to. In order to obtain this information, the efficient way is the description length theory. When the description length is calculated, the corresponding possibility can be found more easily.

### The Description Length Theory

Considering a counting vector such as the testing vector above. It is illustrated by using the mathematics way as follows:

$$\mathbf{V} = \begin{Bmatrix} E & P & F \\ n_e & n_p & n_f \end{Bmatrix} \quad (1)$$

Let  $N$  denote the total number equal to  $n_e + n_p + n_f$ . The description length  $L$  of this counting vector can be calculated by (2)

$$L = N * \log N - (n_e \log n_e + n_p \log n_p + n_f \log n_f) + \delta \quad (2)$$

Where the parameter  $\delta$  is the cost for the description length and is calculated by (3)

$$\delta = \frac{\prod_{i \in \{E, P, F\}} n_i}{\sum_{i \in \{E, P, F\}} n_i} \quad (3)$$

Then for the matrix  $\mathbf{M}$ , it is consist of some counting vectors  $\mathbf{v}_i$  as the form

$$\mathbf{M} = \{\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_n\}^T \quad (4)$$

When this matrix is obtained, its corresponding description length is the sum of the description lengths of each counting vector  $v_i$ .

For analysis by using this matrix, the description length implies that possibility that one result happened. If the description length is small, its meaning is that the results are converged to the hoping results. But actually, this calculation is more complex. Namely, although the optimal description length is existing, it is difficult to find it. The efficient way to resolve this problem is to use the searching approaches. In our work, the Multiple Optimization Algorithm is suggested.

### The Multiple Optimization Algorithm and Repast Implementation

There are two kinds of searching atoms in MOA, one is the global searching atoms (GSA) which are used to search the global optimal results, another is the local searching atoms (LSA) used to find the local optimal results. These two types of atoms work alternately to find the global optimal results and the local optimal results respectively. In iterations, the global searching atoms are construct first, then weather one GSA joins in the queue or quit out to the queue is determined by the strategy of survival of the fittest. The LSA works as the same way as the GSA done. The fitness value for each atom are used to modify the position of this atom in the queue by sequencing. The atom with the lowest fitness value is always listed in the tail of the queue. The GSA in the former position of the queue will form more LSA to ensure the optimal results achieved rapidly.

We first construct the whole structure of the searching atoms, each atom denotes a group of weights, the total adaptive code length after weighting is used as the fitness value of this atom. After iterations, the atom with the minimum fitness value is considered as the optimal result. To implement the updating process adaptively, when some symbols are encoded, the MOA execute again to achieve a group of new weights.

The MOA structure is given in Fig. 1

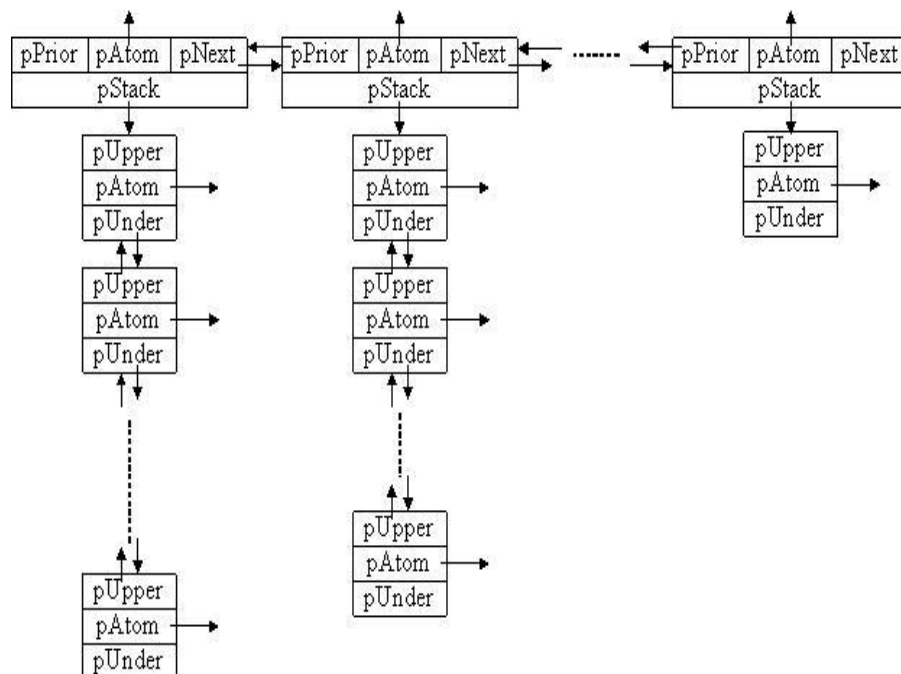


Figure 1. The structure of the MOA

Then combining with the MOA and Repast. We define the corresponding model for Repast by using the description length theory and the structure of MOA. It is easily to construct this program. Due to space limitations, the details of implementation will not give here.

Then the resulted program is used to analyze the physical data of college students from Yunnan.

## Experiments and Results

The number of students is reach to 28213 for our work in this paper. By using the traditional analysis method, the result is near 90% students are reach to Pass level.

In this paper, the same data of these students are used as the testing sources. The number of iterations are set as initializing, 200 and 400. The corresponding results are listed as follows:

For initializing, the state image of Repast is given if Fig. 2

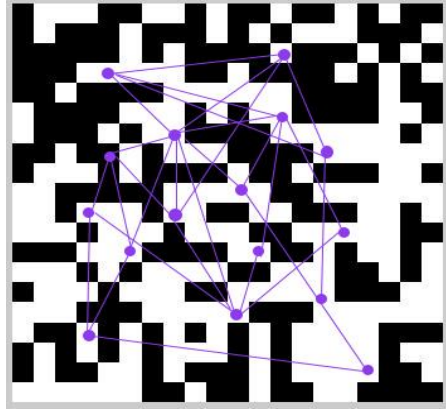


Figure 2. The state image of initializing state

The corresponding results after 200 iterations are given in Fig. 3

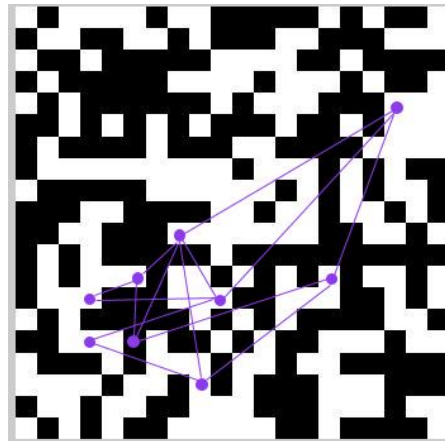


Figure 3. The state after 200 iterations

At last, when the results are converged to the optimized results illustrated by Fig. 4.

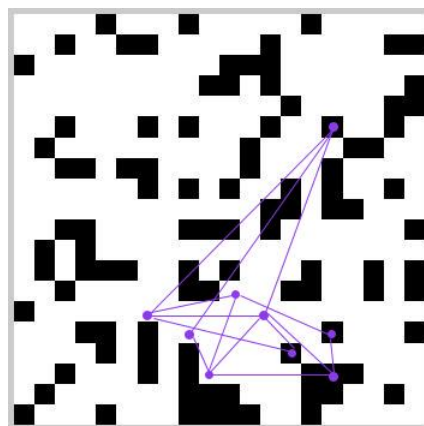


Figure 4. The converged results after 400 iterations

It is obviously that the results are converged. If the number is read, the value of the ratio that these students are near to 91.3%. It implies that the results from the proposed algorithm is near to the results from the traditional method but more accuracy.

## Conclusion

The physical fitness test plays an important role to testify the health level of the campus in our country. These test data is analyzed firstly to guide the training scheme making, which is one of effective approaches to enhance the level of the physical fitness of campus. In this paper, the corresponding data analysis method based on Repast is discussed. The mathematical model for testifying these data is constructed firstly. Then the multiple optimization algorithm is suggested to help the data mining from these data, which aims to form the effective training scheme. The physical fitness data from two universities are used as the experimental data and the corresponding results indicate that the proposed algorithm can obtain the intuitional results to choose the training scheme.

## References

- [1] Jianhua Chen, Yufeng Zhang, Xinling Shi, Image coding based on wavelet transform and uniform scalar dead zone quantizer [J], Signal Processing: Image Communication, vol.21, pp.562-572, 2006.
- [2] X.Wu, P.A.Chou, X.Xue, Minimum conditional entropy context quantization[C], in: Proc. of IEEE Inter. Symposium on Information Theory, Sorrento, Italy, p.43, June 2000.
- [3] Marco Cagnazzo, Marc Antonini, Michel Barlaud, Mutual information-based context quantization [J], Signal Processing Image Communication, vol.25, pp. 64–74, 2010.
- [4] Mahoney, M. Fast Text Compression with Neural Networks. In Proc. FLAIRS, Orlando FL, 2000.
- [5] M. D. Cao, T. I. Dix, L. Allison, and C. Mears, A simple statistical algorithm for biological sequence compression, in Proc. of the Data Compression Conf, DCC-2007, Snowbird, Utah, 2007.
- [6] Armando J. Pinho etc, Bacteria DNA Sequence Compression using a mixture of finite-context models, IEEE Statistical Signal Processing Workshop, Portugal, pp.125-128, 2011.
- [7] X. Li, Niching without niching parameters: particle swarm optimization using a ring topology. Evolutionary Computation, IEEE Transactions on 14(1), 150-169, 2010.
- [8] Min Chen, Jianhua Chen, Mohai Guo, Affinity Propagation for the Context Quantization, Advanced Materials Research, Vols.791, pp 1533-1536, 2013.