Analysis of C-Programming Iteration Type Courseware by Mahalanobis-Taguchi System

Hiroyuki WATANABE[a*] & Susumu SAKANO[a]
[a]*College of Engineering, Nihon University, Japan*
*hiroyuki@ee.ce.nihon-u.ac.jp

Abstract: In this paper, it is proposed for the classification of the Mahalanobis distance of learning data in the C-programming iteration type courseware that a virtual teacher group is used as a standard space and that the membership function is used to convert learning time. Furthermore, a quantitative analysis method of the questions in the courseware is proposed by calculating the SN ratio allocating the Mahalanobis distance into the orthogonal arrays. The validity of the results by these proposed methods is confirmed by the measured values, and the causes for taking a long learning time are cleared.

Keywords: Mahalanobis distance, SN ratio, membership function, courseware

Introduction

Learning data, which consists of scores and learning time, is achieved by a learning management system from the answer of the courseware. Since every learner achieves the passing score in the iteration type courseware, learning time becomes the target of the analysis. In the past, the method by SP-chart or the Mahalanobis-Taguchi System have been reported for the analysis methods considering the cross correlation [1][2]. However, the estimation of learning time which consists of some questions in the courseware has not been reported in both methods. Especially, the later method has a feature that is able to calculate the quantitative value with some combined conditions, but has some problems. The 1st is that the classification of learning data depends on its standard space. The 2nd is that the Mahalanobis distance becomes great for it has an infinite learning time range. The 3rd is that the quantitative analysis method of questions in the courseware isn't known.

In this paper, it is proposed for the C-programming iteration type courseware that learning data of a virtual teacher group is used as a standard space and that the learning time is converted to a limited value by using a membership function, and the Mahalanobis distance of learning data is classified into 2 groups. Furthermore, an analysis method of questions in the courseware is proposed by calculating the SN ratio allocating the Mahalanobis distance into the orthogonal arrays. The validity of the results by these proposed methods is confirmed by the measured values, and the causes for taking a long learning time are cleared.

1. Analysis Methods

The Mahalanobis distance is calculated by the distance between two points: one is measured values and the other is a standard space. In this paper, the standard space is made at random by a virtual teacher group, who answers with a short learning time or high

score. The procedure of the analysis methods by the Mahalanobis-Taguchi System of learning time is shown as the step1~step3 and its score is as the step2~step3.

1.1 Step1: Preprocessing of Learning Time

When t_{ij} is the learning time for question No.j (=1, 2, ... , m) of learner-i (=1, 2, ... , n), t_{ij} is related to thinking time of the question, and takes the range $0 \leqq t_{ij} \leqq \infty$. Therefore, t_{ij} is converted to the value by the following equation using the membership function.

$$v_{ij}(t_{ij}) = \begin{cases} 1 & (t_{ij} \leqq t_{aj}) \\ \dfrac{t_{bj} - t_{ij}}{t_{bj} - t_{aj}} & (t_{aj} < t_{ij} \leqq t_{bj}) \\ 0 & (t_{ij} > t_{bj}) \end{cases} \quad (1)$$

Where, t_{aj} is the smallest learning time of the question No.j in the standard space, and t_{bj} is equal to 2.2 times of the average learning time [1].

1.2 Step2: Calculation of Mahalanobis Distance

When a_{jk} is the reciprocal of the Peason product-moment correlation coefficient between the questions, Mahalanobis distance of learner-i (D_i^2) is calculated by the following equation.

$$D_i^2 = \frac{1}{m} \sum_{j=1}^{m} \sum_{k=1}^{m} a_{jk} V_{ij} V_{ik} \quad (2)$$

Where, V_{ij} is the normalized v_{ij} by the average value of learning data and its standard deviation. The average value of V_j is 0. The average value of D_i^2 is 1.0, which means that the distance from the standard space becomes greater as the value gets bigger.

1.3 Step3: Calculation of SN Ratio

The Mahalanobis distances are allocated as trying the evaluation of learner-i in a courseware into each line of the 2 levels' orthogonal arrays. Where, level-1 and level-2 correspond to “use the question in the courseware” in the orthogonal arrays, and “take it off,” respectively. When the SN ratio (η_j) indicating the smaller is the better response is used for trying the evaluation, it is calculated by the following equation.

$$\eta_j = -10 \log\left(\frac{1}{n} \sum_{i=1}^{n} D_i^2\right) \quad \textbf{(3)}$$

The graph of factorial effect can be made by allocating η_j into the orthogonal arrays.

2. Analysis Results

2.1 Analysis of Learning Data by Mahalanobis Distance

Every leaner answers a question which consist of the fundamental questions (No.A~No.D) and the advanced questions (No.E~No.I) in the C-programming course until he/she achieves the passing score (more than 80%) in each question by using a learning management system [1]. The number of total questions is m=9, the number of virtual teachers is 27 (=3×m), and the number of learners is n=54.

Fig.1 shows the Mahalanobis distances of learning time and scores. The score in each question is calculated by (the sum of score in iteration numbers) / (iteration numbers). The dashed lines show the boundary of the classification by cluster analysis. A short learning

time group is 31%, and they are in the high score group. On the other hand, a long learning time group is 69%, 2/3rd of them take a few iteration numbers with a long learning time, and the rest 1/3rd takes many iteration numbers with a short learning time. Since some learners take the value of more than 2000 in the Mahalanobis distances of learning time, it means that the valid result is obtained using the approximation by Eq.(1).

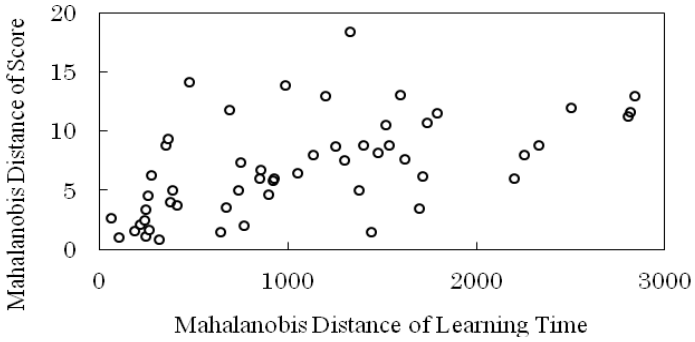

Fig.1: Mahalanobis distances of learning time and scores

2.2 Evaluation of Courseware by Graphs of Factorial Effect

Fig.2 shows the graph of the factorial effect of learning time. It shows that the left side (● mark) is the level-1 in the 2 levels' L_{12} orthogonal arrays, and that the right side (○ mark) is in its level-2. The SN ratio of this courseware is μ_{all}^{1}=-29.6db. This value is almost equal to -30.0db calculated by the average value (1004) from the Mahalanobis distance of learning time in Fig.1 and Eq.(3). The SN ratio of the optimum condition, which consists of the questions No.A, C, H in level-2, and the other questions No. in level-1, is calculated as μ_{opt}^{1}= -25.6db, and the gain is μ_{opt}^{1}-μ_{all}^{1}=+4.0db. The factor of the long learning time's group in Fig.1 is the fundamental questions, since the sum of SN ratio in its level-1 is smaller than that of advanced questions. Furthermore, the graph of the factorial effect of score shows that low-score questions are the questions No.C, G.

Table 1 shows the average learning time in each question No. When the question No. of the SN ratio in Fig.2 is downward to the left, the average learning time of all learners is more than 7 times the one of virtual teachers in Table 1. The cross correlation in Table 1 isn't considered as it is in Fig.2, but the results are almost valid. No.A(hex) takes a long learning time for it is the first question in the courseware. No.C(while) takes a long learning time for they take many iterations within a short time. No.H(pointer) needs a long thinking time, but they don't take many iterations.

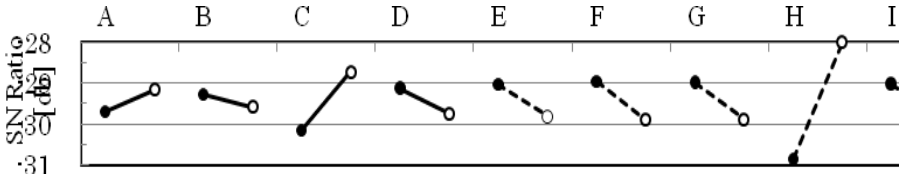

Fig.2: Graph of the factorial effect of learning time

Table 1: Average learning time in each question No. [sec]

Question No.	A	B	C	D	E	F	G	H	I
All learners/Virtual teachers	198/28	168/25	172/13	105/21	134/23	83/21	60/15	263/33	78/24

3. Conclusions

For the analysis of learning data of the C-programming iteration type courseware by the Mahalanobis-Taguchi System, it was proposed that the methods where a virtual teacher group is used as a standard space, and learning time is converted into a limited value by a membership function. The quantitative analysis method that the questions influence in the courseware was proposed by allocating the SN ratio into the orthogonal arrays. By these methods, the SN ratio which consists of many questions in the courseware was able to be estimated, and the causes for taking a long learning time were cleared.

References

[1] Hiroyuki, W., & Katsuhiro, K. (1999). An analysis of response time distribution for CAI courseware. *ICCE*, 2, 125-132. Chiba, Japan.

[2] Elizabeth, A. C. et al. (2007). An evaluation Mahalanobis-Taguchi System and neural network for multivariate pattern recognition. *Journal of Industrial and Systems Engineering*, 1(2), 139-150.