E-Testing Construction Support System with some Prediction Tools

Pokpong Songmuang, Masahiro Ando, Masahito Nagamori, Maomi Ueno and Toshio Okamoto
Graduate School of Information Systems
The University of Electro-Communications
pokpong@ai.is.uec.ac.jp

Abstract

This paper proposes an e-testing construction support system (eTCSS) with the prediction tools for the constructed test. This paper performs some comparison experiments to find the best predictive performances models for the Predictive response-time distribution and the Predictive response-time distribution. Furthermore, the amount of test information based on the Item Response Theory, which is important to improve the measurement efficiency of the constructed test, is applied to be the prediction tool of the eTCSS. Finally, this paper evaluates the system by using actual data. The results show the effectiveness of the system.

1. Introduction

This paper focuses on the e-testing construction support system (eTCSS) which is one of advantages of the e-testing, but there are not so many researches concerning the eTCSS. In some previous papers, they suggested employing the mathematical models as prediction tools (score distribution and response-time distribution) to improve the test efficiency ([1] and [2]). Although the prediction tools of the systems ([1] and [2]) have been developed, the employed mathematical models were not evaluated to find the high predictive performance models for the prediction tools. In addition, it is well known that the Item Response Theory (IRT) has features to describe the relationship between examinees abilities and items responses which important to improve the measurement efficiency of the constructed test.

The goal of this paper is to propose the eTCSS with the prediction tools to support the e-testing construction. The prediction tools are as follows: 1) Predictive response-time distribution (PRD), 2) Predictive score distribution (PSD), and 3) Predictive amount of test information (PATI) based on the IRT. This paper performs some comparison experiment to find the best predictive performance models for the PRD and the PSD of the eTCSS. Moreover, the IRT is applied to be the prediction tool to visualize the amount of test information to improve the measurement efficiency of the constructed test.

Finally, the effectiveness of the proposed system is evaluated by comparing the average test information of the constructed tests with and without the prediction tools. As the results, the prediction tools can improve the measurement efficiency while the time required of the test constructions with the prediction tools is decreased.

2. The Prediction tools of the eTCSS

The eTCSS is developed to support the test construction with the prediction tools, as shown in Figure1. The prediction tools of the eTCSS are developed to interactively visualize the current status of the constructed test to improve the measurement efficiency of the constructed test. The prediction tools are as follows: 1. The PRD visualize the response-time distribution of the constructed test, as shown in the left bottom part of the Figure1, 2. The PSD is developed to assist the test-authors to visualize current score distribution of the constructed test, as shown in the middle bottom part of the Figure1, and 3. The PATI is illustrated in the right bottom part of the Figure1. The test information based on IRT can be used to indicate the measurement efficiency of test which important to test construction. The efficiency test should have high information on the ability level that would be measured. The most widely used IRT models are logistic, which include the two-parameter logistic model (2PL). Therefore, the test information function based on the two-parameter logistic model (2PL) is applied to the PATI.
3. Comparison of Predictive models

This paper compares four predictive models for the PRD. There are the Gamma distribution, Normal distribution, Lognormal distribution and Weibull distribution. The sample data is the Japanese Test history data of 87 examinees. In the experiment, the data is randomly divided into halves; one is training data and the other one is test data (validation data), and the dividing processes are repeated 200, 500 and 1000 times. The efficiency is examined using a root mean square error (RMSE) and standard deviation (SD) between the train data and the test data. The results of the RMSE and the results of the SD of the response-time data of the sample data are shown that the Gamma distribution performs the best predictive performance. Therefore, the Gamma distribution is applied to the PRD of the eTCSS.

The four models in the comparison for the PSD are a binomial distribution model, the mixture model of several binomial distributions, a beta-binomial distribution model and a mixture model of several beta-binomial distributions. The sample data is the score data and the process of the evaluation is same as the previous comparison. The results of the SD and the results of the RMSE of the sample data are shown that the mixture model of several binomial distributions provides the best performance. As the result, the mixture model of several binomial distributions is selected to be the PSD of the eTCSS.

4. System Evaluations

The purposes of this evaluation was to evaluate the prediction tools efficiency by comparing the average information of the constructed tests (which indicates the improving of the tests in aspect of measuring examinees’ ability) and the time required for the test constructions. The test-authors who participated in this evaluation are twenty graduate students of Engineering School. Before the evaluation, we prepared 190 items in the item pool with response data.

The results show that the average information of the tests constructed with the prediction tools is higher than one of the tests constructed without the prediction tools. This paper also compares the average time required of the test constructions with and without the prediction tools. The results show that the test constructions without the prediction tools have average time required more than test construction with tools. From the results, we can say that the prediction tools are effective to make the measurement efficiency of the constructed test increase while the test construction time required is decreased. Therefore, test-authors can construct high measurement efficiency test in short time by the proposed system.

5. References