K-036

Design and Development of e-Testing Construction Support Agent

Pokpong Songmuang †, Masahiro Ando †, Masahito Nagamori †, Maomi Ueno †

Abstract

This paper proposes an e-testing construction support agent based on the Item Response Theory (IRT) in order to improve the examinee's ability measurement efficiency of the constructed test. The agent provides the following two features; 1) the agent recommend item to add into the constructed test from the item pool and to delete from the constructed test based on the maximum information criterion at the examinee's ability level which is expected to be measured, and 2) the monitoring for improvement of the average amount of test information of the constructed test which indicates the test measurement efficiency.

1. Introduction

According to the growth of the Internet, the computer-based testing has been extended to web-based testing which is often called "e-testing". The e-Testing Construction Support System (eTCSS) is one of the e-testing advantages, but there are not so many researches concerning such the system. Thus, this paper focuses on 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 measurement efficiency ([1] and [2]). Although the prediction tools of the systems ([1] and [2]) have been developed, the Item Response Theory (IRT), which has features to indicate the measurement efficiency of test, was unemployed.

The goal of this paper is to propose the agent to support the e-testing construction in the eTCSS. The agent has two features which are designed based on the IRT as following;

1.) The first feature is the item recommendation based on the maximum information criterion at the examinee's ability level which is expected to be measured. In order to maximize the information of the constructed test, the agent recommends the appropriate item from the item pool which should be added into the constructed test and the agent recommends the item which should be deleted from the constructed test.

2.) The second feature of the agent is to assist test-authors to monitor the improvement of the average amount of test information of constructed test and report to the test-authors.

Finally, the effectiveness of the proposed agent is evaluated by comparing the average amount of test information of the constructed tests with and without the agent. Higher level of the test information indicates improvement of the measurement efficiency of the constructed test. As the results, the agent can improve the measurement efficiency of the constructed test.

2. e-Testing Management System

The e-Testing Management System (eTMS) has been developed [3]. The eTCSS is the main features of the eTMS which supports test-authors construct test with the two prediction tools and with the agent as shown in Figure1. The frame 1 (Figure1) shows the constructed test details (test name, test id, and test-authors name). The frame 2 (Figure1) is the items list of the item pool, this feature facilitates the test-authors to evaluate each item difficulty (which is calculated according to rates of correct answers of history data) by illustrate the item IDs in different colors. The frame 3 (Figure1) illustrates the selected items properties; the items difficulty, the selected frequencies and the average of response-time. The tools in frame 3 (Figure1) are provided to delete items and determine the item orders. The prediction tools of the eTCSS are shown in frame 5 of the Figure1. The prediction tools are as follows: 1) predictive response-time distribution, 2) predictive score distribution. The frame 4 (Figure1) is discussion board which the agent uses to communicate with the test-authors. The details of the agent are described in the next section.

†The University of Electro-Communications

3. Agent

At the first step of test construction process, the test-authors have to define the examinee's ability level which is expected to be measured. This ability level will be calculated in the two agent features (the item recommendation and the monitoring of the average amount of constructed test information). Both of the agent features are developed base on the two-parameter logistic model which is one of the popular models in the IRT. The details of the agent features are described as follows:

1) The agent uses the maximum-information criterion to recommend the remaining items in the item pool which appropriate to add into the constructed test and the selected items in the constructed test which should be deleted from the constructed test. The recommended items are ignored in the recommendation process. The test-authors can select or ignore the recommended items.

To describe the recommendation process, the following notation and concepts are need. The items in the pool are denoted by i=1,...,I, where as the rank of the recommended items which should be add in to the constructed test are denoted by x=1,...,X and the rank of the recommended items which should be add in to the constructed test are denoted by y=1,..., Y. Thus, i_x is the index of item in the pool that is administered in the recommendation process and the i_y is the index of item in the constructed test that is administered in the recommendation process. The selected items in the constructed test are denoted by k=1,...,K. The set of recommended items is $S_x=\{i_1,...,i_{x-1}\}$, $S_y=\{i_1,...,i_{y-1}\}$ and the set of selected items is $S_k=\{i_1,...i_k\}$. The set of the remaining items in the item pool after x-1 items have been recommended and k items have been selected is $R_x=\{1,...,I\}/(S_k \cup S_{x-1})$. The set of the remaining items in the

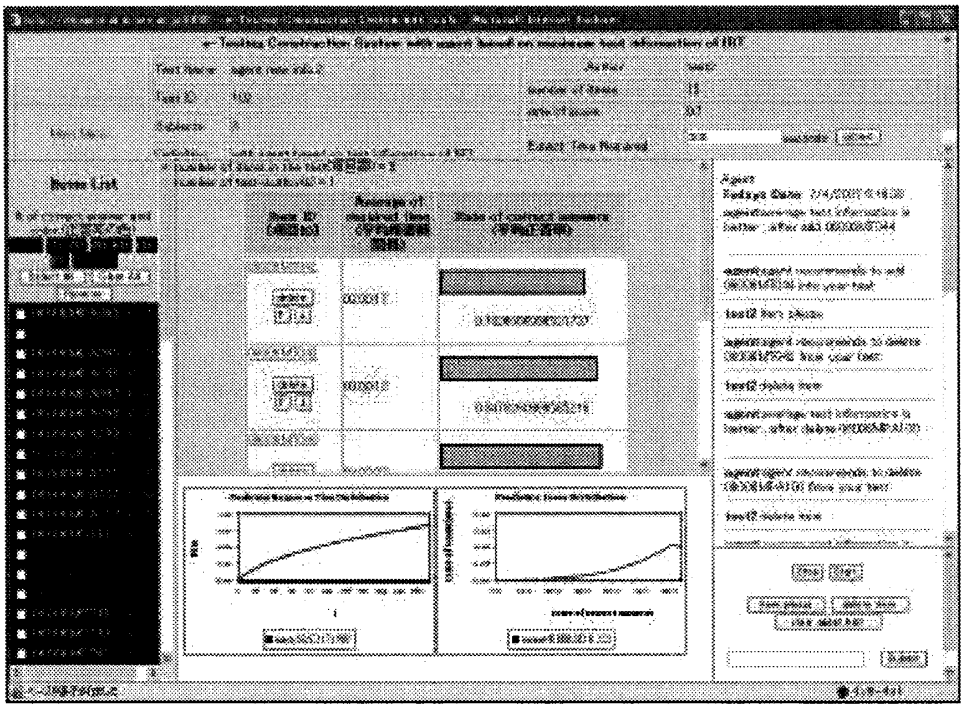

Figure1 e-Testing Construction Support System

constructed test after y-1 items have been recommended and k items have been selected is $R_y=S_k-S_{y-1}$.

The function of agent to recommend item from the item pool which should be added into the constructed test is given as;

$$i_x \equiv \arg\max_j \{I_{S_k+\{j\}}(\theta) : j \in R_x\}$$

where θ is the examinee's ability level which is expected to be measured. The function of agent to recommend item of the constructed test which should be deleted is given as;

$$i_y \equiv \arg\min_j \{I_{S_k-\{j\}}(\theta) : j \in R_y\}$$

2) The agent assists test-authors to monitor and compare the average mount of test information between before and after add/delete item of the constructed test.

The test-authors send the requests to agent via the discussion board and the agent recommendation massages of the both agent features are illustrated in the frame 4 (Figure1).

4. System Evaluations

The purposes of this evaluation were to evaluate the agent efficiency by comparing the average information of the constructed tests (which indicates the improving of the tests in aspect of measuring examinee's ability) between with and without the agent. The test-authors who participated in this evaluation are twenty graduate students of Engineering School. Before the evaluation, we prepared 80 items in the item pool with response data.

The results in Figure2 show that the average amount information of the tests constructed with the agent is higher than one of the tests constructed without it. Furthermore, the average times required of the test constructions with the agent are lower than with out the agent.

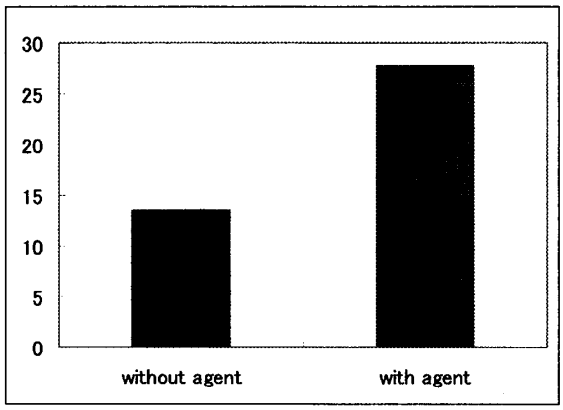

Figure 2 Average amount of test information

The results of the questionnaires, which were provided to evaluate the system, confirm that the agent support the test-authors to construct test base on the maximum information criterion.

Reference

[1] K. Nagaoka, "On Development of Computerized Testing System for Practical Use with Estimation Function", Japan Journal of Education Technology, 24(1), 2000, pp.63-72.

[2] M. Ueno, "Web-based computerized testing system for distance education", Educational Technology Research, 28, 2005, pp. 59- 69

[3] P. Songmuang & M. Ueno, "e-Testing Management System", In G. Richards (Ed.), Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2005, 2005, pp. 3139- 3148