# Clique Algorithm to Minimize Item Exposure for Uniform Test Forms Assembly

Takatoshi Ishii[1](✉) and Maomi Ueno[2]

[1] Tokyo Metropolitan University, Tokyo, Japan
t_ishii@tmu.ac.jp
[2] University of Electro-Communications, Tokyo, Japan
ueno@ai.is.uec.ac.jp

**Abstract.** Educational assessments occasionally require "uniform test forms" (or parallel test forms), in which each test form consists of a different set of items, but the forms are equivalent (i.e., equivalent quality based on test information function of item response theory). However, the construction of uniforms tests often suffers bias of item exposure frequency. Ideally, the item exposure frequency should have a uniform and low distribution. For this purpose, we propose a clique algorithm for uniform test forms assembly with low item exposure. We formalize this test assembly as a searching the clique that has minimum item exposure in the maximum cliques. As the results, the proposed method utilizes the item pool more efficiently than traditional methods do. We demonstrate the effectiveness of the proposed method using simulated and actual data.

**Keywords:** Uniform tests assembly · Item exposure control · Clique problem

## 1 Introduction

ISO/IEC 23988:2007 provides global standards on the use of IT to deliver assessments. For high-stakes test, this standard is recommending to use "uniform test forms", in which each test form consists of a different set of items, but the forms are equivalent (i.e., equivalent quality based on test information function of item response theory). In practical use, the test administrator is required to assemble as many tests as possible for saving test security. To increase the number of assembled test, some test assembly methods allow that any two tests in uniform tests can include common items less than user allowed as test constraint (and this situation is called overlapping condition). However, these methods on overlapping condition do not control how many times each item used in assembled uniform test forms. Accordingly, item use count of each item is not equivalent (e.g., one item is included in 15% tests but another item is included in only less than 1% tests). This exposure deviation of items is called item exposure problem [6]. This problem decreases reliability of items and tests. Ideally, the item exposure of each item should have a uniform, and be as low as possible.


C. Conati et al. (Eds.): AIED 2015, LNAI 9112, pp. 638–641, 2015.
DOI: 10.1007/978-3-319-19773-9_80

To reduce this problem, we propose a clique algorithm for the uniform test forms assembly with low item exposure. We formalize the test assembly as searching the clique that has minimum item exposure in maximum cliques. As the results, the proposed method assembles uniform tests with less item exposure, and more number of uniform tests than the traditional methods [2–5] do. Thus, the proposed method utilizes the item pool more efficiently than traditional methods do. Finally, we demonstrate the effectiveness of the proposed method using simulated and actual data.

# 2 Clique Problem for Uniform Test Assembly with Low Item Exposure

## 2.1 Clique Problem for Uniform Test Assembly with Low Item Exposure

To assemble the uniform test assembly with low item exposure, we formalized test assembly as a clique problem. The first, we define the uniform tests. We employ the definition of "uniform test forms" that is the same as traditional methods (e.g., [2–4]). Accordingly, the uniform test forms is a clique in the following graph:

$$V = \left\{ \begin{array}{c} s : s \in S, \text{“Feasible test form”}, s \text{ satisfies all test constraints} \\ \text{excepting the overlapping constraint from a given item pool} \end{array} \right\}$$

$$E = \left\{ \begin{array}{c} \{s',s''\} : \text{The pair of } s' \text{ and } s'' \text{ has fewer common (overlapping) items} \\ \text{than the allowed number in the overlapping constraint} \end{array} \right\}.$$

The next, we define item exposure. The exposure $Ex_i$ of item $i$ in uniform tests $U$ is formalized as follows:

$$Ex_i = \sum_{r=1}^{|U|} x_{i,r}, \tag{1}$$

where, $x_{i,r} = 1$ denotes that $i$-th item is selected into $r$-th test form, and $x_{i,r} = 0$ denotes otherwise. Then, we define item exposure $Ex_U$ of uniform test forms $U$ as follows:

$$Ex_U = \max(Ex_1, \cdots, Ex_r, \cdots, Ex_{|U|}). \tag{2}$$

Thus, the item exposure $Ex_U$ of uniform tests $U$ is the maximum exposure of items $Ex_i$ in the uniform tests $U$. Accordingly, the minimizing the item exposure $Ex_U$ is Min-Max for exposure of items $Ex_i$. Therefore, the minimizing the item exposure $Ex_U$ minimizes the worst exposure of item in the uniform tests $U$.

The last, we define the uniform test assembly with low item exposure. We formalize this uniform test assembly as the following clique problem:

$$\textbf{minimize} \frac{Ex_C}{|C|} \tag{3}$$

$$\textbf{subject to} \forall v, \forall w \in C, \{v,w\} \in E \tag{4}$$

The eq:4 restrict $C$ as clique. Increasing $|C|$ decreases term of $\frac{Ex_C}{|C|}$ in eq:3. And decreases $Ex_C$ decreases term of $\frac{Ex_C}{|C|}$. Thus, the eq:3 optimize the number of tests $|C|$ and the item exposure $Ex_C$, simultaneously and indirectly. We call the term $\frac{Ex_C}{|C|}$ as "item exposure rate".

### 2.2 Algorithm for Uniform Test Assembly with Low Item Exposure

We propose an approximate algorithm for uniform test assembly with low item exposure. This algorithm consists of the four steps.

Algorithm for Uniform Test Assembly with Low Item Exposure

**Step 1:** (initialize)
Step 1 sets Clique $C$ as assembled uniform test forms by our previous work [2], and clique set $C_{candidate}$ as empty ($C_{candidate} = \phi$).

**Step 2:** (Add step)
Step 2 assembles feasible test, and adds it to the current searching clique $C$. To assemble feasible test, Step 2 solves following optimization problem:

$$\textbf{maximize} \sum_{i=1}^{n} \lambda_i x_i \tag{5}$$

$$\textbf{subject to} \sum_{i=1}^{n} x_{i,r} x_i \leq \text{(overlapping constraint)} \tag{6}$$

$$(r = 1 \cdots |C|)$$

where $x_i = 1$ denotes that $i$-th item is selected into the assembling test form, and $x_i = 0$ denotes otherwise. All other test constraints (e.g., the information amount based on Item Responce Theory, average of answering time, and so on) are included in the constraints. Therein, coordinates $\lambda_1, \lambda_2, \ldots, \lambda_n$ denote random variables distributed uniformly on $[0, 1]$. $\lambda_i (0 \leq i \leq n)$ are resampled each problem is solved. ( This formulation is generation of [1] to assemble to assemble uniform tests on overlapping condition. ) While this problem has solution, Step 2 repeatedly solves this test assembling problem using LP solver (e.g., ILOG CPLEX ). And Step 2 adds the solution of this problem to the current uniform test forms: clique $C$. If the current clique $C$ is changed, Step 2 adds the current clique $C$ to clique set $C_{candidate}$.

**Step 3:** (Delete Step)
Step 3 deletes tests from the current clique $C$ for reduce item exposure $Ex_C$. The deleted test includes the item $i$ with maximum exposure $Ex_i = Ex_C$. By repeating deletion of test, Step 3 reduces the current clique size to 90% of it self. If the current clique is changed, Step 3 adds the current clique to clique set $C_{candidate}$. If the computation time is less than given time, then jump to Step 2

**Step 4:** (Output)
Step 4 finds out the clique that has minimum item exposure rate $\frac{Ex_C}{|C|}$ from clique set $C_{candidate}$, and outputs it.

## 3 Experiments and Results

To demonstrate the performances of the proposed method, we compared the number of assembled test forms and the item exposure with proposed method and the traditional methods ( "BST" : [5], "GA" : [4], "BA" : [3], and "RndMCP" : [2]). For this comparison, we used ILOG CPLEX 11.0 for solving the liner programming in [5] and the proposed method.

We used an actual item pool that was used for the Synthetic Personality Inventory (SPI) examination, which is a popular aptitude test in Japan. In addition, we used three simulated item pools with total numbers of items $I = 500, 1000$, and 2000. The parameters of items in the simulated item pools were set to have similar specification to the actual items. We set the test constraints as (1) the test includes 25 items, and (2) the numbers of allowed overlapping items are 0, 5 and 10.

**Table 1.** The Number of Assembled Test Forms and Item Exposure

| Pool Size | OC | BST | | | GA | | | BA | | | RndMCP | | | Proposal | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\lvert C\rvert$ | $Ex_C$ | $\frac{EX_C}{\lvert C\rvert}$ | $\lvert C\rvert$ | $Ex_C$ | $\frac{EX_C}{\lvert C\rvert}$ | $\lvert C\rvert$ | $Ex_C$ | $\frac{EX_C}{\lvert C\rvert}$ | $\lvert C\rvert$ | $Ex_C$ | $\frac{EX_C}{\lvert C\rvert}$ | $\lvert C\rvert$ | $Ex_C$ | $\frac{EX_C}{\lvert C\rvert}$ |
| 500 | 0 | 12 | 1 | 8.3% | 3 | 1 | 33.3% | 5 | 1 | 20.0% | 10 | 1 | 10.0% | 17 | 1 | 5.9% |
| | 5 | 20 | 6 | 30.0% | 23 | 5 | 21.7% | 96 | 16 | 16.7% | 4380 | 362 | 8.3% | 10736 | 823 | 7.7% |
| | 10 | 20 | 12 | 60.0% | 21 | 7 | 33.3% | 107 | 15 | 14.0% | 99983 | 12995 | 13.0% | 100975 | 13141 | 13.0% |
| 1000 | 0 | 21 | 1 | 4.8% | 4 | 2 | 50.0% | 6 | 2 | 33.3% | 17 | 1 | 5.9% | 33 | 1 | 3.0% |
| | 5 | 40 | 14 | 35.0% | 17 | 5 | 29.4% | 104 | 11 | 10.6% | 46305 | 3399 | 7.3% | 48055 | 3504 | 7.3% |
| | 10 | 40 | 21 | 52.5% | 19 | 4 | 21.1% | 105 | 14 | 13.3% | 100000 | 8705 | 8.7% | 101000 | 8630 | 8.5% |
| 2000 | 0 | 53 | 1 | 1.9% | 8 | 1 | 12.5% | 12 | 1 | 8.3% | 32 | 1 | 3.1% | 70 | 1 | 1.4% |
| | 5 | 80 | 27 | 33.8% | 22 | 4 | 18.2% | 104 | 6 | 5.8% | 96876 | 3935 | 4.1% | 97826 | 3829 | 3.9% |
| | 10 | 80 | 43 | 53.8% | 23 | 4 | 17.4% | 103 | 7 | 6.8% | 100000 | 4013 | 4.0% | 100957 | 4014 | 4.0% |
| 978 (actual) | 0 | 24 | 1 | 4.2% | 9 | 1 | 11.1% | 9 | 1 | 11.1% | 16 | 1 | 6.3% | 31 | 1 | 3.2% |
| | 5 | 39 | 10 | 25.6% | 283 | 23 | 8.1% | 371 | 23 | 6.2% | 40814 | 2177 | 5.3% | 44105 | 2163 | 4.9% |
| | 10 | 39 | 13 | 33.3% | 286 | 22 | 7.7% | 381 | 24 | 6.3% | 100000 | 5598 | 5.6% | 101000 | 5274 | 5.2% |

Table 1 shows the number of test forms $|C|$, the item exposure $Ex_C$, and the item exposure rate $\frac{EX_C}{|C|}$, for the proposed method and traditional methods.

In any case, proposed method assembled more number of test forms $|C|$ , lower item exposure $Ex_C$, and lower item exposure rate $\frac{EX_C}{|C|}$ than traditional methods[2–5] did.

## 4 Conclusion

We proposed a uniform test assembly method to maximize the number of uniform test forms, and to minimize the item exposure. To archive this, the proposed method minimizes the item exposure rate = (item exposure)/(the number of tests).

To demonstrate the performance of proposed method, we conducted an experiment using simulated and actual data. The result was summarized that the proposed method assembled a greater number of uniform test forms with lower item exposure rate than the traditional methods did. Future work will include (1) revealing the feature of our method such as how long time cost and how much space cost does the proposed method requires and etc, and (2) assessing this method in various situations.

## References


1. Belov, D.I.: Uniform test assembly. Psychometrika **73**(1), 21–38 (2008)
2. Ishii, T., Songmuang, P., Ueno, M.: Maximum clique algorithm and its approximation for uniform test form assembly. IEEE Transactions on Learning Technologies **7**(1), 83–95 (2014)
3. Songmuang, P., Ueno, M.: Bees algorithm for construction of multiple test forms in e-testing. IEEE Transactions on Learning Technologies **4**, 209–221 (2011)
4. Sun, K.T., Chen, Y.J., Tsai, S.Y., Cheng, C.F.: Creating irt-based parallel test forms using the genetic algorithm method. Applied Measurement in Education **2**(21), 141–161 (2008)
5. van der Linden, W.J.: Liner Models for Optimal Test Design. Springer (2005)
6. Wainer, H.: Educational Testing Service: CATS: Whither and Whence. Educational Testing Service (2000)