

Т. В. Лобова, А. Н. Ткачев (Новочеркасск, ЮРГТУ (НПИ)).
Многокритериальная оптимизация контрольно-измерительных материалов для тестирования.

В настоящее время технологии тестового контроля обучаемых находят широкое применение в учебном процессе на всех уровнях образования, используются для диагностики текущей успеваемости, а также при итоговой аттестации, например, в форме ЕГЭ и при аккредитации образовательных учреждений высшего и среднего профессионального образования. В связи с этим задача повышения надежности оценки уровня подготовки обучаемых с использованием технологий тестирования является актуальной и требует разработки специальных процедур формирования контрольно-измерительных материалов (КИМ).

Рассматривается следующая задача проектирования КИМ. Пусть тестирование проводится по отдельной дисциплине (возможно, теме), содержащей m дидактических единиц (ДЕ). Формирование тестовых заданий осуществляется путем выбора конкретных тестовых заданий (ТЗ) из сформированного заранее банка. Пусть n_j — число тестовых заданий, соответствующих j -й ДЕ ($1 \leq j \leq m$), n_{0j} — максимально допустимое число ТЗ по этой ДЕ. Примем сквозную нумерацию для содержащихся в банке ТЗ в соответствии с порядком следования ДЕ. При этом будем считать, что отдельные ТЗ могут иметь комплексный характер и содержательно использоваться для одновременного тестирования в рамках нескольких ДЕ.

Предположим, что для формирования КИМ привлекаются высококвалифицированные преподаватели, которые должны ответить на вопрос, следует ли включать каждое из имеющихся ТЗ в КИМ или нет, а также оценить степень полезности и надежности принятия такого решения. Принимаемые экспертами решения идентифицируются с помощью следующих показателей: $\Theta_{ij} = 1$, если i -е ТЗ следует использовать для оценки подготовки обучаемых (в противном случае $\Theta_{ij} = 0$); $p_{ij} \in [0, 1]$ и $q_{ij} \in [0, 1]$ отражают степень уверенности в принимаемом решении о включении с вероятностью p_{ij} или исключении с вероятностью q_{ij} ; $a_{ij}^{(1)}, a_{ij}^{(2)} \in [0, 100]$ описывают качество и степень полезности i -го ТЗ для тестирования по j -й ДЕ.

Далее вероятности p_{ij} , q_{ij} и показатели $a_{ij}^{(1)}$, $a_{ij}^{(2)}$ усредняются для каждого тестового задания с учетом мнений всех экспертов. После этого, учитывая вероятностный характер принимаемых решений, находятся математические ожидания $A_{ij}^{(1)}$, $A_{ij}^{(2)}$ выигрышей, достигаемых при включении i -го ТЗ в КИМ или его исключении. В результате для каждого i -го задания формулируется платежная матрица выигрышей размером $2 \times m$, которой соответствует функция выигрыша $Z_i(u_i, j) = A_{ij}^{(1)} u_i + A_{ij}^{(2)} (1 - u_i)$, $1 \leq j \leq m$.

В результате решения игры $2 \times m$ находятся значения u_i^* согласно равенствам $u_i^* = \arg \max_{0 \leq u_i \leq 1} z_i(u_i)$, $z_i(u_i) = \min_{1 \leq j \leq m} Z_i(u_i, j)$, $i = 1, 2, \dots, n$.

Принятие решения о включении ($x_i = 1$) или исключении ($x_i = 0$) i -го ТЗ в

КИМ осуществляется путем решения двухкритериальной задачи оптимизации

$$\sum_{i=1}^n c_i x_i \rightarrow \max, \quad \sum_{i=1}^n d_i x_i \rightarrow \max, \quad x_i \in \{0, 1\},$$
$$\sum_{i=l_1}^{l_2} x_i \leq n_{0j}, \quad l_1 = \sum_{k=1}^{j-1} n_k, \quad l_2 = \sum_{k=1}^j n_k, \quad j = 1, 2, \dots, m,$$

где $c_i = z_i(u_i^*) / \{\sum_{k=1}^n z_k(u_k^*)\}$, $d_i = u_i^* / \{\sum_{k=1}^n u_k^*\}$.