

УДК 004.81:159.953.5(045)

Шибицька Н.М. канд. техн. наук
Тимофієва Г.А.

ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ВИЗНАЧЕННЯ СКЛАДНОСТІ ПИТАНЬ В ТЕСТОВИХ ЗАВДАННЯХ

Національний авіаційний університет

Запропоновано методика визначення складності питань в тестових завданнях. Досліджено ефективність її використання в автоматизованих системах діагностики знань. Використання такої технології дозволяє скоротити кількість питань у тестових завданнях та більш об'єктивно визначити рівень знань студента на різних етапах навчання. Доведено, що попереднє тестування значно скорочує кількість експериментальних даних на відміну від інших методик визначення складності питань тестового завдання

Вступ

Інтенсивне впровадження інформаційних технологій в освіті дозволяє перейти до диференційованого оцінювання знань студентів з різним рівнем базової підготовки [1]. Як наслідок, широке використання в процесі навчання досягнень комп'ютерних технологій зумовило виникнення одного із важливих елементів сучасного навчального процесу – автоматизованої системи навчання та контролю знань, яка, у той же час, є засобом керування навчальним процесом.

Для рішення задачі діагностики знань на сучасному етапі розвитку освіти використовуються тестові завдання (ТЗ), які дозволяють оцінити рівень знань та вмінь студента, що він отримав при вивченні учбового матеріалу. В залежності від того, наскільки успішно виконане завдання, робиться висновок о результатах навчання та рівні отриманих знань. При цьому зміст ТЗ визначається цілями навчання, що реалізується на цьому чи іншому етапі навчання. Для досягнення мети навчання на кожному з етапів навчального процесу необхідно визначити відповідні рівні засвоєння дидактичного матеріалу. При цьому для кожного рівня відповідає своя шкала оцінок.

При створенні тестів необхідна розробка математичного апарату для обробки результатів тестування та визначення їх якості [2]. Переважна більшість існуючих систем контролю знань проводить лише оцінку знань і не проводить жодних

оцінок якості самого тесту, не має можливості корегувати тестові завдання, не враховують таке поняття, як складність питань ТЗ. Відомо, що безпідставне збільшення складності ТЗ призводить до збільшення часу на проведення тестування, що в свою чергу, призводить до швидкого стомлення студентів, а це збільшує кількість помилок.

Таким чином, пошук нових методів удосконалення систем автоматизованого контролю знань має теоретичне і прикладне значення.

Аналіз досліджень і публікацій

Останнім часом приділяється багато уваги питанням визначення складності як самих ТЗ, так і питань, що є частиною ТЗ [1 – 3]. У класичній теорії тестів розглядаються лише емпіричні показники складності із спрощеним підрахунком частки неправильних відповідей [3]. Недолік такого методу в тому, що він характеризує рівні складності ТЗ лише як відношення неправильних відповідей до кількості учасників тесту, тому не може бути об'єктивною оцінкою. У системах дистанційного навчання часто використовують оцінку міру складності ТЗ на основі передбачуваного числа операцій, необхідних для вдалого виконання завдання. А саме, кожна форма завдання має свій вже визначений рівень складності, але така характеристика не дає змоги корегувати ТЗ.

На даний час найбільш поширеними та відомими є методи експертного оціню-

вання [4], які вимагають використання експертів із складу викладачів, спеціалістів даної області, які мають достатні знання за всіма питаннями теми, запропонованих для включення в ТЗ. Кількість експертів не регламентується, але при збільшенні їх кількості статистична передбачуваність визначення вагових коефіцієнтів збільшується. Таким чином, цей метод вимагає великих затрат. Однак, на практиці лише один викладач визначає вагові коефіцієнти, що не дає можливість адекватного оцінювання складності питання, бо є суб'єктивною думкою.

Постановка завдання

Розробити технологію визначення складності питань в ТЗ. Дослідити ефективність її використання в автоматизованих системах контролю знань.

Визначення складності питань в ТЗ

При розробці ТЗ необхідно врахувати їх складність, яка залежить від специфіки тесту. В залежності від місця проведення, цілей та задач діагностики знань розрізняються такі види контролю: вхідний, поточний контроль, тематичний контроль, підсумковий контроль.

Схема елементарного контролю має вигляд (рис. 1):

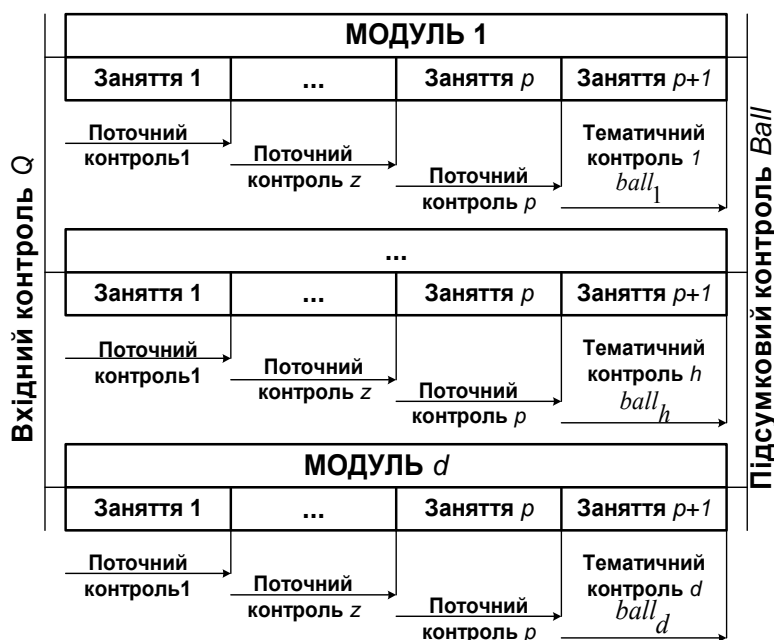


Рис. 1. Схема елементарного контролю знань

Де $Q = (q_1, q_2, \dots, q_i, \dots, q_n)$ – початковий рівень знань, $i = \overline{1, n}$ – кількість студентів, $z = \overline{1, p}$ – кількість занять за модулем. Таким чином, кінцевим результатом буде оцінка $Ball$, що складається із результатів тестування на кожному етапі навчання та визначається

$$Ball = \sum_{h=1}^d ball_h,$$

де $h = \overline{1, d}$ – кількість тестувань.

Розглянемо етапи створення тесту (рис. 2). Для структурування навчального матеріалу необхідно скласти логіко-

структурну схему чи реєстр елементів знань, з яких складається навчальний матеріал. При цьому для проведення експериментального тестування для корегування ТЗ необхідна велика вибірка студентів, що дає можливість при статистичній обробці результатів визначати якість ТЗ. Але така кількість експериментальних даних не завжди виправдовує себе, при цьому виникає питання скорочення як пілотних тестувань так і кількості самих питань. Для цього необхідно врахувати складність самих питань, бо збільшення складних чи простих питань знижує валідність тесту.

Для покращення достовірності результатів тестування при проектному розрахунку пропонується визначати вагові коефіцієнти за допомогою експериментального тестування (рис. 3).

Розглянемо методику розрахунку вагових коефіцієнтів, які характеризують ступінь складності ТЗ $k^{(n)}$ (де n – номер тесту) та визначають достовірність результатів контролю знань. За врахуванням вагових коефіцієнтів враховується ступінь складності питань ТЗ по відношенню до інших елементів знань, задається максимальна кількість балів, яку може отримати студент при правильному виконанні завдання. А це у свою чергу визначає достовірну кількісну та якісну характеристику знань. Налаштування ТЗ полягає в визначенні вагових коефіцієнтів в процесі експериментального тестування.

Перед початком пілотного тестування необхідно провести вхідний контроль для визначення базового рівня знань q_i , де $i = \overline{1, n}$, n – кількість студентів в групі. Така характеристика має бути об'єктивною оцінкою рівня підготовленості студентів. Оскільки останній час проводиться незалежне тестування, то можна враховувати ці результати як початковий рівень знань підтверджений на зовнішньому незалежному оцінюванні.

Нехай початкове значення вагових коефіцієнтів складності кожного питання $(w_{01}, w_{02}, \dots, w_{0s}) = 1$ під час першого експериментального тестування, але викладач може сам врахувати суб'єктивний рівень складності завдання та встановити максимальну кількість балів, що можна отримати за кожне питання

$$(ball_{01}, ball_{02}, \dots, ball_{0s}) = 1,$$

де s – кількість питань.

В результаті тестування отримаємо кількість балів за кожне питання для студентів $ball_{ij}$, де $i = \overline{1, n}$, $j = \overline{1, s}$. Знайдемо відносний рівень знань $q'_{ij} = \frac{q_i}{ball_{ij}}$ та проведемо нормалізацію необхідну для роз-

рахунку складності $w_{norm_{ij}}$ кожного питання для кожного студента групи:

$$w_{norm_{ij}} = \frac{q'_{ij}}{\max q'_i},$$

де $i = \overline{1, n}$, $j = \overline{1, s}$ (табл. 1). Таким чином, вагові коефіцієнти для кожного питання, що включені в ТЗ

$$w_j = \frac{\sum_{i=1}^n w_{norm_{ij}}}{n}. \quad (1)$$

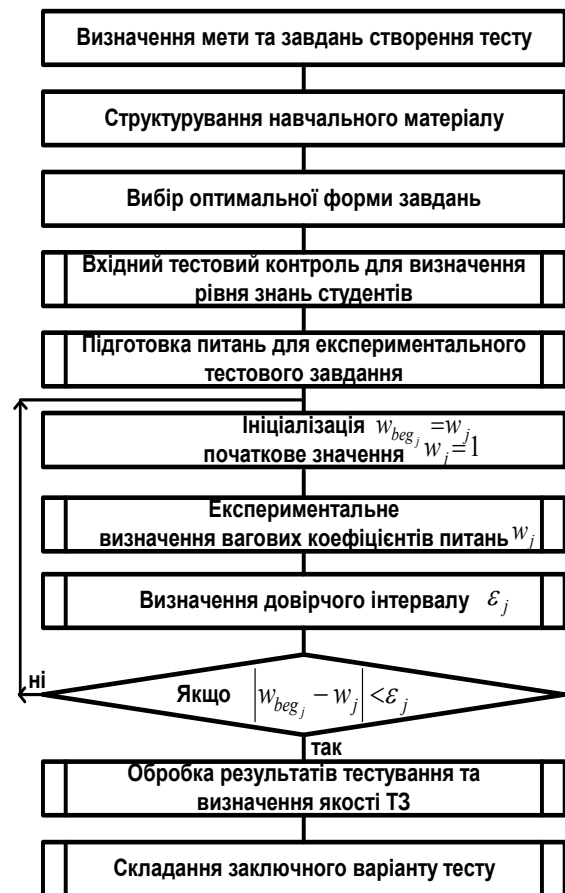


Рис. 2. Схема визначення та корегування вагових характеристик питань ТЗ

Виходячи з того, що ваговий коефіцієнт є чисельною мірою складності питання, можна розрахувати вагові коефіцієнти w_n^{rang} ТЗ в цілому:

$$w_n^{rang} = \frac{\sum_{j=1}^s w_j}{s},$$

де n – номер ТЗ.

Таблиця 1.

Код студента	№ питання (j)				
	1	...	j	...	s
1	$w_{norm_{11}}$		$w_{norm_{1j}}$		$w_{norm_{1s}}$
...
i	$w_{norm_{i1}}$		$w_{norm_{ij}}$		$w_{norm_{is}}$
...
n	$w_{norm_{n1}}$		$w_{norm_{nj}}$...	$w_{norm_{ns}}$
Результат	w_1	...	w_j	...	w_s

Перед використанням нових ТЗ необхідно проводити тривалу апробацію та корегування тестів на репрезентативних вибірках студентів для того, щоб отримати адекватність емпіричних даних тестування вимогам моделі вимірювання.

Розглянута методика визначення вагових коефіцієнтів використовує експериментальні дані, а отже є емпіричною системою, для корегування якої необхідно скорочена кількість дослідних тестувань. Визначимо довірчі інтервали для попередньо знайдених вагових коефіцієнтів.

$$\varepsilon_j = t_\varphi \frac{D_j}{\sqrt{n}}, \quad (2)$$

$$\begin{aligned} \varepsilon_j^B &= w_j + \varepsilon_j, \\ \varepsilon_j^H &= w_j - \varepsilon_j, \end{aligned} \quad (3)$$

$$D_j = \sqrt{\frac{\sum_{i=1}^n (w_{norm_{ij}} - w_j)^2}{n-1}},$$

де ε_j – довірчий інтервал, t_φ – довірча вірогідність, D_j – дисперсія експериментальних оцінок експертів, ε_j^B та ε_j^H – відповідно верхня та нижня довірчі границі вагових коефіцієнтів.

У випадку, коли розрахункові значення довірчої нижньої границі ε_j^H не перевищує 0 ($\varepsilon_j^H \leq 0$), то завдання виключається із переліку ТЗ або йому признача-

ється мінімальне значення вагового коефіцієнта складності w_j .

По результатам розрахунку вагових коефіцієнтів знайдемо різницю Δ_j :

$$\Delta_j = |w_{beg_j} - w_j|. \quad (4)$$

Якщо $\Delta_j < \varepsilon$, то тестування можна припинити і зафіксувати знайдене значення вагових коефіцієнтів. Повторне тестування з метою апробації тесту проводяться для питань в яких

$$\Delta_j > \varepsilon, \quad (5)$$

що дозволяє скоротити час, необхідний для відлагодження тесту. Ті питання, в яких різниця $\Delta_j < \varepsilon$, не включаються в повторне тестування.

Розглянемо застосування запропонованої методики на наступному прикладі. Для дослідження її адекватності було проведено тестування чотирьох груп студентів із різним рівнем підготовки. Нехай під час першого експериментального тестування початкове значення вагових коефіцієнтів складності $w_{0j} = 1$. Знайдемо вагові коефіцієнти w_j (1) кожного питання для першої групи студентів. За формулами (2) та (3) розрахуємо довірчий інтервал. Повторимо цю процедуру для наступних груп студентів, але вже $w_{beg_j} = w_j$. Після кожного тестування знаходимо різницю Δ_j (4) для всіх питань та перевіряємо умову (5).

Розглянемо збіжність методу на прикладі одного вибіркового питання (рис. 2).

При повторному тестуванні врахуємо попередній результат.

Таким чином достатньо трьох груп для визначення складності даного питання.

Порівняємо збіжність методу при використанні початкового рівня знань для нормалізації даних та без базового рівня знань.

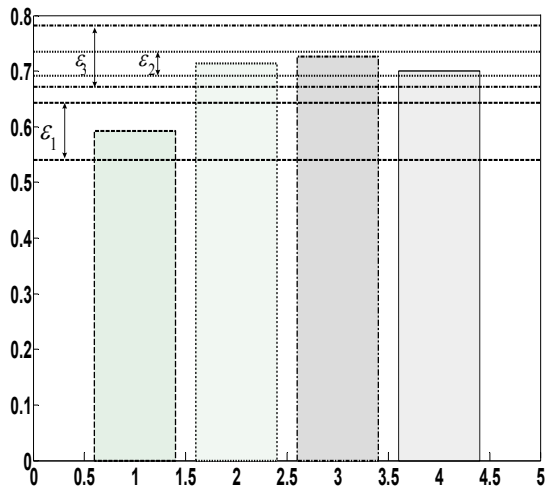


Рис. 4. Значення вагових коефіцієнтів складності отриманих в результаті експериментального тестування

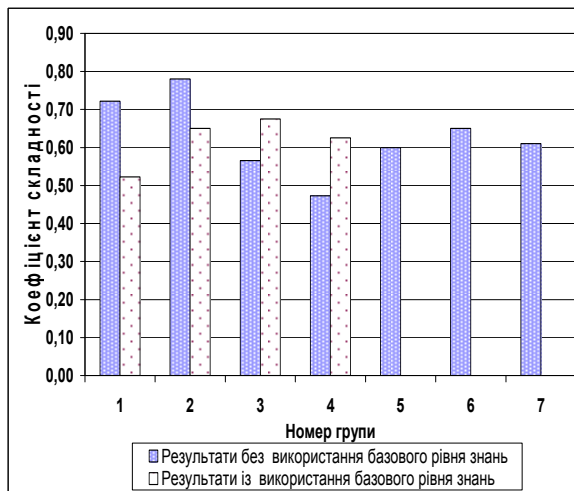


Рис. 5. Результати розрахунку коефіцієнтів складності в залежності від базового рівня знань

Результати вибірки, що приведені на рис. 5, дозволяють зробити висновок що необхідно було сім груп студентів, у той час коли при визначеному початковому рівня знань достатньо було лише трьох груп студентів для апробації ТЗ на прикладі одного питання.

Висновки

Розроблено інформаційну технологію для рішення задачі діагностики знань, відмінною характеристикою якої є визначення складності питань в ТЗ. Досліджено ефективність її використання в автоматизованих системах контролю знань.

Запропонована технологія дозволяє скоротити кількість питань, що в свою

чергу скорочує час необхідний для тестування. Розроблений алгоритм реалізації базової функції визначення та корегування вагових характеристик питань ТЗ дає можливість оперативно корегувати характеристики складності питань в автоматичному режимі.

При цьому методика не вимагає достатньо великої експериментальних даних на відміну від інших відомих параметричних моделей. Врахування базового рівня знань для визначення складності питання надає можливість більш об'єктивно визначити складність питань ТЗ, а це в свою чергу дозволяє отримати більш точний результат тестування студентів.

Використання запропонованої методики в автоматизованих системах діагностики знань надає можливість більш точно визначити результат тестування, що покращує ефективність навчального процесу.

Список літератури

1. Аванесов В.С. Основы научной организации педагогического контроля в высшей школе. – М.: МИСиС, 1989. – 168 с.
2. Олійник М.М. Тест як інструмент кількісної діагностики рівня знань в сучасних технологіях навчання – Донецьк: ДонНУ, 2001. – 84 с.
3. Федорчук П.І. Адаптивні тести: статистичні методи аналізу результатів тестового контролю знань // Математичні машини і системи. – К.: 2007. – №3,4. – С. 122–138.
4. Волков В.І., Алексєєв О.М., Алексєєв М.О. Тестовий контроль знань: Навчальний посібник. – Суми: ВТД «Університетська книга», 2004. – 100 с.