

Оцінка повноти відповідей в автоматизованих системах контролю знань

Запропоновано новий метод аналізу повноти відповідей на тестові запитання. Цей метод базується на використанні Декартової відстані, відстані Левенштейна. Метод призначено для використання у автоматизованих системах контролю знань.

Ключові слова: відстань Левенштейна, відстань Хеммінга, декартова відстань, аналіз результатів тестування, повнота відповіді.

Вступ

Постановка проблеми. Перевірка знань за допомогою тестування є досить поширеним засобом контролю знань, що використовується майже в усіх навчальних закладах. Спочатку тестування проводилось за допомогою паперових носіїв, що роздавались учням. З одного боку, такий підхід не потребує використання технічних засобів під час перевірки знань, але може призвести до помилок при перевірці тестових завдань та примушує викладача додатково витратити час на перевірку завдань. Залучення до процесу тестування комп'ютерів значно полегшило та дозволило автоматизувати, як сам процес тестування, так і процес оцінювання отриманих відповідей. Однак, однією з найбільш важливих проблем у тестуванні залишається отримання об'єктивних результатів, що максимально достовірно відображають реальну картину знань [1-6].

Аналіз останніх досліджень. Проблема аналізу результатів тестування досліджується вже не один рік [1-6]. Так у роботі [6] досліджується процес формування відповіді на тестове запитання у часі, а у роботах [2-4] пропонуються різні моделі оцінки відповідей та складності тестових запитань.

Виділення невирішених раніше частин загальної проблеми. Очевидно, що існуючі методики оцінки результатів тестування та складності тестових запитань мають як переваги, так і недоліки. Тому, є актуальною розробка нових методів оцінювання повноти наданих відповідей, складності тестових запитань та комп'ютерних систем тестування, в яких використовуються такі методи.

Формулювання цілей статті. Поширимо запропонований у роботі [3] метод оцінки складності запитання закритого типу на запитання з обрання правильної послідовності та обрання правильної відповідності, та відкриті запитання.

Основні визначення та загальні положення

Зупинимось спочатку на розгляді запитання з обрання правильної послідовності. У якості прикладу візьмемо наступне запитання:

В якій послідовності адміністратор даних має виконувати вказані нижче дії

1. передати інформацію розробникам для первинного ознайомлення
2. визначитись з розмірами інформатизації підприємства
3. узагальнити отриману інформацію
4. опитати структурні підрозділи з метою отримання відповіді на необхідні запитання.

Правильна відповідь на вказане запитання має виглядати наступним чином

- 1) передати інформацію розробникам для первинного ознайомлення – **4-й етап**
- 2) визначитись з розмірами інформатизації підприємства – **1-й етап**
- 3) узагальнити отриману інформацію – **3-й етап**
- 4) опитати структурні підрозділи з метою отримання відповіді на необхідні запитання – **2-й етап**

Фактично у підсвідомості студента, що відповідає, це запитання умовно можна розділити на чотири запитання закритого типу, а саме

I) Адміністратор даних має передати інформацію розробникам для первинного ознайомлення

- 1) на першому етапі
- 2) на другому етапі
- 3) на третьому етапі
- 4) на четвертому етапі

II) Адміністратор даних має визначитись з розмірами інформатизації підприємства

- 1) на першому етапі
- 2) на другому етапі
- 3) на третьому етапі
- 4) на четвертому етапі

- III) Адміністратор даних має узагальнити отриману інформацію
- на першому етапі
 - на другому етапі
 - на третьому етапі
 - на четвертому етапі
- IV) Адміністратор даних має опитати структурні підрозділи з метою отримання відповіді на необхідні запитання
- на першому етапі
 - на другому етапі
 - на третьому етапі
 - на четвертому етапі

Будемо вважати, що правильному варіанту відповіді відповідає 1, а неправильному 0 [3, 6]. Тоді шаблони правильних відповідей на запитання I-IV матимуть вигляд наведений у табл. 1.

Таблиця 1. Шаблони правильних відповідей

Запитання	Шаблон відповіді
I	(0, 0, 0, 1)
II	(1, 0, 0, 0)
III	(0, 0, 1, 0)
IV	(0, 1, 0, 0)

Об'єднуючи послідовно шаблони правильних відповідей з табл. 1 ми отримаємо точку, яка відповідає шаблону правильної відповіді [3] на наше початкове запитання – $A(0,0,0,1,1,0,0,0,0,0,1,0,0,1,0,0)$.

Надалі, не обмежуючи загальності висновків, будемо вважати, що будь-яке запитання на обрання правильної послідовності містить N подій. У цьому випадку шаблону правильної відповіді на запитання буде відповідати точка $A_i(a_{i1}, a_{i2}, a_{i3}, \dots, a_{iM})$ [3], де i – номер запитання, $M = N^2$, а структуру шаблону правильної відповіді зображено на рис. 1.

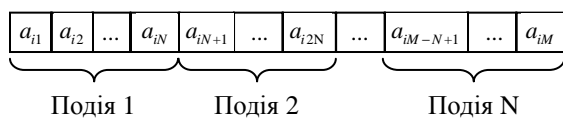


Рисунок 1 – Структура шаблону правильної відповіді.

Відповідь на це запитання можна представити у вигляді точок $B_{ij}(b_{ij1}, b_{ij2}, b_{ij3}, \dots, b_{ijM})$, де $a_{ij} \in \{0;1\}$, $b_{ijk} \in \{0;1\}$, $j = \overline{1, N_i}$, $k = \overline{1, M}$, N_i - кількість відповідей на i -те запитання.

Аналогічним чином, можна побудувати шаблон правильної відповіді для запитань на обрання відповідності та представити його у вигляді точки $A_i(a_{i1}, a_{i2}, a_{i3}, \dots, a_{iM})$, а відповіді на

це запитання у вигляді точок $B_{ij}(b_{ij1}, b_{ij2}, b_{ij3}, \dots, b_{ijM})$.

Тоді, для запитань на обрання правильної послідовності та обрання правильної відповідності, так само, як і для запитань закритого типу [2, 3], можна визначити близькість наданої відповіді до правильної обчисливши відстань між точками B_{ij} та A_i за наступною формулою

$$\rho_{Dek}(A_i, B_{ij}) = \sqrt{\sum_{k=1}^M (a_{ik} - b_{ijk})^2}. \quad (1)$$

Слід відзначити, що у даному випадку квадрат Декартової відстані між точкою A_i та точкою B_{ij} , чисельно дорівнює відстані Хеммінга між шаблоном правильної відповіді та наданою відповіддю [7].

Розглянемо тепер відкриті запитання та відповіді на них. Ці запитання мають найменшу вірогідність вгадування правильної відповіді, але водночас більш складні з точки зору аналізу результатів і програмної реалізації систем тестування.

Відкриті запитання можуть мати більше ніж один шаблон правильної відповіді, на відміну від запитань іншого типу (закриті, з обрання послідовності та обрання відповідності). У якості прикладу розглянемо наступне запитання:

В топології «зірка» усі вузли під'єднуються до інтелектуального пристрою, що має назву ...

Правильні відповіді на вказане вище відкрите запитання наведені у табл. 2. Так еталонний рядок T_1 являє собою назву пристрою англійською мовою, еталонні рядки T_2 та T_3 є транслітерацією рядку T_1 українською та російською мовами відповідно. Еталонні рядки T_4 та T_5 – альтернативні назви пристрою українською та російською мовами.

Таблиця 2. Приклад еталонів правильної відповіді на відкрите запитання

Еталон	Правильна відповідь
T_1	switch
T_2	світч
T_3	свитч
T_4	комутатор
T_5	коммутатор

Таким чином, для вказаного запитання ми маємо п'ять еталонних рядків, кожен з яких є вірною відповіддю, тому необхідно перевіряти надану відповідь на збіжність (ідентичність) з кожним із еталонних рядків.

Надалі, не обмежуючи загальність висновків, будемо вважати, що для будь-якого

відкритого запитання під номером i , існує множина шаблонів правильних відповідей $\Omega_i = \{T_{ir}, |T_{ir}| > 0, r = 1, Q\}$, де $|T_{ir}|$ – довжина рядку, Q – кількість шаблонів відповідей. Тоді відстань між правильною відповіддю та наданою, для відкритих запитань можна обчислити за формулою

$$\rho(\Omega_i, S_{ij}) = \min_{T_{ir} \in \Omega_i} (\rho_{Lev}(T_{ir}, S_{ij})), \quad (2)$$

де ρ_{Lev} – відстань Левенштейна між двома рядками [7], що показує мінімальну кількість операцій вставки одного символу (Insert), видалення одного символу (Delete), заміни одного символу на інший (Change), необхідних для перетворення першого рядку в другий.

Надалі, для будь-якого запитання під номером i та відповіддю на нього під номером j будемо обчислювати відстань між шаблоном відповіді та відповіддю за формулою

$$d(i, j) = \begin{cases} \rho_{Dek}^2(A_i, B_{ij}), & qtype(i) \neq 3 \\ \rho(\Omega_i, S_{ij}), & qtype(i) = 3 \end{cases}, \quad (3)$$

яку з урахуванням формул (1) та (2) можна представити у вигляді

$$d(i, j) = \begin{cases} \sum_{k=1}^M (a_{ik} - b_{ijk})^2, & qtype(i) \neq 3 \\ \min_{T_{ir} \in \Omega_i} (\rho_{Lev}(T_{ir}, S_{ij})), & qtype(i) = 3 \end{cases}, \quad (4)$$

де $qtype(i) = 0$ якщо i -ме запитання закрите, $qtype(i) = 1$ для запитання на обрання правильної послідовності, $qtype(i) = 2$ для запитання на відповідність, $qtype(i) = 3$ для відкритого запитання.

З іншого боку, формули (3) та (4) показують кількість помилок, припущених у j -й відповіді на i -ме запитання [3], тобто, чим більше помилок, тим більше відстань між шаблоном відповіді та відповіддю.

На практиці тестові запитання мають досить різну структуру. Так, наприклад, у запитаннях закритого типу може змінюватись загальна кількість варіантів відповідей та кількість варіантів, що є вірними. У відкритих запитаннях може змінюватись довжина еталонного рядку. Умовно кажучи, дві помилки у рядку довжиною в три символи – це набагато гірше ніж дві помилки у рядку довжиною 15 символів. Тому, для визначення повноти наданої відповіді у подальшому необхідно використовувати відносну кількість помилок $\varepsilon(i, j)$, яку будемо обчислювати за наступною формулою:

$$\varepsilon(i, j) = \begin{cases} \frac{\sum_{k=1}^M (a_{ik} - b_{ijk})^2}{\sum_{k=1}^M a_{ik}^2}, & qtype(i) \neq 3 \\ \min_{T_{ir} \in \Omega_i} \left(\frac{\rho_{Lev}(T_{ir}, S_{ij})}{|T_{ir}|} \right), & qtype(i) = 3 \end{cases}, \quad (5)$$

де $\varepsilon(i, j) = 0$, якщо відповідь на запитання є повною (цілком правильною). Чим більше $\varepsilon(i, j)$, тим менш правильною (повною) є надана відповідь.

Зазвичай, на практиці, для визначення складності запитання, обчислюється відсоток правильних відповідей від загальної кількості наданих відповідей на це запитання. Для визначення складності тестових запитань будемо використовувати відсоток цілком правильних відповідей, та середнє значення відносної кількості помилок

$$\varepsilon_{Ci} = \frac{\sum \varepsilon(i, j)}{N_i}. \quad (6)$$

Приклад використання методу оцінки повноти відповіді

Запропонований у цій статті метод оцінки повноти відповідей було реалізовано програмно в автоматизованій системі контролю знань (АСКЗ) та використано на практиці більше року.

У табл. 3 наведено деякі варіанти відповідей студентів, на вказане раніше відкрите запитання, а також чисельне значення відстані, відносної кількості помилок та номер шаблону до якого ця відповідь є найближчою.

Таблиця 3. Приклади відповідей на відкрите запитання

№	Відповідь	$d(i, j)$	$\varepsilon(i, j)$	T_i
1	switch	0	0,000	1
2	комутатор	0	0,000	4
3	свитч	0	0,000	3
4	svitch	1	0,167	1
5	switc	1	0,167	1
6	swich	1	0,167	1
7	свич	1	0,200	3
8	коммунікатор	3	0,300	5
9	концентратор	5	0,500	5
10	т-конектор	6	0,667	4
11	процесор	7	0,778	4
12	сервер	5	0,889	–
13	кільце	6	0,889	–
14	топология	8	0,889	–
15	интернет	8	0,900	–
16	loopback	7	1,000	–
17	витая пара	8	1,000	–
18	локальная сеть	12	1,200	–

Так відповіді під номерами 1–3 є повними і цілком співпадають з еталонними відповідями наведеними у табл. 2. Відповіді під номерами 4–7 є майже повними і різниця з еталонною відповіддю полягає лише в одному символі ($d(i, j)=1$). Ці помилки є граматичними, і при усній відповіді викладач, міг би їх взагалі не помітити. З даних, наведених у табл. 2, можна побачити, що, чим більше значення $\varepsilon(i, j)$, тим менше надана відповідь збігається з еталонною. Так для відповідей №12–№18 взагалі важко зрозуміти з якою еталонною відповіддю вони збігаються.

У табл. 4 наведено 10 запитань закритого та відкритого типу у порядку зростання відсотку цілком правильних відповідей. Так, запитання №1 та №2 мають самий низький рівень цілком правильних відповідей. Якщо в даному випадку для визначення складності запитання враховувати лише відсоток цілком правильних відповідей, можна зробити висновок, що запитання №1 є найскладнішим з десяти наведених. Однак запитання 1 має більш ніж вдвічі менше значення ε_{Ci} порівняно з запитанням №2 і за цим показником має бути розташованим між запитаннями №8 та №9. Цей факт можна дуже легко пояснити. При створенні запитання №1 викладач у якості одного з еталонних рядків обрав наступний – «избыточность данных». Як показав аналіз результатів, під час тестування, студенти дуже часто вводили наступні відповіді:

- «избыточность»;
- «избыточностью»;
- «избыточность данных».

Кожна з цих відповідей не збігається на 100% з еталонним рядком. Однак неможна вважати що вказані відповіді свідчать про незнання студентами матеріалу. Як показала співбесіда, ці студенти або не звернули увагу на те у якому відмінку треба вводити відповідь, або пропустили слово «даних», вважаючи що це є очевидним.

Висновки

Таким чином запропонований у статті метод дозволяє ефективно оцінювати повноту відповідей на тестові запитання. При необхідності цей метод може бути вдосконаленим за рахунок використання у формулах (4) та (5) однієї з модифікованих відстаней Левенштейна.

Наукова новизна роботи полягає у новому методі аналізу повноти відповідей, який базується на геометричному підході та використанні

Подібна ситуація склалась і при відповіді на запитання №3, яке мало 5 варіантів відповідей, 3 з яких були вірними. Як показав детальний аналіз, студенти обирали вірні відповіді але не всі. Цей факт не дозволяє вважати це запитання дуже складним.

Таблиця 4. Складність запитань різного типу

№	Запитання	ε_{Ci}	%
1	Наявність в базі даних інформації, що повторюється називається	0,481	6,4
2	Оберіть які з вказаних запитань повинні вирішуватись адміністратором бази даних	1,092	13,0
3	Які основні елементи мають ОС с графічним інтерфейсом, що відрізняють їх від ОС без графічного інтерфейсу:	0,602	16,7
4	Для того щоб два комп'ютери мали змогу здійснювати обмін інформацією один з одним у локальній мережі, маски підмережі (X.X.X.X) на кожному з цих комп'ютерів	1,518	20,9
5	IP-адреса комп'ютера має відповідати наступним вимогам	0,562	21,3
6	В топології зірка усі вузли під'єднуються до інтелектуального пристрою, що має назву	0,609	28,2
7	Яка з IP-адрес записана неправильно	1,342	32,9
8	У реляційних БД рядок таблиці має назву	0,512	46,3
9	Розширення txt мають	0,341	79,0
10	Розмір файлу вимірюється у:	0,119	90,2

відстані Левенштейна.

Практична значимість запропонованого методу полягає у нових та більш ефективних можливостях аналізу повноти відповідей та складності тестових запитань у АСКЗ.

Подальший розвиток досліджень пов'язан із використанням цієї методики у експертних системах [1].

Список використаної літератури

1. Антонов Ю.С. Використання експертних систем у комп'ютерних системах тестування / Ю.С. Антонов // Сучасні тенденції розвитку інформаційних технологій в науці, освіті та економіці: матеріали V Всеукраїнської науково-практичної конференції: в 2-х т. Луганськ 7-9 квітня 2011р. – Луганськ: Phoenix, 2011. – Т. 1. – С 22-23.

2. Антонов Ю.С. Деякі аспекти створення комп'ютерних систем тестування та використання нової методики аналізу результатів / Ю. С. Антонов // Сучасні тенденції розвитку інформаційних технологій в науці, освіті та економіці: матеріали IV Все-української науково-практичної конференції. 15-17 квітня 2010 р., м.Луганськ. – Луганськ: Phoenix, 2010. – С 153-155.
3. Антонов Ю.С. Методика аналізу тестових завдань на основі отриманих результатів тестування: [Елек-тронний ресурс] / Ю.С. Антонов, О.М. Космінська // Інформаційні технології і засоби навчання. – 2009. – № 4. – Режим доступу: <http://www.nbu.gov.ua/e-journals/ITZN/em12/emg.html>. – Заголовок з екрана.
4. Белоус Н.В. Дифференциальное оценивание знаний при дистанционном тестировании / Н.В. Белоус, И.В. Куцевич // Штучний інтелект. – 2009. – № 4. – С. 63-73.
5. Алексеев Н.А. К вопросу о количественном оценивании результатов тестового контроля знаний / Н.А. Алексеев // Открытое образование. – 2006. – №4. – С. 45-51.
6. Рыбанов А.А. Моделирование динамики процесса формирования ответов на тестовые задания закрытой формы при дистанционном тестировании знаний / А.А. Рыбанов // Открытое образование. – 2006. – №6. – С. 43-50.
7. Гасфилд Дэн. Строки, деревья и последовательности в алгоритмах: информатика и вычислительная биология / Гасфилд Дэн. – СПб.: Невский Диалект; БХВ-Петербург, 2003. – 654 с.

Надійшла до редакції 30.01.2012

Ю.С. АНТОНОВ

Донецкий государственный университет
управления

Y.S. ANTONOV

Donetsk State University of Management

ОЦЕНКА ПОЛНОТЫ ОТВЕТОВ В АВТОМАТИЗИРОВАННЫХ СИСТЕМАХ КОНТРОЛЯ ЗНАНИЙ

ANSWERS COMPLETENESS ESTIMATION IN KNOWLEDGE CONTROL AUTOMATED SYSTEM

Предложен новый метод анализа полноты ответов на тестовые вопросы. Этот метод основывается на использовании Декартового расстояния, расстояния Левенштейна. Метод предназначен для использования в автоматизированных системах контроля знаний.

New method of answers completeness estimation analysis for test questions are proposed. This method is based on the use of Cartesian's distance, distance of Levenshtein A method is intended for the use in knowledge control automated system.

Ключевые слова: *расстояние Левенштейна, расстояние Хэмминга, Декартово расстояние, анализ результатов тестирования, полнота ответа.*

Keywords: *distance of Levenstein, Hemming distance, Dekart distance, test results analysis, answers completeness.*