

УДК 004.93

DOI: 10.31498/2225-6733.50.2025.336246

МОДЕЛЮВАННЯ ТА РЕАЛІЗАЦІЯ БАГАТОРІВНЕВОГО АДАПТИВНОГО ТЕСТУВАННЯ У ІНТЕЛЕКТУАЛЬНІЙ СИСТЕМІ

Федосова І.В.	д-р пед. наук, професор, ДВНЗ «Приазовський державний технічний університет», м. Дніпро, ORCID: https://orcid.org/0000-0003-3923-8270 , e-mail: fedosova_i_v@pstu.edu ;
Проніна О.І.	канд. техн. наук, доцент, ДВНЗ «Приазовський державний технічний університет», м. Дніпро, ORCID: https://orcid.org/0000-0001-7085-8027 , e-mail: pronina_o_i@pstu.edu ;
Веремій В.О.	магістр, ДВНЗ «Приазовський державний технічний університет», м. Дніпро

Стаття присвячена розгляданню підходу до підвищення якості освітнього процесу шляхом впровадження комп'ютерного адаптивного тестування. В роботі акцентується увага на проблемах традиційного оцінювання знань, таких як відсутність індивідуалізації та суб'єктивність викладача, й пропонують адаптивні методи формування тестів, які враховують рівень підготовки здобувача освіти в режимі реального часу. Описано різні алгоритми адаптивного тестування. Також описано метод багаторівневого тестування з блоковою структурою, де завдання змінюються відповідно до попередніх відповідей. Цей метод був покладений в основу інформаційної системи для перевірки якості знань здобувачів. Було наведено формулу для розрахунку переходу між блоками, що окреслюють рівень знань в тесті. Також було описано інтерфейс користувача, типи питань, подано реалізацію системи тестування, а також механізм зміни рівнів складності в процесі проходження тесту. Інтерфейс системи є зручним для користувача і передбачає підтримку різних типів тестових завдань, що робить її універсальною для використання в закладах освіти. Розроблена багаторівнева структура тесту дозволяє гнучко адаптувати не лише складність, а й логіку переходу між завданнями, підвищуючи точність і об'єктивність оцінювання, а також зменшуючи загальну тривалість тестування. Реалізована методика демонструє практичну доцільність і гнучкість впровадження адаптивного підходу в систему автоматизованого контролю знань, що може стати основою для створення більш ефективних освітніх технологій.

Ключові слова: адаптивне тестування, інтелектуальна система, індивідуальна траєкторія тесту, оцінка знань.

Постановка проблеми

У сучасній системі освіти пріоритет набуває не лише доступність, а й якість навчання. Питання якості набуває особливої ваги як у педагогічній науці, так і в суспільному дискурсі [1]. Одним із способів визначення ефективності освітнього процесу є аналіз того, наскільки добре здобувачі освіти опановують навчальний матеріал. Рівень знань здобувачів оцінюється шляхом спеціально організованих перевірок, які мають педагогічну основу. Зазвичай ці перевірки передбачають аналіз відповідей на завдання різного рівня складності. Щоб мінімізувати вплив особистих оцінок викладача, все частіше використовуються автоматизовані засоби перевірки знань, зокрема тестові системи, які дозволяють забезпечити більш об'єктивне оцінювання.

Адаптивне тестування з'явилося як відповідь на потребу зробити процес оцінювання знань більш ефективним. Основна ідея полягала в тому, щоб скоротити кількість завдань і час, витрачений на тестування, водночас підвищивши точність визначення рівня підготовки здобувачів освіти. Ключовим принципом цієї технології є індивідуальний підхід до формування тесту – завдання підбираються відповідно до знань конкретного учасника, що дозволяє формувати більш результативну систему перевірки.

Головною рисою комп'ютерного адаптивного тестування є побудова індивідуальної послідовності завдань безпосередньо під час проходження тесту. Ця

послідовність залежить від рівня підготовки кожного учасника, який оцінюється в режимі реального часу – іноді навіть із можливістю прогнозування подальших відповідей. У результаті можна говорити про те, що:

- кожен конкретний тест унікальний і не був раніше опублікований, що також підвищує його секретність (і надійність);

- ефективність комп'ютерного адаптивного тестування вище звичайного комп'ютерного тестування, так як труднощі запропонованих завдань прагнуть відповідати рівню досягнень тестованого, отже, підвищується інформативність тесту і швидше досягається задана точність оцінювання результату, що також може скоротити тривалість тесту;

- комп'ютерне адаптивне тестування дозволяє розпізнавати не тільки тестованих з середнім рівнем досягнень, але і дає можливість виявити найбільш яскраві, видатні особистості.

Також однією з основних проблем є той факт, що заздалегідь невідомо, скільки запитань треба поставити тестованому, щоб визначити рівень його знань. Це пов'язано з тим, що в основі адаптивного підходу лежить індивідуалізація процедури відбору завдань тесту, яка за рахунок оптимізації складності завдань стосовно до рівня підготовленості здобувачів освіти забезпечує генерацію ефективних тестів [2].

Аналіз останніх досліджень та публікацій

Існує багато публікацій на дану тематику, оскільки це питання є актуальним на поточний момент. Так, наприклад, автори роботи [3] пропонують метод формування оцінювання, котрий є альтернативою традиційним підходам до оцінювання знань та системам IPT (Item Response Theory). Алгоритм може зменшити кількість завдань, які має вирішити здобувач вищої освіти, і забезпечити більш швидкий спосіб визначення відповідного персоналізованого шляху навчання для здобувачів вищої освіти. Алгоритм визначає найслабші сфери знань здобувачів та вказує найпростіші пункти, з яких можна розпочати корекційне навчання.

В роботі [4] досліджувався вплив різних методів вибору завдань на інформаційну функцію тесту (TIF) та ефективність тесту в комп'ютерному адаптивному тестуванні (CAT). Дослідження проводилося з використанням змодельованих даних, а константами дослідження є розмір вибірки, розподіл параметрів здібностей, розмір пулу завдань, модель теорії відповідей на завдання (IRT) та розподіл параметрів завдань, метод оцінки здібностей, правило початку, контроль експозиції завдань та правило зупинки. Методи вибору завдань, які є незалежними змінними цього дослідження, – це критерій інтервальної інформації, збалансована за ефективністю інформація, значення -b відповідності, інформація Кульбака-Лейблера, максимальна інформація Фішера, інформація, зважена за правдоподібністю, та випадковий вибір. При порівнянні цих методів найкращі показники в аспекті TIF забезпечує метод максимальної інформації Фішера.

Робота [5] має на меті представити завдання, які статистично оптимізують процес оцінювання, враховуючи відповіді випробуваного та оцінені рівні. Нещодавні розробки в навчанні з підкріпленням та глибоких нейронних мережах надають адаптованому тестуванню потенціал для вибору завдань, які використовують більше інформації з усіх завдань, що залишилися, а не лише зосереджуються на наступних кількох завданнях, які потрібно вибрати. Автори переробили CAT відповідно до структури навчання з підкріпленням та пропонують нову стратегію вибору завдань, засновану на методі глибокої Q-мережі (DQN).

У статті [6] наводиться багатовимірне комп'ютеризоване адаптивне тестування для завдань з примусовим вибором (MFC-CAT), що поєднує переваги багатовимірних завдань з примусовим вибором (MFC) та комп'ютеризованого адаптивного тестування (CAT). Це дослідження пропонує набір нових методів вибору завдань на основі інформації KL для MFC-CAT (а саме MFC-KI, MFC-KB та MFC-KLP) на основі моделі Терстона IRT (TIRT). Було проведено три симуляційні дослідження, включаючи одне на основі реальних даних, для порівняння ефективності запропонованих методів вибору завдань на основі KL з існуючими методами на основі FI у тривимірних та п'ятивимірних сценаріях

MFC-CAT з різною тривалістю тестування та міжзнаковими кореляціями.

В роботі [7] розглядається питання незацікавлених відповідей в комп'ютеризованих адаптивних тестах (КАТ). Незацікавлені відповіді (наприклад, швидкі здогадки), які не відображають справжній рівень здібностей учнів, можуть призвести до вибору менш інформативних завдань і тим самим забруднити процедури вибору завдань та оцінки здібностей. Це дослідження пропонує альтернативний метод вибору завдань, заснований на максимальному показнику інформації Фішера (MFI), який розглядає залученість до тестування як вторинну латентну ознаку для вибору найоптимальніших завдань на основі як здібностей, так і залученості. Результати досліджень постфактумного моделювання показали, що запропонований метод може оптимізувати вибір завдань та підвищити точність остаточних оцінок здібностей, особливо для учнів з низьким рівнем здібностей.

Важливою особливістю проектування як комп'ютеризованого адаптивного тестування, так і багатоетапного адаптивного тестування є використання відповідного методу вибору елементів. Очікується, що метод вибору елементів вибиратиме найоптимальніші елементи залежно від рівня здібностей випробуваних, враховуючи інші особливості проектування (наприклад, експозицію елементів та використання банку елементів). У роботі [8] було представлено колаборативну фільтрацію (CF) як новий метод вибору елементів у багатоетапній адаптивній тестовій системі, що збирається на льоту. Результати моделювання показали, що метод UBCF перевершує традиційні методи вибору елементів щодо точності вимірювання. Також метод IBCF продемонстрував найвищу продуктивність з точки зору використання банку елементів. Обговорюються обмеження поточного дослідження та напрямки майбутніх досліджень.

У дослідженні [9] проаналізовано можливість запровадження адаптивного комп'ютерного тестування з урахуванням рівня складності та часу виконання завдань. Запропоновано алгоритм, який змінює складність наступного завдання залежно від інтегральної оцінки попереднього, з порогом ефективності q . Імітаційне моделювання показало, що методика є гнучкою та ефективною, і може бути налаштована під індивідуальні особливості студентів і специфіку дисциплін.

Отже, використання методу адаптивного тестування як основи системи перевірки якості знань здобувачів є актуальним практичним завданням.

Мета статті

Побудова інтелектуальної системи оцінки якості знань здобувачів на основі адаптивного тестування.

Виклад основного матеріалу

В адаптивному тестуванні існує кілька підходів до формування завдань, які відрізняються за

складністю, гнучкістю та алгоритмічною побудовою. Одним із популярних підходів є теорія реагування на завдання, де завдання формуються на основі моделей, які враховують складність завдань, рівень підготовки тестованого та ймовірність правильної відповіді. Після кожної відповіді система оцінює рівень знань і підбирає наступне завдання з урахуванням оновленої оцінки.

Також слід зазначити класичний адаптивний алгоритм. В ньому завдання підбираються поступово – якщо відповідь правильна, то дається складніше завдання; якщо неправильна, то легше завдання. Існує також багаторівневе тестування: в ньому тест поділяється на блоки (модулі), кожен з яких має фіксовану складність. Після проходження першого блоку система обирає наступний блок відповідно до результату. Є також більш простий варіант поточного методу заснований на евристичці. В ньому перехід на більш складні рівні відбувається за правилами, наприклад, «при трьох підряд правильних відповідей перейти на більш складний рівень».

Окремо можна виділити ще аналіз статистичних результатів після проходження кожного питання тесту. Після кожної відповіді обчислюється апостеріорна ймовірність рівня знань і на її основі обирається наступне завдання.

В роботі було вирішено використовувати багаторівневе тестування, що складається з блочного методу.

В середині блоку на основі відповіді на попереднє завдання обирається наступне – з відповідним рівнем складності. У результаті кожен наступний крок формує індивідуальну структуру тесту, де адаптується не лише складність, але й розмір зміни між завданнями. Для переходу між блоками необхідно кожного разу розраховувати вагу питань для кожного типу блоку. Для розрахунку ваги питання кожного блоку використовується формула:

$$w = \frac{qt + 1}{qt} > 1,$$

де w – вага питання; qt – загальна кількість питань.

Вага питання повинна бути більше одиниці, оскільки лише в цьому випадку буде відбуватися перехід на наступний тип блоку. В розробленій системі використовується стандартизована вага питань, оптимальне значення ваги питань блоку було виявлено експериментальним шляхом та складає 1,14. Такий підхід дозволяє скоротити загальну кількість питань, зберігаючи при цьому змістовне наповнення перевірки для кожного блоку. Цей процес реалізується у вигляді циклічного алгоритму, де кожен етап спрямований на уточнення рівня знань на основі результатів поточних відповідей, представлений на рисунку 1.



Рис. 1 – Алгоритм змінно-розгалуженого багатокрокового тестування

Розроблений метод був реалізований у вигляді інтелектуальної системи тестування якості знань здобувачів, інтерфейс системи наведено на рисунку 2.

Основна функція «Тестування» доступна за допомогою команди меню «Тестування» пункту «Доступні тести», що наведено на рисунку 2. Після цього з'явиться вікно «Доступні тести», в якому можна обрати різні варіанти тестів для оцінки якості засвоєння знань студентів та обрати кнопку «Розпочати обраний тест». Після проходження тесту можна переглянути протокол тестування. Ця функція доступна через пункт меню «Тестування» команду «Протоколи тестів». Кожне питання з'являється у окремому вікні. Питання можуть бути чотирьох видів: тестові питання закритого типу з однією правильною відповіддю; тестові питання закритого типу з багатьма правильними відповідями; тестові завдання відкритого типу з можливістю

дописати слово; тестові завдання відкритого типу з можливістю дописати число.

Також можна переглянути, як працював адаптивний метод при формуванні питань в залежності від відповіді. Для цього є кнопка «Зміна рівнів». Рівень тестування змінюється залежно від кількості правильних відповідей у визначеній порції запитань. У випадку рисунка 3, перша відповідь була неправильною, тому навіть після трьох запитань підвищення не відбулося. Проте наступні три правильні відповіді (питання 4-6) забезпечили перехід на вищий рівень. Подібна ситуація повторюється з питаннями 7-9: одна помилка призупинила перехід, але три наступні правильні відповіді (10–12) дозволили перейти на рівень «добрий». Оскільки далі всі відповіді (13-15) були неправильні, система автоматично знизила рівень до «середній». Подальші зміни рівнів відбуваються за тим же принципом.

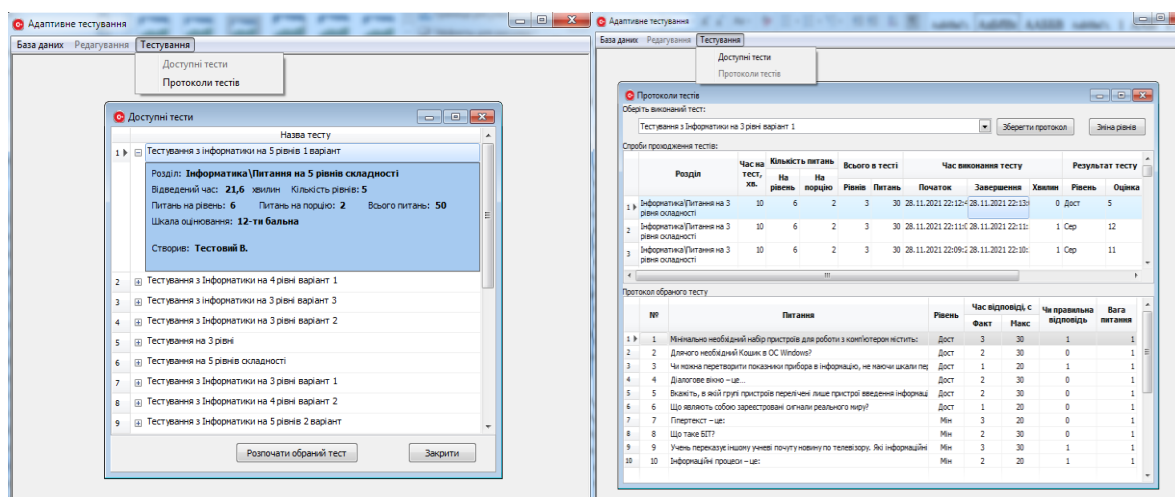


Рис. 2 – Інтерфейс розробленої системи

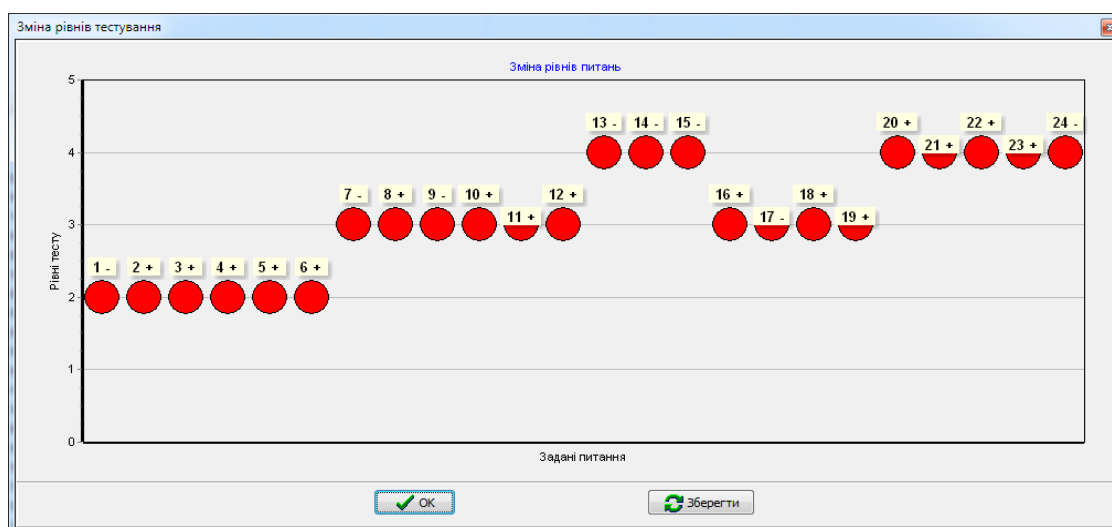


Рис. 3 – Приклад зміни рівнів при 4-х рівнях складності для тесту №5

В ході досліджень було виявлено, що є вплив кількості автоматично сформованих питань при адаптивному тестуванні на оцінку проходження тестів. Була замічена закономірність, що якщо здобувач має нестійкі знання та допускає часто помилки, то йому задається більше питань.

Таким чином, незалежно від того, яку стратегію побудови завдань обрано чи скільки їх включено до тесту, адаптивне оцінювання має гарантувати, що кожен сформований тест повною мірою відповідає змісту навчального матеріалу і точно відображає рівень підготовки здобувача

Висновки

Інформаційні технології поширилися і стали невід'ємною частиною всіх сфер життя сучасного суспільства, в тому числі освітнього середовища. З метою повноцінної реалізації існуючих можливостей повинні розроблятися особливі методи навчання, які розширять коло застосовуваних дидактичних прийомів.

Наразі комп'ютерне тестування має низку переваг перед традиційним тестуванням. Для контролю знань використовуються автоматизовані системи контролю знань і тестування, які дозволяють: полегшити роботу викладача при проведенні екзаменів і проміжних контролів; автоматизувати обробку результатів; зменшити витрати на організацію і проведення тестування; проводити тестування великої кількості здобувачів освіти з використанням широкого кола питань. Але поряд з цими перевагами мають місце і такі недоліки: відсутність залежності між використаними питаннями і відповідями, які отримуються від тестованого; питання вибираються з бази випадковим чином, тому що отримана оцінка ніяк не оцінює рівень підготовки здобувача вищої освіти з певної тематики і дуже залежить від випадковості.

Розроблена адаптивна система тестування адаптується до рівня знань здобувача з метою більш повної оцінки. Традиційно учасниками адаптивної системи є викладачі та здобувачі освіти. Викладачі заповнюють базу даних питаннями трьох основних типів (легкий, середній, складний) і встановлюють основні настройки тесту, такі як: кількість питань, кількість додаткових питань, мінімальний відсоток проходження, мінімальну і максимальну оцінку за тест. Адаптація по глибині опитування реалізується шляхом розбиття на рівні.

У подальшому планується удосконалити валідність тесту, що є найважливішим критерієм якості методу вимірювання. Валідність використовується для загальної характеристики тесту в аспекті відповідності одержаних результатів меті та вимогам оцінювання. А також застосувати наступні прийоми для визначення

валідності: аналіз бісеріального розподілу та аналіз індексу дискримінації.

Перелік використаних джерел

- [1] Radkevych O. Adaptive testing in the context of using electronic learning tools: essence, development and assessment. *Professional Pedagogics*. 2023. Vol. 1(26). Pp. 58-73. DOI: <https://doi.org/10.32835/2707-3092.2023.26.58-73>.
- [2] Василенко Я.П., Олексюк В.П. Особливості організації адаптивного тестування. *Сучасні цифрові технології та інноваційні методики навчання: досвід, тенденції, перспективи* : матеріали VII Міжнародної науково-практичної інтернет-конференції, м. Тернопіль, 8 квітня, 2021. Тернопіль : ТНПУ ім. В. Гнатюка, 2021. С. 11-14.
- [3] Meleško J., Ramanauskaitė S. Time saving students' formative assessment: Algorithm to balance number of tasks and result reliability. *Applied science*. 2021. Vol. 11, iss. 13. Article 6048. Pp. 1-15. DOI: <https://doi.org/10.3390/app11136048>.
- [4] Kürşad M.Ş. The Effects of Different Item Selection Methods on Test Information and Test Efficiency in Computer Adaptive Testing. *Journal of Measurement and Evaluation in Education and Psychology*. 2022. Vol. 14, iss. 1. Pp. 33-46. DOI: <https://doi.org/10.21031/epod.1140757>.
- [5] Wang P., Liu H., Xu M. An adaptive testing item selection strategy via a deep reinforcement learning approach. *Behavior Research Methods*. 2024. Vol. 56. Pp. 8695-8714. DOI: <https://doi.org/10.3758/s13428-024-02498-x>.
- [6] Item selection methods in multidimensional computerized adaptive testing for forced-choice items using Thurstonian IRT model / Q. Wang et al. *Behavior Research Methods*. 2023. Vol. 56. Pp. 600-614. DOI: <https://doi.org/10.3758/s13428-022-02037-6>.
- [7] Gorgun G., Bulut O. Incorporating test-taking engagement into the item selection algorithm in low-stakes computerized adaptive tests. *Large-scale Assessments in Education*. 2023. Vol. 11. Article 27. DOI: <https://doi.org/10.1186/s40536-023-00177-5>.
- [8] Xiao J., Bulut O. Item Selection With Collaborative Filtering in On-The-Fly Multistage Adaptive Testing. *Applied Psychological Measurement*. 2022. Vol. 46, iss. 8. Pp. 1-15. DOI: <https://doi.org/10.1177/01466216221124089>.
- [9] Моделювання адаптивного тестування знань: поріг ефективності, рівень складності та час виконання завдань / О.Ф. Шевчук та ін. *Вісник Вінницького політехнічного інституту*. 2025. № 1. С. 104-112. DOI: <https://doi.org/10.31649/1997-9266-2025-178-1-104-112>.

MODELING AND IMPLEMENTATION OF MULTI-LEVEL ADAPTIVE TESTING
IN AN INTELLIGENT SYSTEM

- Fedosova I.V.** D.Sc. (Pedagogical Sciences), professor, SHEI «Priazovskyi state technical university», Dnipro, ORCID: <https://orcid.org/0000-0003-3923-8270>, e-mail: fedosova_i_v@pstu.edu;
- Pronina O.I.** PhD (Engineering), associate professor, SHEI «Priazovskyi state technical university», Dnipro, ORCID: <https://orcid.org/0000-0001-7085-8027>, e-mail: pronina_o_i@pstu.edu;
- Veremiy V.O.** M.Sc., SHEI «Priazovskyi state technical university», Dnipro

The article is devoted to considering an approach to improving the quality of the educational process by implementing computer adaptive testing. The work focuses on the problems of traditional knowledge assessment, such as the lack of individualization and subjectivity of the teacher, and offers adaptive methods of forming tests that take into account the level of preparation of the education seeker in real time. Various algorithms of adaptive testing are described. A method of multi-level testing with a block structure is also described, where tasks change in accordance with previous answers. This method was used as the basis of an information system for checking the quality of knowledge of applicants. A formula for calculating the transition between blocks that outline the level of knowledge in the test was given. The user interface, types of questions were also described, the implementation of the testing system was presented, and the mechanism for changing difficulty levels during the test was also presented. The system interface is user-friendly and provides support for various types of test tasks, which makes it universal for use in educational institutions. The developed multi-level structure of the test allows you to flexibly adapt not only the complexity, but also the logic of the transition between tasks, increasing the accuracy and objectivity of the assessment, as well as reducing the overall duration of testing. The implemented methodology demonstrates the practical feasibility and flexibility of implementing an adaptive approach into the system of automated knowledge control, which can become the basis for creating more effective educational technologies.

Keywords: adaptive testing, intelligent system, individual test trajectory, knowledge assessment.

References

- [1] O. Radkevych, «Adaptive testing in the context of using electronic learning tools: essence, development and assessment», *Professional Pedagogics*, vol. 1(26), pp. 58-73, 2023. doi: 10.32835/2707-3092.2023.26.58-73.
- [2] Ya.P. Vasylenko, and V.P. Oleksiuk, «Osoblyvosti orhanizatsii adaptivnoho testuvannia» [«Features of organizing adaptive testing»], in Proc. of the VII Int. Sci.-Pract. Internet-Conf. «Modern digital technologies and innovative teaching methods: experience, trends, prospects», Ternopil, 2021. Pp. 11-14. (Ukr.)
- [3] J. Meleško, and S. Ramanauskaitė, «Time saving students' formative assessment: Algorithm to balance number of tasks and result reliability», *Applied science*, vol. 11, iss. 13, article 6048, pp. 1-15, 2021. doi: 10.3390/app11136048.
- [4] M.Ş. Kürşad, «The Effects of Different Item Selection Methods on Test Information and Test Efficiency in Computer Adaptive Testing», *Journal of Measurement and Evaluation in Education and Psychology*, vol. 14, iss. 1, pp. 33-46, 2022. doi: 10.21031/epod.1140757.
- [5] P. Wang, H. Liu, and M.Xu, «An adaptive testing item selection strategy via a deep reinforcement learning approach», *Behavior Research Methods*, vol. 56, pp. 8695-8714, 2024. doi: 10.3758/s13428-024-02498-x.
- [6] Q. Wang, Y. Zheng, K. Liu, Y. Cai, S. Peng, and D. Tu, «Item selection methods in multidimensional computerized adaptive testing for forced-choice items using Thurstonian IRT model», *Behavior Research Methods*, vol. 56, pp. 600-614, 2023. doi: 10.3758/s13428-022-02037-6.
- [7] G. Gorgun, and O. Bulut, «Incorporating test-taking engagement into the item selection algorithm in low-stakes computerized adaptive tests», *Large-scale Assessments in Education*, vol. 11, article 27, 2023. doi: 10.1186/s40536-023-00177-5.
- [8] J. Xiao, and O. Bulut, «Item Selection With Collaborative Filtering in On-The-Fly Multistage Adaptive Testing», *Applied Psychological Measurement*, vol. 46, iss. 8, pp. 1-15, 2022. doi: 10.1177/01466216221124089.
- [9] O.F. Shevchuk, A.A. Yarovyi, Yu.M. Panochyshyn, S.I. Petryshyn, and O.A. Kozlovskiy, «Modeliuvannia adaptivnoho testuvannia znan: porih efektyvnosti, riven skladnosti ta chas vykonannia zavdan» [«Modeling of Adaptive Knowledge Testing: Efficiency Threshold, Task Complexity and Completion Time»], *Visnyk Vinnytskoho politekhnichnoho instytutu – Visnyk of Vinnytsia Polytechnical Institute*, № 1, pp. 104-112, 2025. doi: 10.31649/1997-9266-2025-178-1-104-112.

Стаття надійшла 13.05.2025

Стаття прийнята 02.06.2025

Стаття опублікована 30.06.2025

Цитуйте цю статтю як: Федосова І.В., Проніна О.І., Веремій В.О. Моделювання та реалізація багаторівневого адаптивного тестування у інтелектуальній системі. *Вісник Приазовського державного технічного університету*. Серія: Технічні науки. 2025. Вип. 50. С. 25-30. DOI: <https://doi.org/10.31498/2225-6733.50.2025.336246>.