

Л. І. Гладка, Черкаський національний університет імені Богдана Хмельницького

СИСТЕМНИЙ ПІДХІД ДО ОЦІНКИ ЯКОСТІ ЗНАНЬ У ФОРМІ КОМП'ЮТЕРНОГО ТЕСТУВАННЯ

Гладка Л. І.

Системний підхід до оцінки якості знань у формі комп'ютерного тестування

Робота присвячена реалізації системного підходу до організації процесу тестування для забезпечення якісного контролю знань студентів. Проведено огляд онлайн сервісів, призначених для формування банку завдань та організації тестового контролю знань. Вперше запропонована структура пакету тестових завдань для контролю знань, яка узгоджена зі шкалою ECTS. Використовуючи онлайн систему тестування „Online Test Pad”, проведено апробацію запропонованої методики організації процесу тестування на прикладі організації контролю знань з дисципліни „Алгоритмізація та програмування”.

Ключові слова: тести, типи тестових завдань, диференційована оцінка, онлайн система тестування.

Гладка Л. И.

Системный подход к оценке знаний в форме компьютерного тестирования

Работа посвящена реализации системного подхода к организации процесса тестирования для обеспечения качественного контроля знаний студентов. Проведен обзор онлайн сервисов, предназначенных для формирования банка заданий и организации тестового контроля знаний. Впервые предложена структура пакета тестовых заданий для контроля знаний, которая согласована со шкалой ECTS. Используя онлайн систему тестирования „Online Test Pad”, проведена апробация предложенной методики организации процесса тестирования на примере организации контроля знаний по дисциплине „Алгоритмизация и программирование”.

Ключевые слова: тесты, типы тестовых заданий, дифференцированная оценка, онлайн система тестирования.

За нинішніх умов організації навчального процесу, що передбачають систематичний контроль знань, виникає потреба автоматизації процесу

контролю знань за рахунок використання відповідного програмного забезпечення. Як неодноразово відзначалося дослідниками, комп'ютерне тестування – це ефективний спосіб перевірки [1 – 4]. Однією з переваг комп'ютерного тестування є можливість організації контролю без обмеження його періодичності, забезпечуючи миттєве отримання достовірних результатів контролю відразу після закінчення процесу тестування. Від традиційного контролю знань – тести відрізняються об'єктивністю вимірювання результатів навчання, оскільки вони спираються не на суб'єктивну думку викладачів, а на об'єктивні критерії. До того ж, результати автоматизованого тестування краще піддаються аналізу.

Перевагами вимірювання та формалізованого оцінювання знань з використанням комп'ютерних систем тестування є: можливість автоматизації контролю знань і оцінювання результатів тестування; зручність підготовки тестів; велика кількість запитань, якими можна охопити дисципліну в цілому; довільна тривалість тестування; інтеграція систем тестування з відокремленими базами даних тощо. Тому, важливу роль при організації проведення комп'ютерного тестування відіграє вибір програмного забезпечення як для формування банку тестових завдань так і для реалізації процедури тестування. Нині стрімкого розвитку та поширення набувають онлайн сервіси, до яких переваг використання яких спонукають користувачів їх мобільність, кросплатформенність та можливість миттєвого використання.

Метою роботи є реалізація системного підходу до оцінки якості знань у формі комп'ютерного тестування.

Питання теорії оцінки якості знань у формі комп'ютерного тестування, структури і форм тестових завдань, алгоритмів обробки результатів тестування та їх інтерпретації досліджено у працях К. Інгенкампа, Дж. Равена В. С. Аванесова, Ю. М. Богачкова, Я. С. Бродського, І. Є. Булах, А. Н. Землякова, Ю. І. Іванова, І. Алексейчук, Н. Тализіна, В. Бочарникова, Н. Шиян та ін. [4, 5, 6, 7].

Розглянемо **системний підхід** до організації комп'ютерного тестування.

Виділимо основні етапи системного підходу до організації комп'ютерного тестування:

1. Підготовчий етап.

✓ На матеріалах навчально-методичного комплексу дисципліни сформувані систему контролю знань: кількість контрольних заходів, форми контролю, шкала оцінювання.

✓ Відповідно до системи контролю знань розробити пакети завдань та питань: теоретичні питання, практичні завдання (різномірні), тестові завдання (деталі формування пакету тестових завдань розглянуті нижче).

✓ Відповідно до розробленого пакету завдань сформувані базу тестових завдань для самопідготовки (деталі формування пакету тестових завдань для самопідготовки розглянуті нижче).

2. Уточнюючий етап. Уточнюється методика контролю, склад тестових завдань, здійснюється попередній розрахунок показників і параметрів контролю.

На уточнюючому етапі необхідно здійснити наступні кроки:

➤ Апробація тесту викладачем-розробником у формі самоперевірки на виявлення технічних, синтаксичних помилок.

➤ Уточнення розробленої системи тестових завдань групою експертів (наприклад, методом Дельфі).

Дельфійський метод полягає в проведенні незалежного і неодноразового анкетування експертів. Як експертів пропонується залучити членів навчально-методичної ради відповідного напрямку підготовки. Перша анкета містить загальні питання. Після заповнення експертами анкет, вони повертаються розробникам тесту і обробляються. На основі результатів першої анкети формується друга. У ній експерти ранжують окремі елементи, відображають згоду або незгоду з оцінкою параметрів тестового завдання розробниками. На основі результатів другого анкетування складається третя анкета. Експерти розглядають їх попередні відповіді на фоні колективної думки. Третє анкетування є заключним.

3. *Етап організації контролю знань в студентських групах.* Після проведення тестування, обробити результати тестування та розрахувати наступні параметри:

- коефіцієнт складності завдання K_c (відношення кількості студентів, які не вирішили відповідне завдання, до загального числа студентів групи). На даному етапі необхідно переглянути та доопрацювати занадто прості тестові завдання (значення K_c близьке до нуля) і складні (значення K_c близьке до одиниці);

- дискримінаційний індекс. Значення дискримінаційного індексу лежить в діапазоні між $(-1 \div +1)$. У випадку від'ємного значення слід відбракувати завдання, оскільки від'ємне значення дискримінаційного індексу означає, що на дане питання тесту слабкі студенти відповідають краще, ніж сильні;

- дискримінаційний коефіцієнт - показник, що дозволяє оцінити якість питання. Він являє собою коефіцієнт кореляції між балами, отриманими тестованим студентом з конкретного завдання, і його загальною оцінкою за проходження всього тесту. Від'ємне значення дискримінаційного коефіцієнта свідчить про некоректність аналізованого питання з точки зору правильності оцінки рівня засвоєння навчального матеріалу.

Визначимо основні положення пропонованого підходу до формування пакету тестових завдань:

1. Спочатку рекомендується визначити модель знань студента.

2. Вкрай необхідно, щоб при побудові пакету тестових завдань виконувались загальні вимоги до побудови педагогічних тестів та відомі чотири складові якості тестів [8].

3. Форма подання тестових завдань має бути різноманітною, різнорівневою та враховувати можливість їх використання для різних видів контролю поточного стану знань студента: вхідного, поточного (самоконтролю та модульного контролю) та підсумкового.

4. При розробці пакету тестових завдань повинні використовуватись вікові нормативи та інші стандарти, оформлені документально.

5. Обов'язковим є застосування принципів змістової валідності. Валідність за змістом вважається врахованою, якщо 70% змісту курсу передбачає наявність фактичних знань, то, відповідно, 70% тестових завдань повинні перевіряти саме фактичні знання.

Рекомендується використовувати п'ятикомпонентну модель знань студента. Така модель має враховувати поділ знань на декларативні та процедурні.

Зміст пакету тестових завдань повинен повністю відповідати меті його створення, а також документам, що визначають зміст навчальної дисципліни.

Для врахування якості тестів розглянемо кожен складову окремо.

Інтелектуальна складова якості тестів включає:

- повноту охоплення тестовими запитаннями усіх положень навчального курсу;
- кількість помилок (логічних, мовленнєвих) у формулюваннях запитань та відповідей;
- охоплення запитаннями рівнів освоєння дисципліни: володіння термінологією, розуміння законів, вирішення практичних завдань.

Для її врахування необхідно:

- ставити запитання до кожного положення лекційного матеріалу;
- перед запровадженням тестів у використання потрібно провести їх апробацію;
- будувати запитання на знання термінології, законів, вміння розв'язувати практичні завдання.

Статистична складова якості тестів дає можливість врахувати ймовірність випадкового попадання на правильні варіанти відповідей: наприклад, при 10 запитаннях типу „Одна відповідь з багатьох” з 5-ма варіантами відповідей, з яких одна правильна, у 2-х запитаннях можна випадково потрапити на правильні відповіді. Для врахування статистичної складової якості тестів необхідно застосовувати у тестах не менше чотирьох-п'яти варіантів відповідей; використовувати типи запитань „багато з багатьох”, „впорядкування”,

„відповідність” та „відкрите запитання”; враховувати ймовірність випадкового попадання на правильну відповідь, встановивши нижній прохідний результат на рівні не менше 60% від максимально можливого. Наприклад:

1. Як оголосити макрос, який замінює SIZE на значення 5 на етапі прекомпіляції? Завдання на 1 бал, базовий рівень, розділ „Елементи мови Cі”.

- #define SIZE = 5
- #define SIZE 5
- const SIZE = 5;
- #const SIZE 5
- define SIZE 5

2. У яких рядках коду містяться помилки? Завдання на 3 бали, високий рівень, розділ «Типи даних в Cі».

```
struct A {
    char a[];      // 1
};
struct B {
    char a;
    char b[];     // 2
};

struct C {
    char a[];     // 3
    char b;
};
int main() {
    char a[];    // 4
    return 0;
}
void test(char a[]) { // 5
}
```

- // 1
- // 2
- // 3
- // 4
- // 5
- в інших рядках
- помилки відсутні

Психологічна складова якості тестів враховує повтор термінів у формулюваннях запитання та правильних відповідей; легкість формального запам'ятовування правильних відповідей. Тому пакет тестових завдань на модуль повинен перевищувати 100 запитань (а вибірка для кожного студента – близько 25 запитань); при формулюванні неправильних відповідей використовувати слід вживати «правильні» фрази. Наприклад:

1. Яким кодом можна замінити постфіксний інкремент в рядку 1?

```
#include <stdio.h>
main() {
    int a = 1;
    int b = 2;
    int * array[] = { &a, &b };
    int ** pp = array;
    (*++pp)++; // 1
}
```

- a+=1;
- b+=1;
- array[0]+=1;
- **array[1]+=1;**
- pp+=1;

Дисциплінарно-організаційна складова дозволяє врахувати можливість списування та підказок. Тому обов'язковим має бути встановлення ліміту часу на задачу тесту (<1 хв. на просте запитання).

В існуючих комп'ютерних системах тестування необхідно використовувати якомога більшу кількість запитань різних типів [9]:

- з вибором одного варіанту відповіді;
- з множинним вибором;
- на встановлення відповідності;
- на встановлення послідовності;
- на ранжування;
- на доповнення (з закритою та відкритою відповіддю);
- на роботу з графічним зображенням.

Це дозволить адекватно оцінити стан знань студента і відповідно встановити рівень його компетентності відповідно до вимог однієї з головних складових освітнього стандарту – освітньо-професійної програми підготовки фахівця при використанні різних видів контролю, які проводяться з метою:

- вхідний контроль – для визначення рівня підготовки студентів з тих дисциплін, які вивчалися перед вивченням певної дисципліни, або загального рівня підготовки студентів за попередній період навчання;

- поточний контроль – для визначення поточного рівня підготовки студентів та оцінки засвоєння навчального матеріалу, запланованого на самостійне опрацювання;

- підсумковий контроль – для визначення рівня засвоєння студентом теоретичного та практичного матеріалу з окремої навчальної дисципліни.

Значні труднощі у вітчизняних педагогів викликає питання ранжування тестових завдань підсумкового контролю за рівнями складності та їх відповідності параметрам, заявленим у відповідному освітньому стандарті. Подібних труднощів можна уникнути, якщо дотримуватись наступних правил:

1. Тестові завдання слід розділити на три категорії, кожна з яких визначає певний рівень знань з вибраної дисципліни: достатній, середній або високий.

2. Дотримуватись рекомендацій Міністерства освіти і науки України щодо підсумкового оцінювання знань.

3. Узгодити структуру оцінювання з вимогами кредитно-модульної системи організації навчального процесу (КМСОНП), яка ґрунтується на поєднанні модульної організації навчального процесу та вимірі навчального навантаження студентів в кредитах європейської кредитно-трансферної та акумулюючої системи ECTS. Система ECTS класифікує студентів по рівням знань і таким чином робить інтерпретацію ранжирування більш простою. Використовуючи шкалу ECTS для забезпечення перевірки рівня знань, необхідно при формуванні пакета тестових завдань (як і вибірки) використовувати завдання різних рівнів. Для цього пакет завдань пропонується розбити на три групи:

- 60% – завдання достатнього рівня, які оцінюються в 2% кожне завдання від загальної суми, рівної 100% (рівень DE);

- 25% – завдання середнього рівня, які оцінюються в 6% кожне завдання від загальної суми, рівної 100% (рівень BC);

- 15% – завдання високого рівня, які оцінюються в 10% кожне завдання від загальної суми, рівної 100%(рівень A).

Наприклад, якщо вибірка містить 40 тестових запитань, то з них 24 запитання має бути по 1 балу кожне, 10 – по 3 бали кожне, 6 – по 5 балів кожне.

Зазначимо основні результати досліджень. Для підготовки студентів до тестування сформована база тестових завдань для самопідготовки. При формуванні бази проведено огляд онлайн сервісів, призначених для організації тестового контролю знань.

Онлайн сервіси організації тестового контролю знань розділимо на дві категорії:

1. сервіси, що передбачають формування власного банку тестових завдань;
2. системи з вільно доступним банком тестових завдань.

У табл. 1 подано огляд найпопулярніших онлайн сервісів для формування тестів і вправ та організації різних видів контролю (попереднього, поточного, тематичного, періодичного, підсумкового, самоконтролю), які умовно можна віднести до першої категорії.

Таблиця 1

**Онлайн сервіси першої категорії
для організації тестового контролю знань**

Адреса розміщення сервісу / Призначення сервісу	Можливості сервісу
http://www.classmarker.com/ сервіс для формування тестів, з англomовним інтерфейсом	Публікація тестів на сервісі, друк тестів
http://learningapps.org сервіс для підтримки навчального процесу з використанням інтерактивних модулів, інтерфейс підтримує кілька мов, в т.ч. російську	формування різних видів діаграм з використанням вбудованого інструменту Mindmap або експорт готових діаграм, формування аудіо/відео контенту; можливість розробки інтерактивних вправ; можливість перенесення завдань у власну html- сторінку; можливість обміну інтерактивними завданнями
http://master-test.net/uk сервіс для формування тестів з українським інтерфейсом	має конструктор зі стандартним функціоналом
http://quizlet.com/ сервіс для створення і	можливість розробки флеш-карток на у трьох формах: 1) Flashcard Mode – це вид флеш-

використання флеш-вправ і навчальних ігор	карток, схожий на традиційний паперовий тип; 2) Learn Mode – це вид флеш-карток, який відображає правильні/неправильні відповіді і пропущені запитання; 3) Test Mode – вид флеш-карток, який дозволяє вибрати типи питань: вписування відповідей, вибір відповіді, зіставлення, відповідь так/ні.
http://www.quia.com/ міжнародна освітня програма вивчення англійської мови, онлайн конструктор тестів	можливість розробки тестових завдань з використанням аудіофайлів і рисунків, навчальних ігрових завдань
http://onlinetestpad.com багатофункціональний онлайн конструктор тестів «Online Test Pad»	гнучкий підрахунок результатів тестів; можливість побудови графіків і гістограм за результатами обробки тесту.
http://www.banktestov.ru/ сервіс для формування тестів з російськомовним інтерфейсом	має конструктор зі стандартним функціоналом
http://www.qreature.ru сервіс для формування тестів	можливість розробки тестових завдань з використанням аудіо/відео файлів і рисунків, які зберігаються на інших серверах

Табл. 2 містить низку сервісів другої категорії, які дають можливість сформувавши для широкого кола навчальних дисциплін, що стосуються програмування, банк завдань для забезпечення різних видів контролю.

Таблиця 2

Онлайн сервіси другої категорії для організації тестового контролю знань з навчальних дисциплін, що стосуються програмування

Адреса розміщення сервісу / Мови програмування, для яких наявні тести	Особливості
http://www.cquestions.com/ мова програмування C	тестові завдання розділені за темами
http://www.pskills.org/c.jsp мови програмування C/C++, C#, Java, HTML	тестові завдання на знання основ мови програмування, не розділені за темами
http://www.quizful.net/test мови програмування C/C++, C#, Java, JavaScript, HTML, Ruby та інші	тестові завдання на знання основ мови програмування, не розділені за темами

https://www.expertrating.com мови програмування C/C++, C#, Java, JavaScript, HTML, Perl, Ruby та інші	міжнародна система експертної оцінки знань ІТ спеціалістів, безкоштовне тестування лише для жителів США
www.brainbench.com мови програмування C/C++, C#, Java, JavaScript, HTML, Perl, Ruby та інші	міжнародна система експертної оцінки знань ІТ спеціалістів, більшість тестів платні

Принципи використання розглянутих сервісів подібні, тому кожен викладач має змогу вибрати найоптимальніший шлях забезпечення процесу самопідготовки.

У роботі тестування організовано з використанням багатofункціонального онлайн конструктору тестів „Online Test Pad”:

1. Розроблено банк тестових завдань. Для апробації описаної в роботі методики формування структури пакета тестів, розроблено пакет тестових завдань з дисципліни „Алгоритмізація і програмування”, який складається з 60 питань (з 250 питань банку тестових завдань).

2. Підключаємо банк тестових завдань, використовуючи онлайн сервіс „Онлайн тест Pad” (рис. 1).

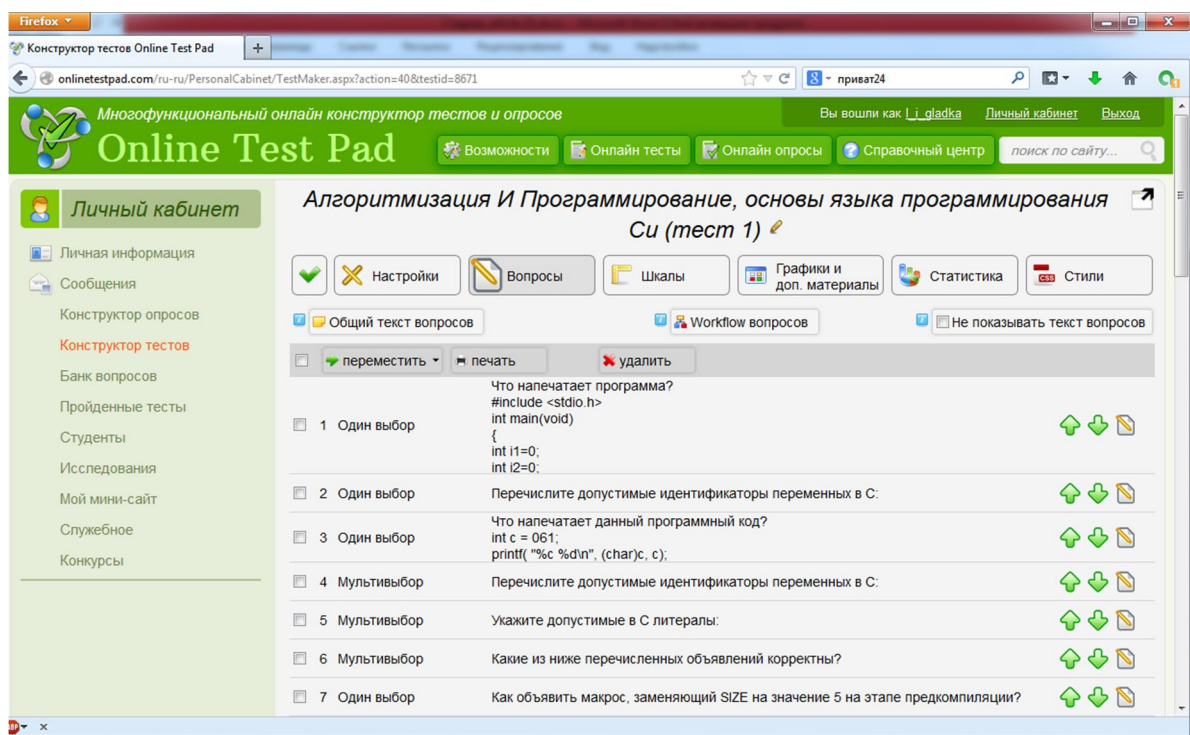


Рис. 1 Розробка тесту з використанням онлайн сервісу „Online Test Pad”

3. Налаштовуємо доступ до тесту. Проводимо тестування.

4. Аналізуємо результати тестування.

Отримані результати тестування корелюють з поточною успішністю. Можливість побудови графіків і гістограм за результатами обробки тесту в системі «Інтернет Тест Рад» дозволяє автоматизувати процес побудови траєкторії успішності кожного студента.

Отже, розглянувши основні положення системного підходу до оцінки якості знань у формі комп'ютерного тестування, можна зробити наступні висновки:

1. конструювання тестів вимагає використання системного підходу і складається з відповідних етапів.

2. розробка досконалого пакету тестових завдань вимагає від викладача не лише прив'язки до контексту та навчальних цілей курсу у вигляді специфічних оперативних термінів, але й врахування моделі знань студента для адекватної оцінки індивідуальних можливостей.

3. широкий спектр онлайн сервісів дає можливість організувати підготовку до підсумкового контролю у формі самопідготовки.

4. використання системного підходу для формування тестових завдань гарантує об'єктивність оцінки.

У подальших дослідженнях системного підходу до оцінки якості знань у формі комп'ютерного тестування планується розробка алгоритму автоматичного формування тестового завдання з використанням банку тестів обраної дисципліни.

Література

1. **The Bologna Declaration on the European space for higher education: an explanation [Electronic Resource].** – Retrieved from <http://ec.europa.eu/education/policies/educ/bologna/bologna.pdf>.

2. **Обучение и искусственный интеллект или основы современной дидактической высшей школы** // Донецк : ДООУ. – 2002. – 504 с.

3. **Ефимов Е. Н.,** Денисов М. Ю., Жилина Е. В. Сравнительный анализ образовательных систем тестирования по критерию функциональной полноты [Электронный ресурс] // Управление экономическими системами. – 2012. – № 4. – Режим доступа : <http://www.uecs.ru/instrumentalnii-metody-ekonomiki/item/1236-2012-04-09-07-23-03>.
4. **Ингенкамп К.** Педагогическая диагностика / К Ингенкамп. – М., 1991. – 240 с.
5. **Иванов Ю. И.** Разработка тестов и методы контроля знаний по дисциплинам // Качество, инновации, образование. – 2008. – № 7. – С. 11 – 17.
6. **Майоров А. Н.** Теория и практика создания тестов для системы образования // М. : Интеллект–центр. – 2001. – 296 с.
7. **Nathan A.** Thompson, David J. Weiss. A Framework for the Development of Computerized Adaptive Tests // Practical Assessment, Research & Evaluatio. Vol. 16. No 1. Retrieved from <http://pareonline.net/pdf/v16n1.pdf>.
8. **Tan Jin,** Barley Mak, and Pei Zhou. Confidence scoring of speaking performance: How does fuzziness become exact? // Language Testing. 2012. No 29 (1). P. 43 – 65.
9. **Hossien** Fallahzadeh New software for computation sensitivity analysis to detect hidden bias for partially order set test statistic in observational studies // Procedia Technology. 2012. Vol. 1. P. 225 – 229.
10. **J. Sua'rez,** A. Franco, R.A. Santos. Order Statistics and Item Bank Analysis in Computer Adaptive Testing // Procedia Technology. 2013. Vol. 7. P. 273 – 281.
11. **Лефрансуа Ги.** Прикладная педагогическая психология. Психология для учителя // Пер. с англ. Волохонский В., Миронов Н. [11–ое изд.]. – СПб. : Прайм-ЕВРОЗНАК. – 2005. – 416 с.
12. **Hensley G.** Creating a Hybrid College Course : Instructional Design Notes and Recommendations for Beginners // Journal of Online Learning and Teaching. 2005. Vol. 1, No. 2. P. 66 – 78.

Hladka L. I.

System Approach to Knowledge Quality Evaluation by Means Of Computer Testing

Article is devoted to the unified development of the methodical approach to generate tests for qualitative monitoring of students knowledge from all forms of learning. The review of online services designed to form a bank of tasks and knowledge testing. In article the structure of the package tests for control knowledge which agreed with the scale of ECTS first proposed. Using online testing system „Online Test Pad”, conducted testing of the proposed method of forming a package of tests on the example of the organization of the control of knowledge with subject „Algorithms and Programming”.

Key words: Tests, Types of Tests, Differentiated Evaluation, Online Testing System.

Відомості про автора

Гладка Людмила Іванівна – кандидат фізико-математичних наук, доцент кафедри автоматизації та комп’ютерно-інтегрованих технологій Черкаського національного університету імені Богдана Хмельницького. Коло наукових інтересів: інформаційно-комунікаційні технології в освіті, проектування баз даних, інтелектуальний аналіз даних.

Стаття надійшла до редакції 03.02.2014 р.

Прийнято до друку 25.02.2014 р.

Рецензент – к. е. н., проф. Гужва В. М.