УДК 378.016:004

Мусієнко М. П., Чорномаз І. К.

МЕТОДИ ФОРМУВАННЯ ЕЛЕКТРОННИХ ПІДРУЧНИКІВ ЗА КЛЮЧОВИМИ СЛОВАМИ ІЗ ЗАДАНИМ ГРАНИЧНИМ РОЗМІРОМ

У статті описано два розроблені методи формування електронних підручників із заданим граничним розміром при запиті по одному з двох можливих варіантів ключових слів: за стандартним (із заздалегідь запропонованого списку) поняттям (терміном); за випадковим поняттям, яке неможливо передбачити на стадії розробки електронного підручника. Описані методи можуть бути використані в комп'ютеризованих системах з відносно невеликим об'ємом пам'яті: КПК, електронні книги тощо, для формування електронних підручників з обмеженим розміром з інформацією про поняття, що цікавить користувача.

Ключові слова: *електронний підручник, базове і випадкове ключове слово, обмежений розмір.*

В статье описаны два разработанных метода формирования электронных учебников с заданным граничным размером при запросе по одному из двух возможных вариантов ключевых слов: по стандартному (из заранее предложенного списка) понятию (термину); по случайному понятию, которое невозможно предусмотреть на стадии разработки электронного учебника. Описанные методы могут быть использованы в компьютеризированных системах с относительно небольшим объемом памяти: КПК, электронные книги и другие, для формирования электронных учебников с ограниченным размером с информацией об интересующем пользователе понятии.

Ключевые слова: *электронный учебник, базовое и случайное ключевое слово, ограниченный размер.*

Two developed methods of electronic textbooks formation with the set of boundary sizes at the query of one or two possible variants of keywords: by the standard concept (from the list, that was offered beforehand); by the casual concept, which is impossible to foresee at the stage of development of electronic textbook, have been developed in the article. The described methods can be used in the computerized systems with the relatively small volume of memory: PC, e-books etc., for the forming of electronic textbooks with the limited measuring and with the information about the requested concept.

Key words: *electronic textbook, base and casual keyword, boundary Sizes*

Аналіз проблеми і постановка завдання

Однією з важливих складових навчально-методичного забезпечення системи освіти є підручник. Проте в Україні існує гостра нестача друкарських видань: аналіз видань наукової і учбової літератури показує, що середньостатистичне забезпечення ВНЗ літературою складає за кількістю екземплярів – 9,3 %, за кількістю найменувань – 5,8 % [1]. Ця проблема може бути частково розв'язана за рахунок випуску, поширення і використання електронних підручників (ЕП). На відміну від паперового варіанту, електронний підручник може мати ряд додаткових переваг: пошук інформації за ключовими словами, гіпертекстовий зміст, глосарій термінів тощо. Особливо актуальна необхідність електронного підручника в системі дистанційного навчання.

Традиційно робота з електронними підручниками здійснювалася на персональному комп'ютері, який характеризувався наявністю достатнього об'єму постійної пам'яті. Проте останнім часом все частіше в повсякденне життя людини входять різні переносні електронні пристрої (так звані гаджети): КПК, смартфони, електронні книги тощо, які використовуються і для зберігання і читання електронних підручників. І якщо зберігання електронного варіанту підручника (як правило, у форматах *.txt, *.doc, *.pdf) не представляє складності з точки зору об'єму, який він займає на носії, то зберігання повноцінного електронного підручника, який включає безліч мультимедійних вставок вже може потребувати достатніх об'ємів постійної пам'яті на носії. В цьому випадку об'єм займаного електронного підручника і методи його зменшення, при необхідності, мають важливе значення.

Таким чином для зберігання ЕП на різних гаджетах потрібна можливість приведення розміру усього ЕП до заданого значення, яка досягається за рахунок зменшення розміру його складових (як правило, за рахунок зменшення якості графічних, аудіо- і відеоматеріалів) або видалення менш значущих компонент (наприклад, видалення яких-небудь додаткових матеріалів, які не несуть основного смислового навантаження). Це накладає жорсткі вимоги до структури ЕП: максимальна деталізація його компонентів, класифікація важливості матеріалу, що представляється тощо.

Зараз існує добре відпрацьована методологія створення ЕП і інших комп'ютерних навчальних систем. Важливим елементом при розробці ЕП є розробка його структури, яка описується за допомогою моделі представлення учбового матеріалу дисципліни. І якщо питаннями розробки структури представлення учбового матеріалу ЕП було приділено достатньо уваги в роботах багатьох вчених [2; 3], то проблема приведення об'ємів підручника до заданого розміру залишалася мало вивченою. Деякі роботи авторів статті в цій області присвячені методам перетворення усього об'єму електронного підручника [4], проте бувають ситуації, коли користувачу немає необхідності вивчати увесь підручник або який-небудь розділ, а його цікавить інформація за якимось конкретним поняттям або терміном. Наприклад, якщо розглядати ЕП з медичної тематики, то це може бути «електрофорез», «електрокардіограф» і так далі, які не є назвами розділів і підрозділів, а містяться тільки в якихось невеликих кадрах (під кадром тут розуміється будь-який структурний елемент ЕП, що містить учбовий матеріал з/без посиланнями на інші структурні елементи без їх змісту в поточному кадрі). Отже це питання залишалося не вивченим.

Метою роботи є розробка методів формування електронних підручників, що містить інформацію за якимось поняттям (за ключовим словом, яке задає користувач) із заданим обмеженим розміром підручника.

Виклад основного матеріалу

Можливі дві ситуації запиту за ключовим словом:

– запит на частину ЕП, який містить інформацію за якимось стандартним (із заздалегідь запропонованого списку) поняттям (терміном), тобто запит за базовим ключовим словом (БКС);

– запит на частину ЕП, який містить інформацію за якимось випадковим поняттям (терміном), тобто запит за випадковим ключовим словом (ВКС).

Базове ключове слово створюється при розробці ЕП і визначається автором цього підручника. Увесь список таких БКС пропонується користувачеві перед отриманням ним ЕП. Випадкові ж ключові слова формуються самим користувачем за його бажанням і не можуть бути визначені на стадії розробки ЕП.

Розглянемо спочатку метод формування ЕП для першого випадку – при запиті за БКС. Список таких базових слів формується автором (методистом) і є одним з базових складових ЕП. Загальна схема трансформації ЕП для такого випадку показана на рис. 1.

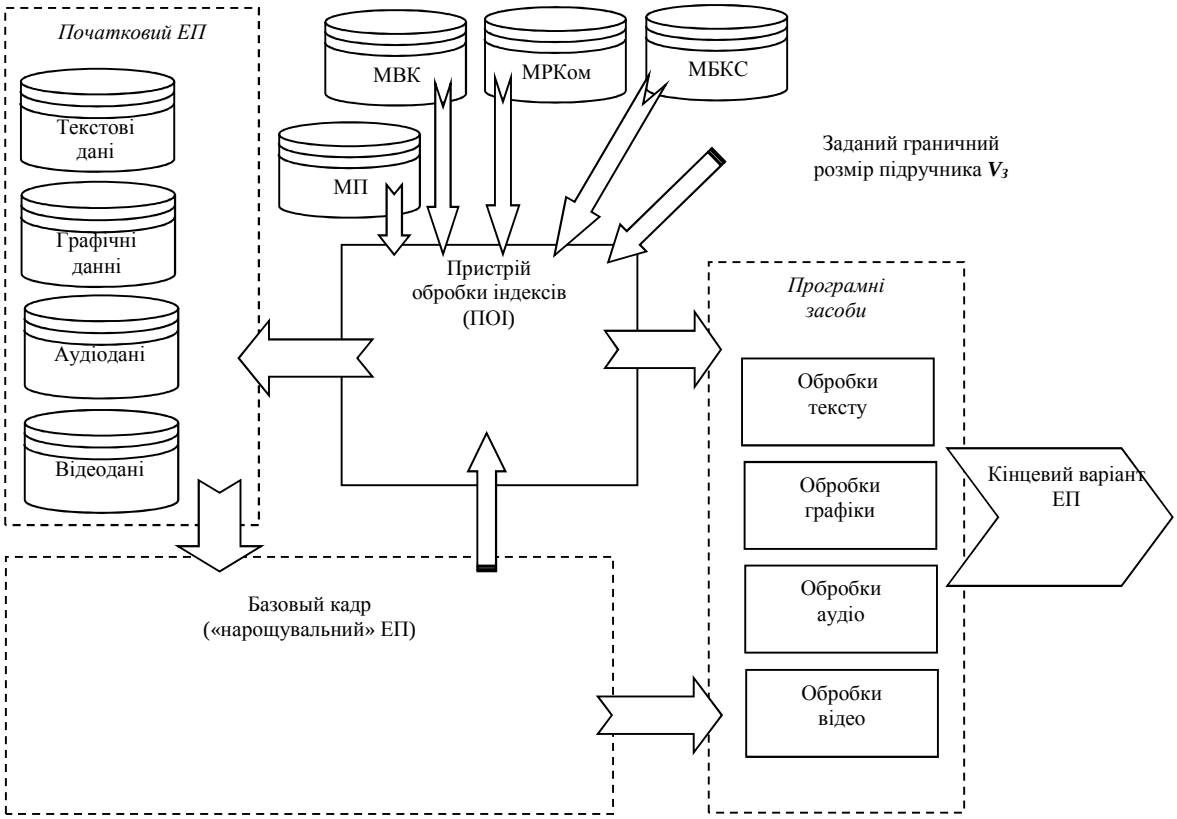

Рис. 1. Схема формування ЕП при запиті за базовим ключовим словом

Ця схема і метод трансформації схожі на описані в роботі [4] для отримання усього ЕП із заданим граничним розміром. Проте тут додатково використовується матриця базових ключових слів (МБКС), яка не має аналога у відомих системах розробки ЕП (приклад заповнення матриці показаний на рис. 2). МБКС визначає важливість того або іншого розділу для кожного ключового слова. Ця матриці близька до матриці логічних зв'язків [5], проте по вертикалі відкладені не індекси елементів учбового матеріалу (як по горизонталі), а наявні базові ключові слова. В таблиці проставляються вагові коефіцієнти, які вказують важливість (за пріоритетом) того або іншого розділу для вивчення матеріалу за ключовим словом. Вагові коефіцієнти не можуть бути рівними в межах одного рядка.

	1	1.1	1.1.1	1.1.2	1.1.3	1.2.1	…	2.1.1	2.1.2	…	*n.m.l-2*	*n.m.l-1*	*n.m.l*
$БКС_1$													
$БКС_2$	1	3	2	4	5	6		44	45		93	92	91
$БКС_3$	2	1	3	5	6	8		56	55		82	83	84
$БКС_4$	42	43	45	47	48	49		78	78		12	11	10
$БКС_5$	25	26	27	28	29	31		54	55		38	37	36
…													
$БКС_{N-1}$	74	75	77	76	78	79		14	15		41	42	44
$БКС_N$	19	21	23	25	20	22		45	47		2	3	5

Рис. 2. Приклад заповнення матриці базових ключових слів

Укрупнена схема алгоритму роботи пристрою обробки індексу (ПОІ) при реалізації методу формування ЕП за запитом за БКС показано на рис. 3. Як видно з рис. 3, в ньому використовується алгоритм (АЛГ 1), описаний в [4], який використовують при перетворенні підручника при запиті на весь ЕП.

Якщо розмір базового кадру ЕП буде більше заданого VЗ, то відбувається перший процес – зменшення. В цьому випадку використовується алгоритм для трансформації за запитом на увесь підручник [4].

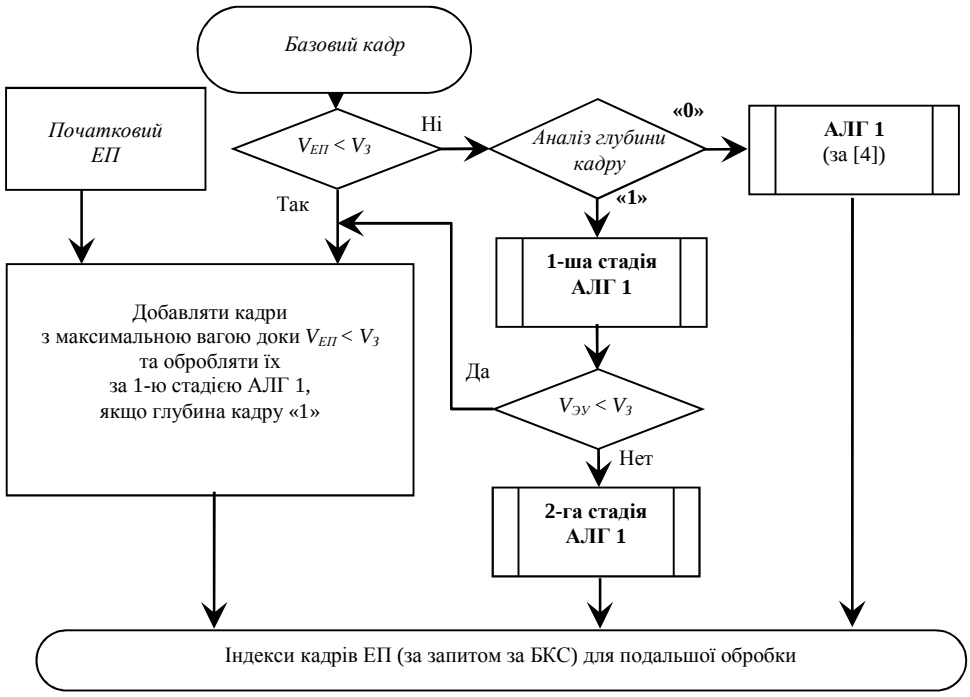

Рис. 3. Укрупнена схема алгоритму роботи пристрою обробки індексів (ПОІ) при реалізації методу формування ЕП за запитом за БКС

Процес перетворення відбувається до тих пір, доки розмір перетвореного ЕП буде не менше заданого значення $V_З$, тобто доки не виконуватиметься умова:

$$\sum_{<ind>} P_{T_{t_1,t_2}^{<ind>}} + \sum_{<ind>} \sum_{n_{<ind>}=1}^{N_{<ind>}} P_{G_{g_1,g_2,g_3,g_4}^{<ind>,n_{<ind>}}} + \sum_{<ind>} \sum_{m_{<ind>}=1}^{M_{<ind>}} P_{A_{a_1,a_2,a_3,a_4,a_5}^{<ind>,m_{<ind>}}} +$$

$$+ \sum_{<ind>} \sum_{l_{<ind>}=1}^{L_{<ind>}} P_{V_{v_1,\ldots,v_9}^{<ind>,l_{<ind>}}} < V_{Ç}, \qquad (1)$$

де *IND* – безліч усіх індексів структурних елементів підручників;

P_T, P_G, P_A, и P_V – розміри текстових, графічних, аудіо- і відеофайлів відповідно;

$N_{<ind>}$, $M_{<ind>}$, $L_{<ind>}$ – кількість графічних, аудіо- і відеокомпонент відповідно в кадрі з індексом <*ind*>;

t_1, t_2 – індекси масивів з розмірами текстових компонент;

g_1, g_2, g_3, и g_4 – індекси масивів з розмірами графічних компонент;

a_1, a_2, a_3, a_4 и a_5 – індекси масивів з розмірами аудіокомпонент;

v_1,…, v_9 – індекси масивів з розмірами відеокомпонент.

Якщо після обробки на якійсь із стадій з'являється вільний об'єм, він заповнюється можливими кадрами з меншим об'ємом. Для цього використовуються перевірки отриманих об'ємів і, залежно від результату, використовується або весь перший алгоритм, або тільки його частини (конвертації або видалення).

Коли ж початковий базовий кадр менше заданого граничного розміру V_3, відбувається другий процес – додавання кадрів до ЕП – «нарощування» кадру за певними правилами (згідно з пріоритетами з матриці властивостей кадрів МВС [4]).

Таким чином, перетворення ЕП з початкового базового кадру шляхом обростання до необхідного об'єму V_3 дозволяє сформувати таку частину ЕП, яка матиме максимальноважливу інформацію щодо базового ключового слова, включаючи усі додаткові складові, які максимально розкриватимуть інформацію за поняттям, яке цікавить користувача.

На рис. 3 зазначена умова «глибини» кадру, що вказує напрям «розростання» ЕП. Значення «1» відповідає напряму в глибину, «0» – на якість. Наприклад, при трансформації ЕП з'являється вибір: зменшити якість компонент кадру (його графічні, аудіо- і відео компоненти тощо) і взяти більше інших кадрів – шлях в глибину; чи зберегти максимальну якість усіх компонент цього кадру – шлях у бік якості. Наприклад, якщо це лінгафонний підручник, де важливою є якість звуку або дизайнерський, де важливою є якість рисунка, то необхідно йти по другому шляху; а якщо це, наприклад, матеріал по математиці або фізиці, то якість рисунків не важлива – важливо максимально охопити матеріал (різні додаткові кадри з доказами теорем тощо).

Розглянемо другу ситуацію. Можуть виникнути випадки, коли користувач хоче отримати інформацію за терміном, який не міститься ні в назві, ні в базі БКС. В цьому випадку розглядається запит на частину ЕП, який містить інформацію за якимось випадковим поняттям (терміном), тобто запит за випадковим ключовим словом (ВКС).

Обробка за цим типом запиту дуже схожа на обробку за БКС (нарощування базового кадру до необхідного об'єму V_3), проте є дві істотні відмінності:

– невідомий базовий кадр, який максимально описує поняття;
– невідомо звідки брати інформацію про пріоритети кадрів відносно базового кадру для нарощування.

При обробці ЕП за запитом за БКС уся ця інформація бралася з МБКС, яка формувалася на стадії створення початкового ЕП. Проте оскільки поняття випадкове, то таку таблицю заздалегідь сформувати неможливо.

Для вирішення першого завдання – визначення базового кадру, в роботі запропоновано використовувати метод релевантного індексованого пошуку учбових матеріалів [6], а саме, частина цього методу – релевантний пошук [7]. У більшості випадків основою релевантного пошуку є TI-IDF метод [8], який використовується у більшості пошукових систем. Його сенс зводиться до того, що чим більше локальна частота терміну (запиту) в документі і більше «рідкість» (тобто чим рідше він зустрічається в інших документах) терміну, тим вище вага цього документу по відношенню до терміну. Також використовується поняття пертинентності [7] тощо. З використанням методу релевантного пошуку визначається базовий кадр, який надалі називатимемо релевантний кадр.

Для вирішення другого завдання – визначення, які файли найбільш важливі для релевантного кадру (для нарощування його), використовуватимемо матрицю зв'язків гіперпосиланнями. Ця матриця схожа на попередню (за рис. 2) і відрізняється тим, що в таблиці ставлять не ваговий коефіцієнт, а кількість гіперзв'язків між файлами. Тут автори відштовхуються від того принципу, що якщо кадр має найбільшу кількість гіперпосилань на інший кадр, то він і найбільш важливий для розглядаємого, адже за гіперпосиланнями під'єднуються інші додаткові кадри, які пояснюють яку-небудь інформацію основного релевантного кадру.

Таким чином, метод формування ЕП при запиті за випадковим базовим словом схожий на попередній – при обробці ЕП при запиті за БКС, тільки базовий кадр замінюється релевантним кадром, а матриця базових ключових слів замінюється матрицею зв'язків гіперпосиланнями.

Висновки

Таким чином, в роботі описано два методи формування ЕП із заданим граничним розміром при запиті за ключовими словами: при запиті на частину ЕП, який містить інформацію за якимось стандартним (із заздалегідь запропонованого списку) поняттям (терміном), тобто при запиті за базовому ключовим словом; і при запиті на частину ЕП, який містить інформацію за якимось випадковим поняттям (терміном), тобто при запиті за випадковим ключовим словом. Обидва методи використовують частину алгоритму, який використовується при перетворенні усього об'єму ЕП під заданий розмір і відрізняються один від одного початковим кадром і типами матриць, які застосовують.

Описані методи можуть бути використані при формуванні електронних підручників з обмеженим розміром з інформацією про поняття, яке цікавить користувача. Це дозволяє їх використовувати в комп'ютеризованих системах з відносно невеликим об'ємом пам'яті: КПК, електронні книги тощо.

ЛІТЕРАТУРА

1. Афонін О. В. Навчальна книга для ВНЗ в контексті книговидавничої справи в Україні [Текст] // Праці І-го міжн. форуму «Інформаційне забезпечення навчального процесу у Вищій школі». – Одеса, 2007. – С. 25–29.
2. Башмаков А. И. Разработка компьютерных учебников и обучающих систем [Текст] / А. И. Башмаков, И. А. Башмаков. – М. : Информационно-издательский дом «Филинъ», 2003. – 616 с.
3. Башмаков И. А. Модель семантической сети для представления учебного материала в компьютерных обучающих средствах [Текст] / И. А. Башмаков, П. Д. Рабинович // Справочник. Инженерный журнал. – М. : Машиностроение, 2002. – № 8. – С. 61–64.
4. Мусиенко М. П. Чорномаз И. К. Разработка метода преобразования электронных учебников к заданным размерам для передачи сервисами GSM сети [Текст] / М. П. Мусієнко, В. І. Томенко // Вісник Черкаського державного технологічного університету. – 2010. – № 3. – С. 72–75.
5. Методы и средства разработки электронных изданий [Электронный ресурс] / Казанский Государственный Технический університет им. А. Н. Туполева. – Режим доступа : http://www.mi.ru/~dupliksv/pauk. – 16.07.2010 г. – Заглавие с экрана.
6. Уніфікований комплекс автоматизації процесу створення та накопичення лекційних матеріалів [Электронный ресурс] // Результати виконання Державної програми «Інформаційні та комунікаційні технології в освіті і науці» на 2006-2010 роки. – Режим доступа : http://www.mon.gov.ua/main.php?query=science/development. – 09.07.2010 г. – Заглавие с экрана.
7. Релевантность [Электронный ресурс] // Википедия. Свободная энциклопедия. – Режим доступа : http://ru.wikipedia.org/wiki/%D0%A0%D0%B5%D0%BB%D0%B5%D0%B2%D0%B0%D0%BD%D1%82%D0%BD%D0%BE%D1%81%D1%82%D1%8C. – 21.07.2010 г. – Заглавие с экрана.
9. TF-IDF [Электронный ресурс] // Википедия. Свободная энциклопедия. – Режим доступа : http://ru.wikipedia.org/wiki/TF-IDF. – 21.07.2010 г. – Заглавие с экрана.

Рецензенти: Кондратенко Ю. П., д.т.н., професор;
Казарєзов А. Я., д.т.н., професор

© Мусієнко М. П., Чорномаз І. К., 2010

Стаття надійшла до редколегії 20.12.2010 р.