

УДК 658.012:681.32:621.38

Л. І. Тимченко

(доктор технічних наук, професор, Державний університет інфраструктури та технологій, м. Київ)

Н. І. Кокряцька

(кандидат технічних наук, доцент, Державний університет інфраструктури та технологій, м. Київ)

І. Д. Івасюк

(кандидат технічних наук, доцент)

Ю. В. Майстренко

(аспірант, Державний університет інфраструктури та технологій, м. Київ)

М. М. Галушко

(аспірант, Державний університет інфраструктури та технологій, м. Київ)

РОЗРОБКА ЙМОВІРНІСНИХ МОДЕЛЕЙ ПАРАЛЕЛЬНОГО МЕТОДА УЩІЛЬНЕННЯ НА ОСНОВІ ПІРАМІДАЛЬНО-ЛІНІЙНОГО ТА ПІРАМІДАЛЬНО-НЕЛІНІЙНОГО КОДУВАННЯ

Запропоновано нові методи ущільнення інформації, а саме ймовірнісні моделі пірамідално-лінійного кодування та пірамідално-нелінійного кодування які дозволять значно покращити системи зберігання інформації. Задача змінити існуючі методи, а саме зменшити час ущільнення та зекономити місце для зберігання інформації на накопичувачі.

***Ключові слова:** пірамідално-лінійне кодування, пірамідално-нелінійне кодування, паралельний метод ущільнення, стиск, кодування, інформація, дані, кодер, декодер, модель.*

Вступ. За останній час об'єми цифрової інформації значно збільшилися. Це обумовлено постійним зростанням обміну інформації, оцифровуванням вже існуючої та стрімкою появою нової інформації в різних галузях формування та розвитку. В зв'язку з цим зростає необхідність зберігати інформацію в цілому. Для вирішення цього питання на даний час існує декілька можливостей та механізмів. По-перше, розробляються новітні технології накопичувачів, по-друге, еволюціонують методи ущільнення для зберігання інформації. Процес розвитку комп'ютерних технологій потребує постійної модернізації в цих напрямках один з

DOI: 10.32703/2617-9040-2019-33-2-1

яких ми розглянемо, а саме паралельний метод ущільнення (ПМУ) на основі пірамідально лінійного (ПЛК) та пірамідально нелінійного кодування.

В умовах стрімкого та динамічного розвитку інформаційних технологій неймовірно зростає кількість інформації, дослідницька тема паралельного методу ущільнення ПМУ (з точки зору забезпечення ефективності за допомогою ПЛК, ніж у відомих методах) є актуальною.

Аналіз останніх досліджень і постановка проблеми. Основною характеристикою алгоритму стиснення є коефіцієнт стиснення. Він визначається як відношення обсягу вихідних незжатих даних до обсягу стислих, тобто: $k = S_0 / S_c$, де k – коефіцієнт стиснення, S_0 – обсяг вихідних даних, а S_c – обсяг стиснутих [1]. Таким чином, чим вище коефіцієнт стиснення, тим алгоритм ефективніше.

Алгоритм RLE (Run Length Encoding, упаковка, кодування довжин серій), є найшвидшим, простим і зрозумілим алгоритмом стиснення даних, іноді виявляється досить ефективним [1]. Значний недолік в тому що від правильного вибору префікса залежить якість самого алгоритму стиснення, так як, якщо в початковому тексті часто зустрічаються поодинокі символи FF, розмір вихідної тексту може перевищити вхідний.

Найбільш ефективним кодом змінної довжини, в якій жодне слово не збігається з початком іншого (тобто префіксний код) є код Хаффмана [2-3]. Недоліком такого кодування є той факт, що разом з закодованим повідомленням необхідно передавати також і побудовану таблицю кодів (дерево), що знижує величину стиснення.

Метод «стопки книг» («MTF-Move To Front»). Кожному символу (букви) присвоюється код залежно від його положення в алфавіті – чим ближче символ до початку алфавіту, тим коротше код. Після кодування чергового символу він поміщається в початок алфавіту, зрушуючи всі інші літери на одну позицію вглиб. Через деякий час символи, що найбільш часто зустрічаються, групуються на початку, що і потрібно для успішного кодування [4-5]. Недоліком такого методу є порівняно невелика ступінь стиснення, низька швидкість і ємні витрати пам'яті.

Арифметичне кодування є методом, що дозволяє упаковувати символи вхідного алфавіту без втрат за умови, що відомий розподіл частот цих символів [6-8]. При розгляді цього методу виникають дві проблеми: по-перше, необхідна матеріальна арифметика, взагалі кажучи, необмеженої точності, і по-друге, результат кодування стає відомий лише при закінченні вхідного потоку.

Класифікаційний аналіз найбільш відомих методів ущільнення з урахуванням запропонованих, приведений на рис. 1.

Розглянувши всі методи стиску інформації, запропонований паралельний метод ущільнення ПМУ дозволяє значно збільшити ефективність ущільнення, а саме значно підвищити якість і знизити час кодування. Лінійні коди, в порівнянні з іншими кодами, дозволяють реалізовувати більш ефективні алгоритми кодування і декодування інформації. Це істотно спрощує реалізацію пристроїв кодування та декодування і робить лінійні коди вельми привабливими з точки зору практичних застосувань [9–11]. Ієрархічний метод класифікації – це метод, при якому задана безліч послідовно ділиться на підпорядковані підмножини, поступово конкретизуючи об'єкт класифікації. Сукупність одержаних угруповань при цьому утворює ієрархічну деревоподібну структуру у вигляді розгалуженого графа, вузлами якого є угруповання.

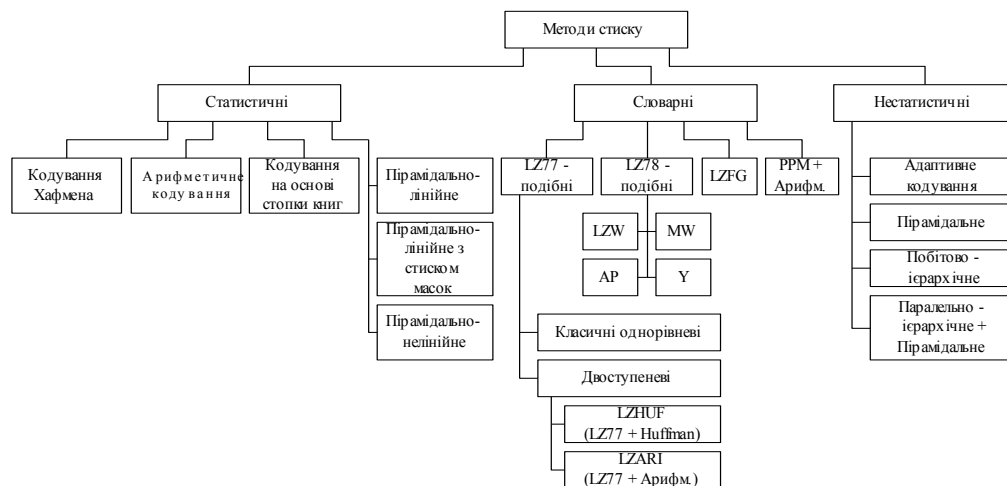


Рис. 1. Класифікаційний аналіз методів ущільнення

Вибір послідовності ознак залежить, перш за все, від характеру інформації. При побудові класифікації вибір послідовності ознак залежить від ймовірності звернення до тієї чи іншої ознаки. Основними перевагами ієрархічного методу є велика інформаційна ємність, традиційність і звичність застосування, можливість створення для об'єктів класифікації мнемонічних кодів, що несуть смислове навантаження [12–14].

Основні методи ущільнення інформації розглянуті в цьому дослідженні. Оцінка методів, представлених в цьому дослідженні, показала що, використання арифметичних, ймовірнісних, словникових, статистичних та нестатистичних методів виключають наступне:

- оцінка стану паралельного символу;
- отримання узагальненої оцінки стану кодування;
- безперервна оцінка декількох характеристик в реальному часі;
- одночасна оцінка як кодованих, так і декодованих символів;
- комбінований метод лінійного кодування при частому зверненні до помилки біта не використовуються в оцінці.

З цією метою вважається за доцільне оцінити переваги ПЛК для зменшення витрат часу і підвищення якості ущільнення.

Мета та задачі дослідження. Метою даного дослідження є розробка паралельного методу ущільнення даних за допомогою пірамідально-лінійного кодування для зменшення об'єму та підвищення якості і ефективності процесу зберігання та передачі інформації.

Для досягнення мети були поставлені такі завдання:

- проаналізувати наявні матеріали та дослідження в напрямку ущільнення інформації;
- розглянути системи побудови пірамідально-лінійних та пірамідально-нелінійних методів ущільнення;
- зробити аналіз отриманих результатів, зокрема місця та доцільність їх використання;
- охарактеризувати висновки з результатів дослідження та аналізу отриманих результатів.

Матеріали та методи дослідження. Дослідження проводились, починаючи з аналізу існуючих матеріалів методів пірамідально-лінійного та пірамідально-нелінійного кодування моделей паралельного метода ущільнення, вивчення принципу побудови якісних характеристик ефективності та місце їх використання [9–11]. На основі аналізу вище вказаних матеріалів, проводилась побудова ймовірнісних моделей паралельного метода ущільнення на основі пірамідально-лінійного та пірамідально-нелінійного кодування. Процес розробки ймовірнісних моделей був побудований на основі емпіричного методу досліджень. Тобто перед побудовою була висунута певна гіпотеза [12] та зроблені прогнозування результатів. Потім проводилась сама побудова моделей з експериментальними розрахунками. В цілому гіпотеза припущення звучала так: «Врахувати досвід та аналіз існуючих матеріалів методів пірамідально-лінійного та пірамідально-нелінійного кодування та моделей паралельного метода ущільнення можна припустити, що розробивши ймовірнісні моделі паралельного метода ущільнення на основі пірамідально-лінійного та пірамідально-нелінійного кодування можуть бути по своїх якісних характеристиках більш ефективними, ніж існуючі методи ущільнення даних, що застосовуються на сьогодні в комп'ютерній техніці, при правильному застосуванні».

Побудова моделей мала поетапний характер. А саме в декілька етапів, на кожному з яких відбувалась побудова певної частини моделі. Також проводили розрахунки та аналіз отриманих результатів на предмет вірності побудови та висунотій гіпотезі. При аналізі отриманих результатів побудови моделей та розрахунків, було враховано такі фактори, як складність побудови, швидкодія, якість ущільнення даних, кількість операцій за одиницю часу, місце можливого використання.

Розробка ймовірнісної моделі паралельного метода ущільнення на основі пірамідально-лінійного кодування. Застосували як критерій вибору статистичний критерій, тобто на кожному рівні паралельно-ієрархічного перетворення вибирається елемент із найбільшою ймовірністю, побудували ймовірнісну модель паралельного метода ущільнення на основі пірамідально-лінійного кодування. Як свідчать теоретичні дослідження, найбільше доцільне застосування статистичного критерію вибору для рішення задач ущільнення, зокрема для комп'ютерних технологій архівації даних.

Розглянемо більш докладно організацію пірамідально-лінійного кодування на функціональному рівні. Нехай в інформаційному потоку даних n елементів є k різноманітних, де p_i ймовірність появи кожного елемента. Причому через p_1 позначимо ймовірність елемента, що найбільш часто зустрічається, тоді

$$p_1 \geq p_2 \geq \dots \geq p_i \geq \dots \geq p_k. \quad (1)$$

Тобто, розмістимо ймовірності появи елементів у порядку не зростання. Пірамідально-лінійний засіб кодування полягає в наступному. Обраний на кожному кроці ПІ перетворення елемент, кодується одиницею, а інші - нулем. Тому, якщо в початковому потоку є n елементів, то й у першому стовпчику масок теж буде n кодів-масок, із яких $n p_1$ – одиниці. Тоді в другому стовпчику масок уже буде $n - n p_1$ кодів-масок, із яких $n p_1$ – одиниці. Таким чином, утворюється $k-1$ стовпчик масок, де кількість елементів у кожному стовпчику буде таким:

Таблиця 1. Оцінка числа елементів стовпчиків-масок для ПЛК

1 стовпчик	2 стовпчик	3 стовпчик	k-1 стовпчик
n	n-np ₁	n-np ₁ -np ₂	$n\left(1-\sum_{i=1}^{k-2} p_i\right)$

Причому, в останньому стовпчику np_k - нулів, інші одиниці (тобто np_{k-1}). Отже, елемент із найменшою частотою появи кодується k-1 кількістю нулів.

Знайдемо кількість елементів у масках:

$$\begin{aligned}
 S &= n + n - np_1 + n - np_1 - np_2 + n - np_1 - np_2 - np_3 + L + n\left(1 - \sum_{i=1}^{k-2} p_i\right) = \\
 &= n + n(1 - p_1) + n(1 - (p_1 + p_2)) + n(1 - (p_1 + p_2 + p_3)) + \dots + n\left(1 - \sum_{i=1}^{k-2} p_i\right) = \\
 &= n(1 + 1 - p_1 + 1 - p_1 - p_2 + 1 - p_1 - p_2 - p_3 + \dots + \\
 &\quad + 1 - p_1 - p_2 - p_3 - p_4 - \dots - p_{k-2}) = \\
 &= n((k - 1) - (k - 2)p_1 - (k - 3)p_2 - (k - 4)p_3 - \dots - 2p_{k-3} - p_{k-2}) = \quad (2) \\
 &= n((k - 1 - k + 2)p_1 + (k - 1 - k + 3)p_2 + (k - 1 - k + 4)p_3 + \\
 &\quad + \dots + (k - 1 - 2)p_{k-3} + (k - 1 - 1)p_{k-2} + (k - 1)p_{k-1} + (k - 1)p_k) = \\
 &= n(p_1 + 2p_2 + 3p_3 + \dots + (k - 3)p_{k-3} + (k - 2)p_{k-2} + (k - 1)p_{k-1} + (k - 1)p_k) = \\
 &= n\left(\sum_{i=1}^{k-1} ip_i + (k - 1)p_k\right).
 \end{aligned}$$

Якщо через A_n^k позначити вхідний потік із n символів по k різноманітних, M стовпчик-маска, а A_{p_i} – елемент із ймовірністю $p_i, i = \overline{1, k}$, то в загальному вигляді схему пірамідально-лінійного засобу кодування можна записати в такому вигляді:

$$A_n^k = A_{p_1} \left(M_n^1 \right) + A_{p_2} \left(M_{n(1-p_1)}^2 \right) + \dots + A_{p_{k-1}} \left(M_{n\left(\sum_{i=1}^{k-2} p_i\right)}^{k-1} \right) + A_{p_k} \left(M_{n\left(1-\sum_{i=1}^{k-2} p_i\right)}^{k-1} \right). \quad (3)$$

Процес пірамідально-лінійного кодування (ПЛК) можна прискорити в 2 рази, якщо на кожному рівні аналізувати пари сусідніх символів із ймовірностями p_i і p_{i+1} . Тоді процес ПЛК стає паралельним і його можна представити у вигляді кодового дерева (рис. 2) і записати схему кодування в такому вигляді:

$$A_n^k = \left(A_{p_1} M_n^1 + A_{p_2} M_{n(1-p_1)}^1 \right) + \left(A_{p_3} M_{n(1-p_1-p_2)}^2 + A_{p_4} M_{n(1-\sum_{i=1}^3 p_i)}^2 \right) + \dots + \left(A_{p_{k-1}} M_{n(1-\sum_{i=1}^{k-2} p_i)}^{k/2} + A_{p_k} M_{n(1-\sum_{i=1}^{k-1} p_i)}^{k/2} \right). \quad (4)$$

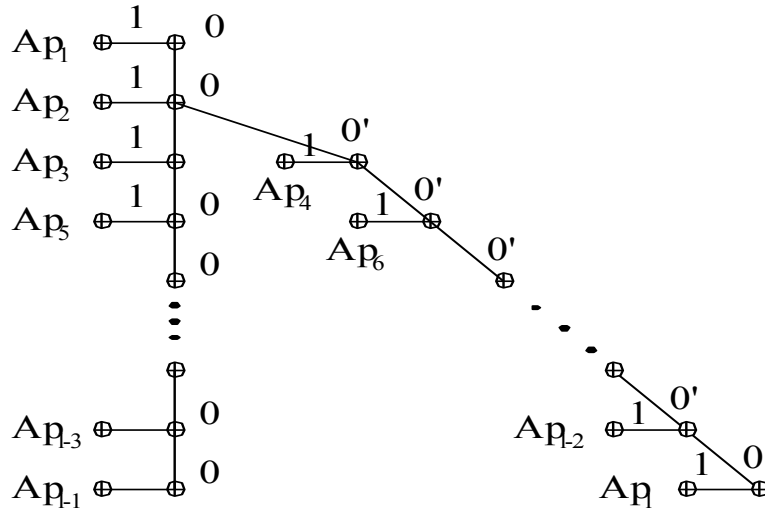


Рис. 2. Граф-схема кодового дерева ПЛК

Аналіз виразу (3) і кодового дерева (рис. 2) показує, що в порівнянні з відомими засобами статистичного кодування за рахунок паралельної організації схеми ПЛК можна домогтися істотного зменшення часу ущільнення. Так, для першої схеми ПЛК цей час визначається шляхом аналізу всіх символів вхідного алфавіту, а для другої схеми ПЛК-половиною. У той же час відомі статистичні засоби, орієнтовані на послідовну вибірку елементів відповідно до ймовірностей появи в потоці даних. У цьому випадку час ущільнення визначається довжиною вхідного потоку даних, а не довжиною алфавіту.

Тобто враховане зауваження, що стосується останнього елемента з найменшою ймовірністю p_k .

Справедливий такий вираз:

$$S = n \left(\sum_{i=1}^{k-1} i p_i + (k-1) p_k \right). \quad (5)$$

Проаналізуємо ефективність ущільнення при реалізації ПЛК. Порівняємо кількість кодів-масок у масках-стовпчиках із початковим потоком інформації, тобто S і $n \log_2 k$. Тоді коефіцієнт ущільнення R :

$$R = \frac{\log_2 k}{\left(\sum_{i=1}^{k-1} i p_i + (k-1) p_k \right)}. \quad (6)$$

Оцінимо, при якій умові $R > 1$, тобто умова, при якій даний метод ущільнення

На першому рівні елементів, сума ймовірностей котрих ≈ 0.5 , сортуються у порядку убавання ймовірності. Ці елементи кодують одиницею, всі інші - нулем. Таким чином, множина всіх елементів розбивається на дві підмножини. На наступному рівні роздільно для кожної підмножини процедура цілком повторюється, тобто процес відбувається рекурсивно, поки не з'явиться підмножина, що складається тільки з одного елемента.

Нехай на першому кроці сума ймовірностей перших l (рахуючи від елемента з найбільшою ймовірністю) буде біля 0.5. Очевидно, що $l = 1, 2, \dots, \frac{k}{2}$; причому $l = 1$

при $p_1 \approx 0.5$ і $l = \frac{k}{2}$, при $p_1 = p_2 = \dots = p_k$.

Тому, якщо в початковому потоці є n елементів, то в першому стовпці-масок теж буде n кодів-масок, із яких $n(p_1, p_2, \dots, p_l)$ – одиниць і $n(1 - \sum_{i=1}^l p_i)$ – нулів.

На другому кроці формується дві підмножини елементів:

I-а з $n \sum_{i=1}^l p_i$ – елементів, II-а з $n \sum_{i=l+1}^k p_i$ – елементів.

Для кожної з цих підмножин, якщо $l \neq 1$, процедура опису для першого кроку кодування цілком повторюється і т.д., доти поки не будуть отримані підмножини, що полягають з одного елемента.

Позначимо через l_2^1 кількість елементів на другому рівні, ймовірності котрих (рахуючи від елемента з найбільшою ймовірністю) у підмножині елементів, закодованих на попередньому рівні нулем, будуть біля 0.5. На цьому рівні ці елементи кодують одиницею. Позначимо через l_2^2 кількість елементів на другому рівні, ймовірності яких (рахуючи від елемента з найбільшою ймовірністю) у підмножині елементів, закодованих на попередньому рівні одиницею будуть біля 0.5.

Тоді для $l_2^1 = 0, 1, 2, \dots, \frac{l}{2}$, причому $l_2^1 = 0$ при $l = 1$, $l_2^1 = 1$ при $p_1 \approx 0.25$ і $l_2^1 = \frac{l}{2}$ при $p_1 = p_2 = \dots = p_l$. Аналогічно, для $l_2^2 = l+1, l+2, \dots, \frac{(k+1)}{2}$, причому $l_2^2 = 0$ при $k = 1$, $l_2^2 = 1$ при $p_{l+1} \approx 0.25$ і $l_2^2 = \frac{(k+1)}{2}$ при $p_{l+1} = p_{l+2} = \dots = p_k$. Тому в другому стовпчику кількість одиниць:

$$n \left(1 - n \sum_{i=l+1}^k p_i \right) \sum_{i=1}^{l_2^1} p_i = n \sum_{i=l+1}^k p_i \sum_{i=1}^{l_2^2} p_i. \quad (9)$$

Співвідношення (9) справедливо для підмножини елементів, закодованих на попередньому рівні нулем.

$$n \sum_{i=1}^l p_i (p_1 + p_2 + \dots + p_{l_2^1}) = n \sum_{i=1}^l p_i \sum_{i=1}^{l_2^1} p_i. \quad (10)$$

Тоді співвідношення (10) справедливо для підмножини елементів, закодованих на попередньому рівні одиницею.

Таким чином, на другому рівні отримано чотири підмножини елементів:

$$\text{I-a: } n \sum_{i=1}^l p_i \sum_{i=1}^{l_2^1} p_i ; \text{ II-a: } n \sum_{i=1}^l p_i \sum_{i=l_2^1+1}^l p_i ; \text{ III-a: } n \sum_{i=l+1}^k p_i \sum_{i=l+1}^{l_2^2} p_i ; \text{ IV-a: } n \sum_{i=l+1}^k p_i \sum_{i=l_2^2+1}^k p_i .$$

У загальному випадку число l_x^y – число елементів із ймовірністю p_i для кожного рівня, де $x = 0$ номер рівня, $y = \overline{1, \dots, 2^{x-1}}$.

Причому, якщо $y = 0$ – непарне, то це означає, що на попередньому рівні дані елементи були закодовані одиницею, якщо парне 0 то на попередньому рівні дані елементи були закодовані нулем. Тоді для третього стовпчика утворяться 8 підмножин елементів.

$$\text{I-a: } n \sum_{i=1}^l p_i \sum_{i=1}^{l_2^1} p_i \sum_{i=1}^{l_3^1} p_i , \text{ II-a: } n \sum_{i=1}^l p_i \sum_{i=1}^{l_2^1} p_i \sum_{i=l_3^1+1}^{l_3^2} p_i , \text{ III-a: } n \sum_{i=1}^l p_i \sum_{i=l_2^1+1}^l p_i \sum_{i=l_3^2+1}^{l_3^3} p_i ,$$

$$\text{IV-a: } n \sum_{i=1}^l p_i \sum_{i=l_2^1+1}^l p_i \sum_{i=l_3^2+1}^l p_i , \text{ V-a: } n \sum_{i=l+1}^k p_i \sum_{i=l+1}^{l_2^2} p_i \sum_{i=l+1}^{l_3^3} p_i , \text{ VI-a: } n \sum_{i=l+1}^k p_i \sum_{i=l+1}^{l_2^2} p_i \sum_{i=l_3^3+1}^{l_3^4} p_i ,$$

$$\text{VII-a: } n \sum_{i=l+1}^k p_i \sum_{i=l_2^2+1}^k p_i \sum_{i=l_3^4+1}^{l_3^5} p_i , \text{ VIII-a: } n \sum_{i=l+1}^k p_i \sum_{i=l_2^2+1}^k p_i \sum_{i=l_3^5+1}^k p_i ,$$

$$\text{де } l_3^1 = 0, 1, \dots, \frac{l_2^1}{2} ; l_3^2 = l_2^1 + 1, l_2^1 + 2, \dots, \frac{(l + l_2^1)}{2} ; l_3^3 = l + 1, l + 2, \dots, \frac{l_2^2}{2} ;$$

$$l_3^4 = l_2^2 + 1, l_2^2 + 2, \dots, \frac{(k + l_2^2)}{2} .$$

Тоді для $p_1 \geq p_2 \geq \dots \geq p_k$ можна записати кількісну оцінку числа елементів стовпчиків-масок, що відображена в табл. 2.

$$M_1 = \left(\begin{array}{c} n \sum_{i=1}^l p_i \\ n \left(1 - \sum_{i=1}^l p_i \right) \end{array} \right) \quad (11)$$

$$M_2 = n \left(\begin{array}{c} \sum_{i=1}^l p_i \\ \sum_{i=1}^l p_i \\ \sum_{i=l+1}^k p_i \\ \sum_{i=l+1}^k p_i \end{array} \right)^T \left(\begin{array}{c} \sum_{i=1}^{l_2^1} p_i \\ \sum_{i=l_2^1+1}^l p_i \\ \sum_{i=l+1}^{l_2^2} p_i \\ \sum_{i=l_2^2+1}^k p_i \end{array} \right) \quad (12)$$

Аналогічно в матричній формі виглядає M_3 і нескладно записати вираз для M_j .

Причому, S і R приймають граничні значення:

- 1) якщо p_1, p_2, \dots, p_k – утворять ряд спадної геометричної прогресії;
- 2) якщо $p_1 = p_2 = \dots = p_k$.

А модель кодування в матричній формі тоді має вигляд:

$$A_n^k = \sum_{j=1}^k \sum_{i=1}^k A_{p_i} M_{p_i}^{n_j}, \quad (17)$$

де n_j – число елементів на j -му рівні, $n_i = n$.

M_{p_i} – матриця-стовпчик масок на j -му рівні елемента з ймовірністю p_i .

Висновки. Виконавши розробку ймовірнісної моделі паралельного метода ущільнення на основі пірамідально-лінійного кодування, якісну умову ущільнення можна сформулювати таким чином: «двійковий логарифм від кількості різноманітних елементів і початкового потоку даних повинний бути більше суми добутоків ймовірності елемента на його порядковий номер вибірки (вибірка здійснюється за не зростанням ймовірностей)».

Аналіз виразу (7) показує, що ефективність засобу ПЛК з ущільнення даних порівняна із сімейством методів ущільнення на основі кодування Хаффмена.

На відміну від методів кодування Хаффмена, де необхідна організація досить складної процедури побудови кодового дерева, для ПЛК його будувати немає необхідності, хоча формальна побудова його можлива.

Зазначена перевага особливо цінна при використанні адаптивного кодування, для якого необхідно постійне коректування дерева відповідно до статистики вхідного потоку, що змінюється. При реалізації це потребує значних витрат на перебалансування кодового дерева відповідно до нових частот символів на кожному кроці.

Оскільки реалізація ПЛК не передбачає побудови кодового дерева, то при використанні адаптивного кодування не потрібно додаткових витрат на перебалансування кодового дерева, що істотно покращує якість ущільнення.

Також, відзначені переваги пірамідально-лінійного кодування повною мірою відносяться до загального випадку – пірамідально-нелінійного кодування.

Треба відзначити, що метод ущільнення на основі пірамідального кодування має схожість з методом швидкого сортування, запропонованого в роботі. Причому, розроблений метод має асимптотично більш якісні характеристики, ніж усі відомі методи ущільнення даних. Даний метод при використанні оптимальних значень кількості етапів і оптимальних значень p_i може бути застосований до комбінаторних джерел. Якщо робити перетворення блоків ймовірнісних джерел в адаптивні блоки з комбінаторними властивостями, то даний метод також може бути використаний. Але таке–перетворення може бути практично дійсним лише для Марковських джерел низького порядку. Тому цей метод найбільш доцільно використовувати для таких даних, як наприклад, відеосигнали природнього походження, а також текстів, які можна представити як Марковські джерела низького порядку і текстів із порівняно невеликим словником.

Проаналізувавши існуючі методи ущільнення, можна точно стверджувати, що модель ПЛК значно краще за всіма показниками, ніж класичні. Це обумовлено

такими пріоритетами, як скорочення часу кодування до 20 мс на 1 кБ, покращення якості та зменшення об'єму інформації.

Розроблений МПУ має два ПЛК, кожен з яких дає достатньо високі показники щодо ущільнення, але друга модель має переваги стосовно часу ущільнення та якості збереження інформації.

Експериментальне дослідження розробки ймовірнісної моделі паралельного ущільнення було проведено на основі пірамідально-лінійного кодування. Результати якого повністю підтверджують гіпотезу та показують доцільність розробки даної моделі. Порівняння показує, що ПЛК значно швидший і не потребує додаткового навантаження на апаратні ресурси.

Властивості класичних методів ущільнення та пірамідально-лінійного кодування були досліджені методом математичної індукції, порівнянням кількості кодів масок із початковим потоком інформації та методами паралельно-ієрархічних перетворень. Що підвищує якість дослідження в багато разів.

Результати, отримані в роботі, можуть бути безпосередньо використані для таких даних, як сигнали природного походження, для текстів із порівняно невеликим словником. А також для текстів які можна представити як Марковські джерела низького порядку. Ці методи демонструють потенціал збільшення об'єму збереженої інформації не модифікуючи апаратні ресурси та дозволяють скоротити час обробки в 10 разів, а також покращують якість ущільнення приблизно в 4 рази.

ЛІТЕРАТУРА

1. Владимир Цымбал. Теория информации и кодирование. Задачник. Учебное пособие. М.: Ленанд. 2014. 280 с.
2. Симаков А. Код Хаффмана. URL:http://compression.ru/download/articles/huff/simakov_2002_huffcode.html (2002)
3. Алгоритм Хаффмана на пальцах (2012). URL: <https://habr.com/ru/post/144200>
4. Бакнелл Д. М. Фундаментальные алгоритмы и структуры данных в Delphi. Санкт-Петербург. 2006. С. 449 – 469.
5. Juan Francisco, Rodríguez Herrera. URL: https://w3.ual.es/~vruiz/Docencia/Apuntes/Coding/Text/03-symbol_encoding/06-move-to-front_transform/index.html (2016)
6. Moffat A., Turpin A. Compression and Coding Algorithms. University of Melbourne. 2002. С. 145–154.
7. Арифметическое кодирование. URL: <https://habr.com/ru/post/130531> (2011)
8. Б. Я. Рябко, А. Н. Фионов. Эффективный метод адаптивного арифметического кодирования для источников с большими алфавитами. Пробл. передачи информ. 1999. С. 34 – 39.
9. Чобану. М. Многомерные, многоскоростные системы обработки сигналов. Техносфера. Москва. 2009. С. 345 – 356.
10. Д. Ватолин, А. Ратушняк, М. Смирнов, В. Юкин. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео. Диалог-МИФИ. Москва. 2003. С. 94 – 291.
11. Буза М. К. Системы параллельного действия. пособие для студентов спец. «Информатика» Минск. БГУ. 2009. С. 349 – 367.
12. D. Salomon, G. Motta. Data Compression Guide. London. 2010. 1361 с.
13. Тимченко Л.І., Кокряцька Н.І., Герцій О.А., Петровський М.С., Степанюк Д.С. Паралельно-ієрархічні мережі для оброблення біометричних зображень та зображень плям лазерних пучків. Монографія. Полтава. АСМІ. 2017. 363 с.
14. Тимченко Л.І., Кокряцька Н.І., Герцій О.А., Петровський М.С., Степанюк Д.С. Паралельно-ієрархічні мережі для оброблення зображень. Теоретичні дослідження : монографія. Полтава: АСМІ, 2017. 469 с.
15. Пуанкаре А. Наука и гипотеза. Пер. с франц. Либроком. Москва 2015. С. 116 – 130.

REFERENCES

1. Vladimir Tsybmal (2014). Theory of information and coding. Problem. Tutorial. Lendand – 280. [in Russian].
2. Simakov A (2002). Code of Huffman. URL: http://compression.ru/download/articles/huff/simakov_2002_huffcode.html
3. Huffman's algorithm on fingers (2012). URL: <https://habr.com/ru/post/144200>
4. Bucknell D. M. (2006). Fundamental algorithms and data structures in Delphi. St. Petersburg. 449-469. [in Russian].
5. Juan Francisco Rodríguez Herrera (2016). URL: https://w3.ual.es/~vruiz/Docencia/Apuntes/Coding/Text/03-symbol_encoding/06-move-to-front_transform/index.html
6. Moffat A., Turpin A. (2002). Compression and Coding Algorithms. University of Melbourne. 145 – 154.
7. Arithmetic coding (2011). URL: <https://habr.com/ru/post/130531>
8. B. Ya. Ryabko, A. N. Fionov. (1999). An effective method of adaptive arithmetic coding for sources with large alphabets. Probl. information transfer. 34-39. [in Russian].
9. Choban. M. (2009). Multidimensional, multi-speed signal processing systems. Technosphere. Moscow. 345-356. [in Russian].
10. D. Vatolin, A. Ratushniak, M. Smirnov, V. Yukin. Data compression methods. Arrangement of archivers, compression of images and video. Dialogue-MIFI. Moscow. (2003). 94-291. [in Russian].
11. Buz M.K. Systems of parallel action. allowance for students special. "Informatics" Minsk. BSU (2009). 349-367. [in Belarus].
12. D. Salomon, G. Motta. (2010). Data Compression Guide. London. 1361.
13. Timchenko L.I., Kokryatskaya N.I., Gertsio O.A., Petrovsky M.S., Stepanyuk D.S. (2017). Parallel-hierarchical networks for processing biometric images and images of stains of laser beams. Monograph. Poltava ASMI 363. [in Ukrainian].
14. Timchenko L.I., Kokryatskaya N.I., Garcia O.A., Petrovsky M.S., Stepanyuk D.S. (2017) Parallel-hierarchical networks for image processing. Theoretical research: monograph / L.I. Timchenko, N.I. Kokryatska, O.A. Gercius, MS Petrovsky, D.S. Stepanyuk - Poltava: ASMI. – 469. [in Ukraine].
15. Poincaré A. (2015). Science and hypothesis. Per. from French Librokom. Moscow. with. 116-130. [in Russian].

Л. И. Тимченко

(доктор технических наук, профессор, ГУИТ, г. Киев)

Н. И. Кокряцкая

(кандидат технических наук, доцент, ГУИТ, г. Киев)

И. Д. Ивасюк

(кандидат технических наук, доцент)

Ю. В. Майстренко (аспирант, ГУИТ, г. Киев)

М. М. Галушко (аспирант, ГУИТ, г. Киев)

РАЗРАБОТКА ВЕРОЯТНОСТНЫХ МОДЕЛЕЙ ПАРАЛЛЕЛЬНОГО МЕТОДА УПЛОТНЕНИЯ НА ОСНОВЕ ПИРАМИДАЛЬНО-ЛИНЕЙНОГО И ПИРАМИДАЛЬНО-НЕЛИНЕЙНОГО КОДИРОВАНИЯ

Предложены новые методы уплотнения информации, а именно вероятностные модели пирамидально-линейного кодирования и пирамидально-нелинейного кодирования которые позволят значительно улучшить системы хранения информации. Задача изменить существующие методы, а именно уменьшить время уплотнения и сэкономить место для хранения информации на накопителе.

Ключевые слова: *пирамидально-линейное кодирование, пирамидально-нелинейное кодирование, параллельный метод уплотнения, сжатие, кодирование, информация, данные, кодер, декодер, модель.*

L. Tymchenko

(doctor of technical sciences, professor, SUIT, Kyiv)

N. Kokryatska

(candidate of technical sciences, associate professor, SUIT, Kyiv)

I. Ivasyuk

(candidate of technical sciences, associate professor)

Y. Maystrenko

(postgraduate student, SUIT, Kyiv)

M. Halushko

(postgraduate student, SUIT, Kyiv)

DEVELOPMENT OF PROBABILITY MODELS OF THE PARALLEL SEALING METHOD BASED ON PYRAMIDAL LINEAR AND PYRAMIDAL NONLINEAR CODING

New methods of information compression have been proposed, namely, probabilistic models of pyramid-linear coding and pyramid-non-linear coding which will significantly improve information storage systems. The task is to change the existing methods, namely, reduce compaction time and save space for storing information on the drive.

In the first model of pyramid-linear coding, the time is determined by analyzing all the characters of the input alphabet, and in the second half. In both cases, a significant reduction in compaction time is observed, but in the second case it is 20 ms per 1 kB of information. Also, pyramid-linear coding does not provide for the construction of a code tree, which significantly improves the quality of compaction.

The results of the study demonstrate the positive nature of the development of models of pyramidal linear coding and pyramidal non-linear coding, which will make it possible to advance in the field of information technologies and significantly improve information storage and transmission systems.

Research in the future will provide significant savings in resources on the use of data storage capacity, as well as solve the issue of time and quality regarding data compression. It should be noted that pyramid-linear coding is the basis for the implementation of completely new and high-quality information processing systems. The development of these systems will fundamentally change the industry and will give a significant impetus to evolution.

Keywords: *pyramid-linear coding, pyramid-non-linear coding, parallel compaction method, compression, coding, information, data, encoder, decoder, model.*