Министерство образования и науки Украины
Харьковский национальный университет радиоэлектроники

НАУКОЕМКИЕ ТЕХНОЛОГИИ В ИНФОКОММУНИКАЦИЯХ:

ОБРАБОТКА И ЗАЩИТА ИНФОРМАЦИИ

Коллективная монография

Харьков
Компания СМИТ
2013

УДК 621.395
ББК 32.973.201-018.2
Н 34

*Печатается по решению Ученого совета
Харьковского национального университета радиоэлектроники
(протокол № 23 от 05.07.2013 г.)*

Рецензенты:

В. И. Хаханов — д-р техн. наук, проф., декан факультета компьютерной инженерии Харьковского национального университета радиоэлектроники;

О. Е. Федорович — д-р техн. наук, проф., зав. кафедрой информационно-управляющих систем Национального аэрокосмического университета «ХАИ» им. М. Е. Жуковского;

О. К. Юдин — д-р техн. наук, проф., зав. кафедрой комплексных систем защиты информации Национального авиационного университета

Н 34 **Наукоемкие** технологии в инфокоммуникациях: обработка и защита информации : коллективная монография / под редакцией В. М. Безрука, В. В. Баранника. — Х. : Компания СМИТ, 2013. — 398 с.

ISBN 978-617-621-021-4

Коллективная монография содержит материалы по актуальным направлениям инфокоммуникационных технологий. Излагаются направления развития таких важных областей как планирование инфокоммуникационных сетей, повышение качества предоставляемых услуг, обработка и защита информации, в том числе, вопросы обеспечения эффективного хранения, обработки, сжатия и безопасности информации, управления инфокоммуникационными сетями с использованием методов прогнозирования временных рядов, принятия оптимальных решений, распознавания образов, распределенной обработки информации и облачных вычислений.

Материалы книги представляют теоретический и практический интерес и могут быть полезными для специалистов, аспирантов, адъюнктов, магистров и студентов, проводящих исследования в области обработки и защиты информации при создании инфокоммуникационных сетей.

УДК 621.395
ББК 32.973.201-018.2

© Коллектив авторов, 2013
© Редакция В.М. Безрука,
В.В. Баранника, 2013

ISBN 978-617-621-021-4

© ООО «Компания СМИТ», 2013

СОДЕРЖАНИЕ

Часть 1. ОБЩИЕ ВОПРОСЫ ИНФОКОММУНИКАЦИЙ

1. Инфокоммуникации. *Гайворонская Г.С.* 7
2. Инфокоммуникации – новая парадигма связи. *Никитюк Л.А.* 26
3. Облачные вычисления в инфокоммуникациях. *Глоба Л.С.* 37
4. Применение интеллектуальных технологий в инфокоммуникационных сетях. *Толюпа С.В.* 54
5. Аспекты исследования слабоструктурированных сред инфокоммуникаций. *Соломицкий М.Ю.* 68
6. Информационные сети и услуги. *Подорожняк А.А.* 86
7. Принятие оптимальных решений в инфокоммуникациях с учетом совокупности показателей качества. *Безрук В.М., Буханько А.Н., Чеботарёва Д.В.* 104
8. Модели оптимального размещения файлов в информационно-телекоммуникационных сетях. *Глоба Л.С.* 126
9. Методы анализа и оптимизации для решения телекоммуникационных задач. *Глоба Л.С., Лысенко Д.С., Кобзарь Л.С.* 145
10. Беспроводные сенсорные сети как часть инфокоммуникационной структуры. *Зеленин А.Н., Власова В.А.* 184

Часть 2. ОБРАБОТКА ИНФОРМАЦИИ В ИНФОКОММУНИКАЦИОННЫХ СЕТЯХ

11. Нейросетевые методы сжатия изображений. *Руденко О.Г., Бессонов А.А., Сныткин М.С., Мирошниченко С.В., Руденко С.О.* 195
12. Сжатие НН квадратуры двумерного дискретного вейвлет-преобразования в двухуровневом полиадическом пространстве. *Баранник В.В., Ширяев А.В., Харченко Н.А.* 214
13. Одномерное плавающее структурное кодирование двоичных матриц. *Баранник В.В., Хаханова А.В., Сидченко С.А.* 226
14. Методи адаптивного ущільнення даних на основі лінійної форми Фібоначчі. *Лужецький В.А., Савицька Л.А.* 241
15. Прогнозирование временных рядов со степенными регрессорами при коррелированном шуме наблюдения. *Омельченко А.В., Роздымаха Е.А., Федоров А.В.* 260
16. Нелинейное оценивание параметров полиномиальных трендов при негауссовской стохастической компоненте. *Селин В.О., Заболотный С.В.* 271
17. Распознавание заданных образов на основе информации в виде случайных сигналов. *Безрук В.М.* 281

18. Разработка алгоритма оценки вероятности риска развития ИБС для региональной системы «Телемедицина». *Анохин Д.А., Никитин В.М., Довгаль В.М., Липунова Е.А., Кочеткова И.А.* 290
19. Применение линейного предсказания речи для идентификации абонентов инфокоммуникационных сетей по особенностям их голоса. *Федоров А.В., Омельченко А.В.* 299

Часть 3. **ЗАЩИТА ИНФОРМАЦИИ В ИНФОКОММУНИКАЦИОННЫХ СЕТЯХ**

20. Перспективные технологии защиты инфокоммуникационных сетей. *Толюпа С.В.* 323
21. Анализ кибератак на информационно-телекоммуникационные системы. *Ларин В.В., Рябуха Ю.Н., Тарнополов Р.В.* 336
22. Применение операций матричного криптографического преобразования для построения систем защиты информации повышенной криптостойкости. *Рудницкий В.М., Бабенко В.Г., Рудницкий С.В.* 344
23. Методов повышения помехоустойчивости систем скрытой передачи информации. *Астраханцев А.А., Дорожан А.В.* 364
24. Повышение энергетической и спектральной эффективности радиоканала сети LTE. *Стрихалюк Б.М., Яремко О.Н., Максимюк Т.А* 375
25. Оцінка ефективності криптоаналізу симетричних примітивів із використанням fuzzy-rainbow-таблиць. *Олійников Р.В.* 385

ПРЕДИСЛОВИЕ

Внедрение информационно-коммуникационных технологий в производство, управление, образование, медицину, финансы, банки, торговлю и быт людей в настоящее время свидетельствует о переводе всех видов деятельности на качественно новый индустриальный уровень, который можно характеризовать появлением отрасли инфокоммуникации. Понятие «инфокоммуникации» является доминирующим при обсуждении современного уровня научно-технического прогресса общества. Важную роль он играет и при определении направлений подготовки нового поколения специалистов в области информационных и телекоммуникационных технологий. Инфокоммуникации – это совокупность методов и средств накопления, хранения, обработки, защиты и передачи информации в пространстве. Развитие инфокоммуникаций является небходимым условием создания информационного обмена и построения информационной структуры общества. Это дает толчок развитию наукоемких технологий в инфокоммуникациях. Наукоемкость технологий обусловлена использованием сложного математического аппарата и высокоинтегрированных вычислительных средств при решении задач обработки и защиты информации в инфокоммуникациях. Отражению вопросов разработки и внедрения в инфокоммуникации новых наукоемких технологий и посвящена данная коллективная монография.

В коллективную монографию вошли материалы ведущих специалистов из Киева, Харькова, Одессы, Львова, Черкасс, Винницы, Белгорода, развивающих наукоемкие технологии обработки и защиты информации в области инфокоммуникаций. В книге нашли отражение как общие вопросы развития инфокоммуникаций, так и частные вопросы хранения, обработки и защиты информации в инфокоммуникационных сетях. Материал монографии условно разделен на три части, содержащих самостоятельные разделы отдельных авторов. Первая часть книги включает разделы по общим вопросам инфокоммуникаций, предоставления информационных услуг, принятия оптимальных решений в задачах инфокоммуникаций, распределенной обработки информации и облачных вычислений. Вторая часть включает разделы по разным методам обработки информации, в частности, прогнозирования временных рядов и распознавания образов при обработке информации, сжатия данных и видеоизображений. Третья часть книги содержит разделы по обеспечению безопасности информационных ресурсов в инфокоммуникациях.

Следует отметить, что из-за ограниченности объема в монографии не затронуты очень важные вопросы передачи информации в инфокоммуникационных сетях.

Научные редакторы В.М. Безрук, д-р. техн. наук, проф.
В.В. Баранник, д-р техн. наук, проф.

14

МЕТОДИ АДАПТИВНОГО УЩІЛЬНЕННЯ ДАНИХ НА ОСНОВІ ЛІНІЙНОЇ ФОРМИ ФІБОНАЧЧІ

В. А. Лужецький, Л. А. Савицька
Вінницький національний технічний університет
lva_zi@mail.ru

Теоретичні дослідження і практика застосування архіваторів показали, що не існує універсального методу ущільнення, що забезпечував би однаковий степінь ущільнення для різних типів даних. Тому наукові дослідження спрямовані на створення ефективних методів ущільнення певних типів даних. Однак дані навіть одного типу, з погляду ущільнення, мають різні властивості і характеристики. З огляду на це, останнім часом прагнуть до створення адаптивних алгоритмів ущільнення даних [1,2].

Одним з найважливіших положень теорії ущільнення інформації є висловлена в [3] ідея поділу процесу ущільнення на дві процедури: моделювання і кодування. Моделювання визначає характеристики джерела даних, що ущільнюються, а кодування перетворює символи у послідовність бітів відповідно до отриманих характеристик. Незважаючи на даний поділ, множина існуючих методів ущільнення настільки різноманітна, що не існує єдиної теоретичної основи, яка дозволяла б описати їх усіх. Для усунення цього недоліку пропонується узагальнена модель адаптивного ущільнення даних, яка забезпечить опис з єдиних позицій існуючих методів ущільнення і розробку нових методів.

Існуючі методи й алгоритми ущільнення даних, в основному, враховують статистику символів у повідомленні [4–6] або базуються на побудові словника [7,8]. І хоча на практиці широко використовуються архіватори, створені на основі саме цих методів, пошуки нових підходів до ущільнення даних продовжуються.

Одним із цікавих підходів є пропозиція використовувати для ущільнення даних оптимізуючі властивості чисел Фібоначчі [9,10]. Суть підходу полягає в тому, що в процесі ущільнення інформації блок цифрових даних будь-якої довжини розглядається як надвелике ціле додатне число, що представляється набором із трьох невеликих чисел. Таке представлення чисел називається лінійною формою Фібоначчі. Дослідження неадаптивного методу ущільнення даних на основі лінійної форми Фібоначчі [10] показали, що коефіцієнт ущільнення суттєво залежить від типів даних і, крім того, має відносно невелике значення. Тому пропонуються нові методи ущільнення даних на основі лінійної форми Фібоначчі, які мають різні рівні адаптації до типу ущільнюваних даних.

1.1. Узагальнена модель адаптивного ущільнення

Вихідні дані, що підлягають ущільненню, будемо розглядати як послідовність символів 0 і 1. Для подальших досліджень пропонується використовувати числову модель джерела даних. Відповідно до цієї моделі послідовність символів розбивається на блоки, що містять деяку кількість символів, і кожному символу в блоці присвоюється своя певна вага. Якщо вага i-го символу блоку дорівнює 2^i (i=0,1..., l–1), то блок є двійковим кодом деякого числа. Таким чином, вихідні дані представляються як послідовність чисел з деяким законом формування.

Блоку з номером i (i=1;2;..K) відповідає число N_i, яке обчислюється за формулою:

$$N_i = \sum_{j=0}^{l-1} a_{ij} 2^j \ ,$$

де a_{ij} – символи i-го блоку.

Розбиття вихідної послідовності даних здійснюється на блоки, що містять або однакову кількість символів, або різну.

У випадку нерівномірного розбиття довжина блоку знаходиться в межах від $l_{\min}$ до $l_{\max}$ і попередньо не задається. Вона є результатом оптимального вибору серед різних значень довжин блоку. Для блоків довжини, що змінюється від $l_{\min}$ до $l_{\max}$ з певним кроком, виконується ущільнення і вибирається та довжина, при якій забезпечується найбільший коефіцієнт ущільнення.

Пропонується така узагальнена модель процесу адаптивного ущільнення:

$$\mathbf{C}_A = \{\mathbf{P}, \mathbf{A}, \mathbf{M}, \mathbf{C}, \mathbf{D}, \mathbf{P}_M, \mathbf{P}_C, \mathbf{P}^*, S, f\},$$

де $\mathbf{P}$ – вихідна послідовність символів алфавіту $\mathbf{A} = \{0,1\}$;

$\mathbf{M} = \{M_i\}$ – множина правил моделювання джерела даних;

$\mathbf{C} = \{C_j\}$ – множина правил кодування даних;

$\mathbf{D} = \{D_j\}$ – множина правил декодування даних;

$\mathbf{P}_M = \{\mathbf{P}_{M_i}\}$ – множина послідовностей, що є результатом моделювання;

$\mathbf{P}_C = \{\mathbf{P}_{C_{ij}}\}$ – множина послідовностей, що є результатом кодування;

$\mathbf{P}^*$ – послідовність ущільнених даних;

S – правило формування структури послідовності;

f – функція оптимізації.

Вихідна послідовність символів **P** перетворюється в множину послідовностей $\mathbf{P}_M$ з використанням множини правил моделювання **M**. Це описується відображенням $\mathbf{M}: \mathbf{P} \to \mathbf{P}_M$. Конкретному правилу моделювання відповідає відображення $\mathbf{P} \xrightarrow{M_i} \mathbf{P}_{M_i}$.

Кожна з послідовностей множини $\mathbf{P}_M$ кодується з використанням множини правил кодування **C**, в результаті чого формується множина кодованих послідовностей $\mathbf{P}_C$

$$\mathbf{C}: \mathbf{P}_M \to \mathbf{P}_C .$$

Конкретне правило кодування описується функцією $\mathbf{P}_{M_i} \xrightarrow{C_j} \mathbf{P}_{C_{ij}}$.

Вибір із множини кодованих послідовностей $\mathbf{P}_C$ єдиної послідовності $\mathbf{P}^*$, яка відповідає найбільшому коефіцієнту ущільнення, здійснюється за правилом, що описується функцією оптимізації f

$$f: \mathbf{P}_C \to \mathbf{P}^* .$$

Формування структури послідовності $\mathbf{P}^*$ з вказівкою додаткової інформації, необхідної для відновлення послідовності **P**, здійснюється на підставі правила S.

Таким чином, процес адаптивного ущільнення описується композицією відображень:

$$C_A = M \circ C \circ f \circ S .$$

Пропонується така узагальнена модель процесу відновлення даних:

$$\mathbf{DC} = \{\mathbf{A}, \mathbf{P}^*, \mathbf{P}, \mathbf{D}, S\},$$

де $\mathbf{A} = \{0,1\}$ – алфавіт;

$\mathbf{P}^*$ – послідовність ущільнених даних;

P – послідовність відновлених (вихідних) даних;

D – множина правил декодування даних;

S – правило формування структури послідовності $\mathbf{P}^*$.

1.2. Правила моделювання джерела даних

Змінюючи кількість символів у блоці, будемо отримувати послідовності чисел з різними законами їх формування. Тобто для різних l будемо мати різні моделі джерела даних. Модель джерела даних можна змінювати, змінюючи не тільки розрядність блоку, але і змінюючи за певним правилом символи в блоці.

Для простоти реалізації пропонується записувати в самий старший розряд блоку символ 1, якщо цей розряд містить 0. Тобто блок виду 0xx…x замінюється на блок виду 1xx ... x, де x означає будь-який символ (1 або 0). Таке правило моделювання будемо позначати m(1).

Можливість зміни моделі джерела даних дозволяє вибирати таку модель, яка забезпечує найбільший коефіцієнт ущільнення для заданого правила кодування. Це є основою методів адаптивного ущільнення даних.

Уведемо такі позначення правил моделювання.

Довжина блоку l може змінюватися від $l_{\min}$ до $l_{\max}$ з деякою дискретністю. Пропонується два підходи до формування дискретних значень l.

Перший підхід передбачає такі значення l:

$$\underset{l_{\min}}{2^a}, 2^{a+1}, 2^{a+2}, \ldots, \underset{l_{\max}.}{2^{a+h}}$$

Відповідно до цього, будемо позначати правила моделювання таким чином:

$$M_i = \mathrm{m}(a, i),\ i = 0,1,\ldots, h.$$

Мінімальне значення a вибирається виходячи з необхідного мінімального коефіцієнта ущільнення. Максимальне значення суми $a + h$ визначається допустимими витратами апаратури і часу ущільнення та відновлення.

Відповідно до другого підходу довжина l приймає такі значення:

$$\underset{l_{\min}}{2^a}, \left(2^a + 1\cdot 2^b\right), \left(2^a + 2\cdot 2^b\right), \ldots, \underset{l_{\max}.}{\left(2^a + g\cdot 2^b\right)}$$

Виходячи з цього, правила моделювання будемо позначати таким чином:

$$M_i = \mathrm{m}\,(a, b, i)\ \ i=0,1,\ldots, g.$$

У даному випадку мінімальне значення a вибирається так саме як і у попередньому випадку. Значення b повинно задовольняти умову: $b \le a$. Значення g вибирається з урахуванням умови, що довжина блоку $l_{\max} = \left(2^a + g\cdot 2^b\right)$ задовольняє вимогам допустимих апаратних і часових витрат.

1.3. Правила кодування та декодування блоків даних

Пропонується набір правил перетворення блоків даних, отриманих у результаті моделювання джерела даних, з метою забезпечення можливості вибору виду перетворення, що дає найбільший коефіцієнт ущільнення.

Для вказівки правила кодування до результату перетворення блоку даних дописуються ознаки.

Пропонується два правила перетворення блоків, що забезпечують ущільнення даних.

Перше правило передбачає відкидання старших нулів у коді. Якщо блок має вигляд:

$$\underbrace{\underbrace{0\,0\ldots0}_{q}1\text{xx}\ldots\text{x}}_{l},$$

то його можна перетворити до вигляду:

$$\underbrace{1\text{xx}\ldots\text{x}}_{l-q}\|\underbrace{\;q\;}_{l_q}$$

за умови $q > l_q + 1$, де $l_q = \left]\log_2 l\right[$.

Друге правило передбачає одержання лінійної форми Фібоначчі для блоку даних. При цьому можливі два підходи.

Перший підхід полягає в безпосередньому перетворенні блоку в лінійну форму Фібоначчі. В цій формі будь-яке ціле додатне число представляється у вигляді [9]:

$$z = a_0 F(j) + b_0 F(j+1),$$

де a_0, b_0 – цілі додатні числа (координати представлення);

j – ціле додатне число (індекс представлення);

$F(j)$ – j-е число Фібоначчі.

Другий підхід передбачає запис одиниці в самий старший розряд блоку вигляду 0xx…x, тобто реалізується правило моделювання m(1), а потім виконується перетворення в лінійну форму Фібоначчі.

Можлива ситуація, коли довжина блоку, що представляє лінійну форму, більше довжини початкового блоку і перетворення за першим правилом недоцільне. У цьому випадку рекомендується представити вихідний блок неперетвореним.

З урахуванням вищезазначеного маємо чотири правила кодування, які породжують відповідні структури кодованих блоків (табл.1). Причому кожне з правил передбачає формування ознак структури p_1 і p_2.

Таблиця 1

Правила кодування і структури блоків

Позначення правила кодування	Ознаки		Позначення структури блоку
	p_2	p_1	
C_0 (0,0)	0	0	STR0
C_1 (0,1)	0	1	STR1
C_2 (1,0)	1	0	STR2
C_3 (1,1)	1	1	STR3

Структура STR0 відповідає правилу, коли вихідний блок представляється неперетвореним.

Вона має вигляд

$$\underbrace{1xx\ldots x}_{l} \| 0 \| 0 \text{ чи } \underbrace{\underbrace{00\ldots0}_{q}1xx\ldots x}_{l} \| 0 \| 0$$

для $q <]\log_2 l[$.

Структура STR1 породжується правилом, що передбачає відкидання q старших нулів у коді вихідного блоку, і має вигляд:

$$\underbrace{1xx\ldots x}_{l-q} \| \underbrace{\quad q \quad}_{l_q} \| 0 \| 1.$$

Нехай $l = 2^n$, тоді $0 \le q \le (2^n - 1)$. Для представлення значення q в такому діапазоні необхідно $l_q =]\log_2 2^n[= n$ розрядів. Виходячи із цього, структура STR1 створюється коли $q > n+1$.

Структура STR2 одержується із блоків вигляду:

$$\underbrace{1xx\ldots x}_{l} \text{ чи } \underbrace{\underbrace{00\ldots0}_{q}1xx\ldots x}_{l},$$

застосуванням правила перетворення в лінійну форму Фібоначчі.

Ця структура описується кортежем $\langle b_0 \| a_0 \| j \| l_{b_0} \| l_{a_0} \| 1 \| 0 \rangle$, де l_{a_0} – розрядність коду числа a_0; l_{b_0} – розрядність коду числа b_0.

Структура STR3 породжується правилом перетворення в лінійну форму Фібоначчі блоку вигляду:

$$\underbrace{\underbrace{00\ldots0}_{q}1xx\ldots x}_{l},$$

після запису одиниці в самий старший розряд. Тобто фактично перетворюється блок вигляду:

$$\underbrace{1\underbrace{0\ldots0}_{q-1}1xx\ldots x}_{l}.$$

Структура STR3 описується кортежем $\langle b_0 \| a_0 \| j \| l_{b_0} \| l_{a_0} \| 1 \| 1 \rangle$.

Розглянемо правила декодування даних для кожної з описаних вище структур. Умовні позначення таких правил наведені в табл. 2.

Вибір правила декодування для заданого блоку здійснюється на основі аналізу ознак p_1 і p_2 структури блоку. Після вибору правила декодування розряди ознак відкидаються і подальші дії виконуються над основною частиною структури блоку.

Якщо p_1=0 і p_2=0, то це означає, що блок має структуру STR0, до якої застосовується правило декодування $D_0(0,0)$. Дане правило полягає в тому, що основна частина структури блоку залишається незмінною.

Таблиця 2

Позначення правил декодування

Позначення структури блоку	Позначення правила декодування
STR0	$D_0\,(0,0)$
STR1	$D_1\,(0,1)$
STR2	$D_2\,(1,0)$
STR3	$D_3\,(1,1)$

Якщо p_1=1 і p_2=0, то до блоку структури STR1 необхідно застосувати правило $D_1(0,1)$. Воно полягає у виконанні таких дій. В основній частині структури

$$\underbrace{1\mathrm{xx}\ \ldots \mathrm{x}}_{l-q}\,\|\,\underbrace{\ q\ }_{l_q}$$

відкинути l_q молодших розрядів і до частини, що залишилася, дописати q старших нулів. У результаті цього буде отримано блок вигляду:

$$\underbrace{\underbrace{0\,0\ldots0}_{q}\,1\mathrm{xx}\ \ldots \mathrm{x}}_{l}\ .$$

Якщо p_1=0 і p_2=1, то до блоку структури STR2 застосовується правило $D_2(1,0)$. За цим правилом, виходячи з чисел j, a_0 і b_0, виконується зворотне перетворення Фібоначчі, тобто обчислюється ціле додатне число N, що відповідає даній лінійній формі Фібоначчі. Після цього визначається розрядність l^* двійкового коду числа N і обчислюється різниця $\Delta = l - l^*$. Дана різниця показує, скільки нулів необхідно дописати в старші розряди коду числа N.

Якщо p_1=1 і p_2=1, то до блоку структури STR3 застосовується правило $D_3(1,1)$. Воно передбачає обчислення числа за заданою лінійною формою Фібоначчі й відкидання самої старшої одиниці в коді цього числа.

1.4. Функції оптимізації

Означення 1.1. Кодована послідовність $\mathbf{P}_{c_{il}}$, що складається з блоків, закодованих за одним правилом, називається *однорідною послідовністю*.

В якості такого правила будемо використовувати правило $C_2(1,0)$ перетворення в лінійну форму Фібоначчі.

Означення 1.2. Кодована послідовність $\mathbf{P}_{c_{il}}$, що складається з блоків, закодованих за різними правилами, називається *неоднорідною послідовністю*.

В якості таких правил можуть бути використані всі чотири, описані вище правила.

Довжина однорідної послідовності обчислюється за формулою:

$$L_{\text{оп}} = \sum_{i=1}^{K} l_{2,i} \ ,$$

де $l_{2,i}$ – довжина i-го блоку структури STR2.

Неоднорідна послідовність має довжину, яка обчислюється так:

$$L_{\text{нп}} = \sum_{i=1}^{n_0} l_{0,i} + \sum_{j=1}^{n_1} l_{i,j} + \sum_{k=1}^{n_2} l_{2,k} + \sum_{g=1}^{n_3} l_{3,g} \ ,$$

де n_0 – кількість блоків структури STR0;
n_1 – кількість блоків структури STR1;
n_2 – кількість блоків структури STR2;
n_3 – кількість блоків структури STR3;
$l_{0,i}$ – довжина i-го блоку структури STR0;
$l_{1,j}$ – довжина j-го блоку структури STR1;
$l_{2,k}$ – довжина k-го блоку структури STR2;
$l_{3,g}$ – довжина g-го блоку структури STR3;

$$n_0 + n_1 + n_2 + n_3 = K \ .$$

Правило вибору з множини кодованих послідовностей $\mathbf{P}_C$ єдиної послідовності $\mathbf{P}^*$, що відповідає найбільшому коефіцієнту ущільнення, описується функцією оптимізації f.

Пропонуються такі варіанти функції оптимізації.

Перший варіант обмежується вибором на рівні блоків однорідної послідовності. При цьому функція оптимізації має вигляд:

$$f_{\text{бл}}^{\text{оп}} = \min\{l_2^{(i)}\} \ ,$$

де $l_2^{(i)}$ – довжина блоку структури STR2 для i-го правила моделювання.

Якщо правило моделювання $m(a,i)$, то i=0,1,... h. Для правила моделювання $m(a,b,i)$ маємо i=0,1, ... g.

Функція оптимізації на рівні блоків, але для неоднорідної послідовності, має вигляд:

$$f_{\text{бл}}^{\text{нп}} = \min\{l_0, l_1, l_2 l_3\} \ ,$$

де l_0, l_1, l_2, l_3 – довжина блока структури STR0, STR1, STR2, STR3, відповідно.

Функція оптимізації на рівні послідовностей має вигляд:

$$f_{\text{посл}} = \min\{L_{\text{оп}}^{(i)}\} \ ,$$

де $L_{\text{оп}}^{(i)}$ – довжина однорідної послідовності для i-го правила моделювання.

Комбінована функція оптимізації забезпечує вибір блоків мінімальної довжини в межах однієї послідовності і вибір послідовності мінімальної довжини для різних правил моделювання. Вона має вигляд:

$$f_{\text{комб}} = \min\{(\min\{l_0, l_1, l_2, l_3\})^{(i)})\}.$$

Таким чином, пропонується чотири варіанти функції оптимізації.

1.5. Оцінки потенційних можливостей ущільнення на основі лінійної форми Фібоначчі

Лінійна форма Фібоначчі в кодованому вигляді представляється блоком структури STR2 або STR3. Ці структури не відрізняються одна від одної ні складом, ні розрядністю складових частин. Різниця є тільки в цифрі однієї ознаки. Тому будемо розглядати ці дві структури як одну.

Довжина блоку вказаної структури обчислюється за формулою:

$$l_2 = l_3 = l_{a_0} + l_{b_0} + l_j + n_{a_0} + n_{b_0} + 2,$$

де l_{a_0} – розрядність двійкового коду числа a_0;

l_{b_0} – розрядність двійкового коду числа b_0;

l_j – розрядність двійкового коду числа j;

n_{a_0} – розрядність двійкового коду числа l_{a_0};

n_{b_0} – розрядність двійкового коду числа l_{b_0}.

Визначимо максимальні можливі значення зазначених розрядностей, виходячи з умови $l_2^{\max} = l_3^{\max} = l$.

Розряди блоку, що зайняті кодами чисел a_0 і b_0, можуть перерозподілятися таким чином. Коли $a_0 = a_{\max}$ і $b_0 = 0$, то всі розряди належать коду числа $a_{\max}$. В ситуації $a_0 = 0$ і $b_0 = b_{\max}$, всі розряди належать коду числа $b_{\max}$. Звідси випливає, що $l_{a_0} + l_{b_0} = l_{a_{\max}} = l_{b_{\max}}$.

Для представлення числа $l_{a_{\max}}$ потрібно $]\log_2 l_{a_{\max}}[$ двійкових розрядів. Така ж кількість розрядів потрібна для представлення числа $l_{b_{\max}}$.

Кількість розрядів, необхідних для представлення числа $j_{\max}$, визначається за формулою:

$$l_{j\max} =]\log_2(Rl)[,$$

де $R = 1{,}5$ – надмірність коду Фібоначчі.

З урахуванням наведених оцінок розрядностей складових частин структури блоку, маємо співвідношення:

$$l = l_{a_{\max}} + l_{j_{\max}} + 2]\log_2 l_{a_{\max}}[+2,$$

на основі якого визначаються розрядності частин структури для заданої розрядності початкового блоку. Такі значення наведені в табл. 3.

Таблиця 3

Розрядності складових частин структури блоку

l	l_j	$l_{a_0}+l_{b_0}$	n_{a_0}	n_{b_0}	l	l_j	$l_{a_0}+l_{b_0}$	n_{a_0}	n_{b_0}
32	8	12	5	5	2048	17	2007	11	11
64	9	41	6	6	4096	18	4052	12	12
128	11	101	7	7	8192	20	8144	13	13
256	12	226	8	8	16384	21	16333	14	14
512	14	478	9	9	32768	23	32713	15	15
1024	15	987	10	10	65536	24	65478	16	16

Відзначимо, що розрядності кодів чисел j, l_{a_0} і l_{b_0} не залежать від змісту початкового блоку і визначаються тільки максимальними значеннями цих чисел. Це означає, що частина структури перетвореного блоку має незмінну розрядність, а частина – розрядність, що може змінюватися. Останню частину складають коди чисел a_0 і b_0. Розрядність цій частини може визначатися з точки зору коефіцієнта ущільнення.

Максимальний або потенційно можливий коефіцієнт ущільнення визначається за формулою:

$$S_{\max} = \frac{l}{l_{2\min}},$$

де $l_{2\min}$ – мінімально можлива довжина перетвореного блоку;

$$l_{2\min} = l_j + n_{a_0} + n_{b_0} + 3.$$

Мінімальне значення a_0 і b_0 можуть бути такими:

$$a_0 = 1,\ b_0 = 0 \text{ і } a_0 = 0,\ b_0 = 1.$$

З урахуванням цього $l_{a_0} + l_{b_0} = 1$.

Значення максимального коефіцієнта ущільнення для різних розрядностей початкового блоку наведені в табл. 4. Аналіз цієї таблиці показує, що, починаючи з розрядності початкового блоку $l = 32$, коефіцієнт ущільнення більше одиниці. Причому зі збільшенням розрядності в два рази коефіцієнт ущільнення зростає більш ніж у 1,8 рази.

Таблиця 4

Значення $S_{\max}$

l	16	32	64	128	256	512	1024
$S_{\max}$	0,9	1,5	2,7	4,5	8,3	14,6	26,9
l	2048	4096	8192	16384	32768	65536	—
$S_{\max}$	48,4	91,0	167,2	315,1	585,1	1110,8	—

1.6. Метод ущільнення одним проходом з рівномірним розбиттям на блоки

Пропонується метод ущільнення (метод ОПРР), який полягає в тому, що вихідні дані розбиваються на блоки однакової довжини, кожен з яких кодується незалежно за чотирма правилами і з чотирьох результатів кодування вибирається той результат, структура блоку якого має найменшу довжину.

При цьому множина правил моделювання джерела даних складається із двох правил $\mathbf{M} = \{m(a,0), m(1)\}$. Перше правило означає, що вихідні дані розбиваються на блоки довжини 2^a розрядів. Друге правило моделювання джерела даних полягає в записі символу «1» в самий старший розряд блоку, якщо він містить нуль.

Множина правил кодування має вигляд $\mathbb{C} = \{C_0(0,0), C_1(0,1), C_2(1,0), C_3(1,1)\}$. Перетворення, які виконуються над блоком вихідних даних, згідно із цими правилами зазначені вище.

Оскільки блоки перетвореної послідовності даних, у загальному випадку, будуть мати різну структуру, то результатом ущільнення є неоднорідна послідовність. Виходячи із цього, маємо таку функцію оптимізації: $f_{\text{бл}}^{\text{нп}} = \min\{l_0, l_1, l_2, l_3\}$.

Як відзначалося вище, у структурі перетвореної послідовності $\mathbf{P}^*$ необхідно помістити інформацію, що забезпечує правильне відновлення вихідної послідовності $\mathbb{P}$.

В основному полі розміщуються перетворені блоки, структура яких визначена за правилом $f_{\text{бл}}^{\text{нп}}$. У додаткове поле розміщується інформація, яка забезпечує правильне відновлення.

Структури блоків STR1 і STR3 забезпечують правильне відновлення вихідних даних без додаткової інформації.

Якщо згідно зі структурою STR1 до коду 1хх......х розрядності $l-q$ дописати q старших нулів, то в результаті буде отримано вихідний код розрядності l.

Лінійній формі Фібоначчі, яка представляється структурою STR3, після зворотного перетворення буде відповідати число, двійковий код якого буде мати l розрядів.

Структура STR2 не забезпечує однозначного відновлення даних без додаткової інформації. Якщо вихідний блок мав вигляд 1хх......х, то після відновлення буде отриманий код довжини l розрядів. Якщо ж вихідний блок мав вигляд 00...0…0 1хх…х, то після відновлення буде отриманий код виду 1хх......х і довжини $l-q$. Звідси випливає, що необхідно додатково вказати розрядність вихідного блоку. Врахування цієї інформації дозволить отримати правильний результат шляхом дописування певної кількості нулів у старші розряди коду.

Для структури STR0 також необхідна інформація про довжину блоку.

У загальному випадку, при розбитті на блоки довжини l один блок може мати довжину менше l. Тому необхідно або вказати довжину послідовності L і потім обчислити довжину меншого блоку, або вказати довжину l_k цього блоку. Другий варіант вимагає меншої розрядності коду, тому вибираємо його.

Таким чином, додаткова інформація містить у собі значення розрядності l і l_k.

З урахуванням вищесказаного, маємо таке правило формування структури послідовності $\mathbf{P}^*$, яке представляється у вигляді кортежу:

$$\mathbf{S} = \left\{ \text{Бл}^* K \,\|\, \ldots \text{Бл}^* 2 \,\|\, \text{Бл}^* 1 \,\|\, l_K \,\|\, 2^a \right\}.$$

Розрядність полів $\text{Бл}^* i$ $(i = 1 \div K)$ різна. Розрядність вихідних блоків може досягати значення 2^{16}, тому розрядність полів l_K і 2^a дорівнює 16.

Безпосередньо процес ущільнення складається з K циклів. У кожному з них зчитується черговий блок вихідних даних Блi, над яким виконуються перетворення згідно правил кодування $C_0(0,0)$, $C_1(0,1)$, $C_2(1,0)$, $C_3(1,1)$ і створюються структури STR0, STR1, STR2, STR3. Для кожної структури визначається її довжина та відповідно до правила оптимізації $\min\{l_0, l_1, l_2, l_3\}$ вибирається структура найменшої довжини, якою доповнюється послідовність ущільнених даних $\mathbf{P}^*$.

Процес відновлення даних визначається структурою послідовності ущільнених даних $\mathbf{S} = \left\{ \text{Бл}^* K \,\|\, \ldots \text{Бл}^* 2 \,\|\, \text{Бл}^* 1 \,\|\, l_K \,\|\, 2^a \right\}$ і множиною правил декодування $\mathbf{D} = \{D_0(0,0), D_1(0,1), D_2(1,0), D_3(1,1)\}$.

Безпосередньо процес відновлення складається з K циклів. На початку кожного циклу зчитується два наймолодші розряди послідовності $\mathbf{P}^*$, що є кодом ознаки структури. Відповідно до цього коду реалізується певне правило декодування. Результатом є черговий блок відновлених даних, що додається в послідовність **P.** Структура, з якої отриманий цей блок, видаляється з послідовності $\mathbf{P}^*$.

1.7. Метод ущільнення фіксованою кількістю проходів з рівномірним розбиттям на блоки

Пропонований тут метод ущільнення будемо скорочено називати методом ФКПРР. Даний метод полягає в тому, що реалізується (h + 1) методів ОПРР для різних довжин блоків і вибирається один результат ущільнення, що має найменшу довжину послідовності $\mathbf{P}^*$.

Множина правил моделювання джерела даних складається з (h + 2) правил: $\mathbf{M} = \{m(a,0), m(a,1), \ldots, m(a,h), m(1)\}$.

i-й прохід ущільнення описується двома правилами моделювання m(a,i) і m(1).

Оскільки кожен прохід ущільнення реалізується за методом ОПРР, тому використовується така множина правил кодування: $\mathbf{C} = \{C_0(0{,}0), C_1(0{,}1), C_2(1{,}0), C_3(1{,}1)\}$.

Як функція оптимізації використовується комбінована функція $f_{\text{комб}} = \min\{(\min\{l_0, l_1, l_2, l_3\})^{(i)})\}$, що забезпечує вибір блоків мінімальної довжини при реалізації кожного проходу ущільнення і вибір послідовності мінімальної довжини серед результатів усіх проходів ущільнення.

Правило формування послідовності $\mathbf{P}^*$ для i-го проходу описується кортежем: $\mathbf{S} = \{\text{Бл}^* K \,\|\, \ldots \text{Бл}^* 2 \,\|\, \text{Бл}^* 1 \,\|\, l_K \,\|\, 2^{a+i}\}$.

Безпосередньо процес ущільнення складається з (h+1) циклів, у кожному з яких реалізується процес ущільнення за методом ОПРР. На початку циклу визначаються значення параметрів для ініціалізації i-го проходу ущільнення. Якщо довжина L_i^* послідовності $\mathbf{P}_i^*$ менше раніше визначеної мінімальної довжини L*, то відбувається оновлення послідовності ущільнених даних $\mathbf{P}^*$ і значення мінімальної довжини L*. У такий спосіб крок за кроком реалізується функція оптимізації.

Процес відновлення даних за даним методом аналогічний процесу за методом ОПРР.

1.8. Метод ущільнення змінною кількістю проходів з рівномірним розбиттям на блоки

Пропонується метод, суть якого полягає в тому, що реалізується кілька проходів ущільнення за методом ОПРР із порівнянням довжин послідовностей $\mathbf{P}^*$ ущільнених даних. При цьому результат ущільнення i-м проходом $\mathbf{P}_i^*$ є послідовністю вихідних даних $\mathbf{P}_{i+1}$ для (i+1)-го проходу.

Якщо довжина L_i^* послідовності ущільнених даних, отриманих у результаті i-го проходу, більше довжини L_{i+1}^* послідовності, отриманої на $(i+1)$-му проході, то послідовність $\mathbf{P}_{i+1}^*$ вважається кращим результатом ущільнення і досліджується можливість подальшого ущільнення шляхом реалізації $(i+2)$-го проходу. Якщо ж $L_i^* < L_{i+1}^*$, то процес ущільнення закінчується і результатом ущільнення вважається послідовність $\mathbf{P}_i^*$.

Мінімальна кількість проходів може бути 2, а максимальна - обмежується деяким значенням $m_{\max}$.

На рис. 1 показані приклади структур послідовностей даних для трьох проходів.

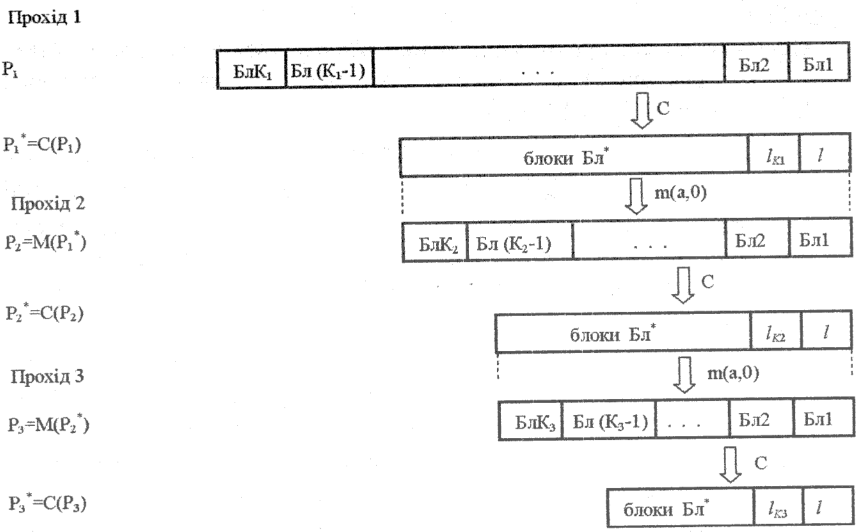

Рис. 1. Структури послідовностей даних для трьох проходів

Послідовність вихідних даних $\mathbf{P}_1$ розбивається на $(K_1 - 1)$ блоків довжини l і один блок довжини l_{K_1}. Кожен з цих блоків піддається ущільненню і до послідовності ущільнених блоків Бл^* дописуються значення l і l_{K_1}. Це є результатом першого проходу $\mathbf{P}_1^*$ і він розглядається як послідовність вихідних даних $\mathbf{P}_2$ для другого проходу і розбивається на $(K_2 - 1)$ блоків довжини l і один блок довжини l_{K_2}. Результат ущільнення цих блоків доповнюється значеннями l і l_{K_2}.

Послідовність $\mathbf{P}_2^*$ далі розглядається як послідовність вихідних даних $\mathbf{P}_3$ для третього проходу та розбивається на $(K_3 - 1)$ блоків довжини l і один блок довжини l_{K_3}. Результат реалізації третього проходу складають ущільнені блоки Бл^* і значення l і l_{K_3}.

Множина правил моделювання джерела даних складається із двох правил: $\mathbf{M} = \{m(a,0), m(1)\}$.

Правила вибору блоків мінімальної довжини при реалізації кожного проходу ущільнення і вибору послідовності ущільнених даних мінімальної довжини описуються комбінованою функцією оптимізації:

$$f_{\text{комб}} = \min\{(\min\{l_0, l_1, l_2, l_3 \ )^{(m)}\}.$$

Правило формування послідовності $\mathbf{P}_i^*$, як результату i-го проходу, описується кортежем $\mathbf{S} = \{Бл^* K_m \| Бл^* 2 \| Бл^* 1 \| l_{km} \| l \| m\}$.

У кожному з m циклів обчислюються значення K_m і l_{K_m}, здійснюється ущільнення за методом ОПРР із формуванням результату $\mathbf{P}_m^*$ і визначається довжина L_m^* послідовності $\mathbf{P}_m^*$. Якщо $L_m^* < L^*$, то реалізується наступний прохід. У випадку $L_m^* > L^*$ процес ущільнення закінчується і до послідовності $\mathbf{P}^*$ дописується кількість проходів, реалізованих для одержання цієї послідовності.

При створенні $\mathbf{P}_m^*$ реалізується складова функції оптимізації $\min\{l_0, l_1, l_2, l_3\}$, а порівняння $L_m^* < L^*$ відповідає складовій $\min\{(\ )^{(m)}\}$.

Перша складова передбачає наявність одночасно чотирьох значень довжин блоків, серед яких вибирається найменше значення.

Друга складова реалізується вибором із двох значень: попереднього і даного. При цьому попереднє значення було мінімальним.

Процес відновлення даних визначається структурою послідовності ущільнених даних $\mathbf{S} = \{Бл^* K_m \| \ldots \| Бл^* 2 \| Бл^* 1 \| l_{K_m} \| l \| m\}$ і множиною правил декодування $\mathbf{D} = \{D_0(0,0), D_1(0,1), D_2(1,0), D_3(1,1)\}$.

Ініціалізація процесу складається з таких дій. Спочатку вводиться послідовність $\mathbf{P}^*$ даних, що підлягають відновленню. Потім зчитуються 4 наймолодші розряди, які є кодом числа m. Після цього дані розряди видаляються з послідовності $\mathbf{P}^*$.

У кожному з m циклів виконуються такі дії. Спочатку зчитуються 16 наймолодших розрядів, які є кодом числа l. Після цього дані розряди видаляються з послідовності $\mathbf{P}^*$. Потім зчитуються і видаляються з $\mathbf{P}^*$ 16 наймолодших розрядів, які є кодом числа l_{K_m}.

Створюється порожня послідовність, призначена для формування послідовності відновлених даних. Формування здійснюється шляхом реалізації K_m циклів, у кожному з яких виконуються такі дії.

Зчитуються два наймолодші розряди, вміст яких визначає правило декодування відповідної структури даних Бл. Структура видаляється з послідовності $\mathbf{P}^*$, а блок заноситься в послідовність $\mathbf{P}$.

Після завершення циклу формування $\mathbf{P}$ послідовність даних, відновлених у даному циклі розглядається як послідовність даних, яка підлягає відновленню в наступному циклі. Приклад перетворення структур послідовностей показаний на рис. 2.

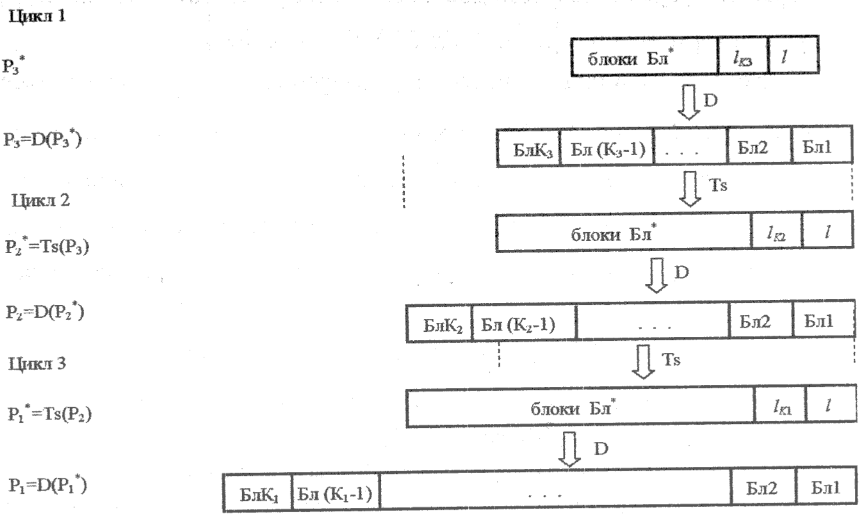

Рис. 2. Структури послідовностей даних для трьох циклів відновлення

У першому циклі відновлення блоки Бл* послідовності $\mathbf{P}_3^*$, що мають певні структури, декодуються за відповідними правилами і формується послідовність $\mathbf{P}_3$ відновлених даних. Вихідна послідовність для другого циклу утворюється шляхом трансформації послідовності $\mathbf{P}_3$ згідно зі структурою послідовності ущільнених даних, тобто $\mathbf{P}_2$=T$_3$($\mathbf{P}_3$). Декодування блоків Бл* послідовності $\mathbf{P}_2^*$ породжує послідовність $\mathbf{P}_2$. Третій цикл також складається із трансформації і декодування послідовностей.

1.9. Метод ущільнення одним проходом з нерівномірним розбиттям на блоки

Пропонований тут метод будемо скорочено називати методом ОПНР.

Суть цього методу полягає в такому (рис. 3). Лінійною формою Фібоначчі представляються числа, що відповідають блокам довжини від $l_{min} = l$ до $l_{max} = l + gl_0$. При цьому довжина блоків змінюється з дискретністю l_0. Серед усіх форм вибирається форма, що забезпечує найбільший коефіцієнт ущільнення ($S_{ущ}$).

На рис. 3 це форма F$_i$. Вона є результатом перетворення блоку Блi вихідних дані довжини ($l+il_0$). Після видалення блоку Блi з послідовності **P** до частини, що залишилася, застосовуються описані дії та визначається блок Блj, лінійна форма якого забезпечує найбільший коефіцієнт ущільнення. При цьому довжина блоку Бл j може відрізнятися від довжини блоку Блi. У такий спосіб довжина блоків, на які буде розбиватися послідовність **P**, буде різною й адаптованою під найбільший коефіцієнт ущільнення.

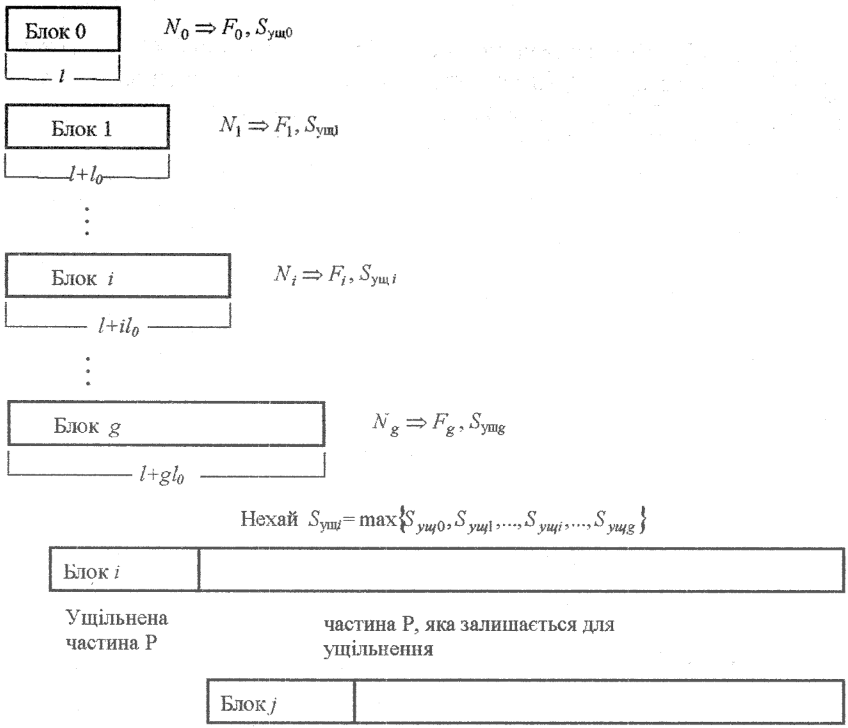

Рис. 3. Схема нерівномірного розбиття на блоки

Множина правил моделювання джерела даних для даного методу складається з g+2 правил:

$$\mathbf{M} = \left\{m(a,b,0), m(a,b,1), \ldots, m(a,b,g), m(1)\right\}.$$

Згідно із цими правилами довжина блоку вихідних даних приймає значення від $l_{\min} = 2^a$ до $l_{\max} = 2^a + g \cdot 2^b$ з дискретністю 2^b. Тобто блоки мають такі довжини:

$$2^a, \left(2^a + 2^b\right), \left(2^a + 2 \cdot 2^b\right), \ldots, \left(2^a + g \cdot 2^b\right).$$

Значення a вибирається виходячи з мінімально припустимого коефіцієнта ущільнення. Значення b повинне задовольняти умову $b \le a$. Чим менше значення b, тим менша дискретність довжини і тим більший коефіцієнт ущільнення може бути отриманий. Але при цьому буде значно збільшуватися час ущільнення.

Оскільки для ущільнення використовується тільки лінійна форма Фібоначчі, тому множина правил кодування складається тільки з двох правил: $\mathbf{C} = \{C_2(1,0), C_3(1,1)\}$.

Правило $C_2(1,0)$ породжує структуру STR2, а правило $C_3(1,1)$ – структуру STR3. Вище було показано, що для відновлення вихідних даних зі структури STR3 не потрібно додаткових даних, а відновлення зі структури STR2 неможливе без знання довжини блоку вихідних даних у випадку, коли він має вигляд 0 . . .0 1x . . .x. Для того, щоб кожен блок ущільнених даних, що має структуру STR2, не доповнювати значенням довжини l, пропонується модифікувати правило $C_2(1,0)$. Зміна полягає в тому, що структура STR2 буде створюватися тільки для блоку вихідних даних вигляду 1 x x...x. З урахуванням цього, у процесі ущільнення блоків вихідних даних буде створюватися тільки одна зі структур: STR2 або STR3. Якщо блок вихідних даних має вигляд 1x x... x, то він перетвориться в структуру STR2. Для блоку вихідних даних виду 0...0 1 xx.. x створюється структура STR3.

За даним методом оптимізація здійснюється на рівні блоків і описується функцією:

$$f_{\text{бл}} = \max\left\{S_{\text{ущ}0}, S_{\text{ущ}1}, \ldots, S_{\text{ущ}g}\right\},$$

де $S_{\text{ущ}i}$ – коефіцієнт ущільнення для i-го правила моделювання, i=0,1,..., g.

$$S_{\text{ущ}\,i} = \frac{l_{\text{бл}\,i}}{l_{\text{S}\,i}},$$

де $l_{\text{бл}\,i}$ – довжина блоку вихідних даних для і-го правила моделювання;

$$l_{\text{бл}\,i} = 2^a + i \cdot 2^b ;$$

$l_{\text{S}\,i}$ – довжина блоку ущільнених даних (структура STR2 або STR3) для i-го правила моделювання.

$$l_{\text{S}\,i} = l_{a_0} + l_{b_0} + l_j + n_{a_0} + n_{b_0} + 2 ,$$

де l_{a_0} – розрядність коду числа a_0 ;

l_{b_0} – розрядність коду числа b_0 ;

l_j – розрядність коду числа j;

n_{a_0} – розрядність коду числа l_{a_0};

n_{b_0} – розрядність коду числа l_{b_0} .

Оскільки для відновлення структури блоків, з яких складається послідовність $\mathbf{P}^*$ не потрібна додаткова інформація, тому правило формування структури послідовності $\mathbf{P}^*$ описується кортежем:

$$\mathbf{S} = \left\{\text{Бл}^* v \,\|\, \ldots \,\|\, \text{Бл}^* 2 \,\|\, \text{Бл}^* 1\right\}.$$

Ініціалізація процесу ущільнення складається з таких дій: введення послідовності **P** вихідних даних, параметрів моделей джерела даних a, b, g, створення порожньої послідовності ущільнених даних $\mathbf{P}^*$=( ) і встановлення початкових значень i=0 та $S_{ущ}$.

Безпосередньо процес ущільнення складається з декількох циклів отримання блоків Бл* ущільнених даних. Оскільки розрядність блоків вихідних даних, що дають найкращий результат, визначається значенням $S_{ущ}$, то і кількість блоків Бл залежить від конкретного змісту послідовності **P**.

Найбільша кількість блоків $N_{бл}^{max}$, на які буде розбиватися послідовність **P**, буде у випадку їхньої мінімальної довжини, тобто $N_{бл}^{max} = L/2^a$.

Найменша кількість блоків $N_{бл}^{min}$ буде у випадку їхньої максимальної довжини $N_{бл}^{min} = L/\left(2^a + g \cdot 2^b\right)$.

Середня кількість блоків $N_{бл}^{ср}$ обчислюється за формулою:

$$N_{бл}^{ср} = \frac{L}{\sum\limits_{i=0}^{g} p_i\left(2^a + i \cdot 2^b\right)},$$

де p_i – ймовірність появи блоку довжини $\left(2^a + i \cdot 2^b\right)$.

Кількість блоків, на які розбивається послідовність **P**, і визначає кількість циклів одержання блоків Бл* ущільнених даних. Кожен блок Бл, що забезпечує найбільший коефіцієнт ущільнення, видаляється з послідовності **P**, і процес ущільнення закінчується при видаленні останнього блоку.

Одержання кожного блоку Бл* ущільнених даних, у свою чергу, складається з (g+1) циклів. На початку кожного циклу обчислюється нове значення довжини блоку Бл. Потім формується цей блок шляхом зчитування l розрядів з послідовності **P.** Якщо старший розряд блоку Бл містить цифру 0, то даний блок ущільнюється шляхом створення структури STR3. Якщо ж старший розряд містить цифру 1, то створюється структура STR2. Для створеної структури визначається її довжина, на підставі якої обчислюється коефіцієнт ущільнення. Його значення присвоюється цикловому коефіцієнту $S_{ущi}$.

Якщо $S_{ущi}$ перевищує найбільший коефіцієнт ущільнення $S_{ущ}$, отриманий на попередніх циклах, то його вважають найбільшим і блок Бл* ущільнених даних заміняють блоком Бл$^*i$. По закінченню циклу блок Бл*, що відповідає найбільшому коефіцієнту ущільнення, записується в послідовність **P*** ущільнених даних.

Процес відновлення даних визначається структурою послідовності ущільнених даних $S = \left\{Бл^* v \,||\, ... \,||\, Бл^* 2 \,||\, Бл^* 1\right\}$ і множиною правил декодування $\mathbf{D} = \{D_2(1{,}0), D_3(1{,}1)\}$.

Ініціалізація процесу складається з двох дій: введення послідовності $\mathbf{P}^*$ і створення порожньої послідовності відновлених даних $\mathbf{P} = (\ )$.

Безпосередньо процес відновлення складається з циклів, в кожному з яких спочатку зчитуються два наймолодші розряди послідовності $\mathbf{P}^*$, що містять ознаки структури блоку ущільнених даних, а потім відновлюється блок даних Бл шляхом реалізації відповідного правила декодування структури. Після цього структура видаляється з послідовності $\mathbf{P}^*$, а блок відновлених даних Бл записується в послідовність $\mathbf{P}$. Цикли закінчуються, коли послідовність $\mathbf{P}^*$ стає порожньою.

Експериментальні дослідження запропонованих адаптивних методів ущільнення показали, що він забезпечують підвищення коефіцієнта ущільнення для різних типів файлів від 2 до 6 разів порівняно з неадаптивним методом ущільнення на основі лінійної форми Фібоначчі.

Висновки

Запропоновано нові методи ущільнення даних на основі лінійної форми Фібоначчі, які передбачають використання g+2 (g=0,1,2,…) моделей джерела даних, чотирьох правил кодування даних і функції оптимізації (адаптації). Вибір на основі функції оптимізації моделі джерела даних і правила кодування даних, які спільно забезпечують найбільший коефіцієнт ущільнення, сприяє адаптації виконуваних перетворень до конкретного змісту ущільнюваних даних.

Адаптація забезпечує підвищення коефіцієнта ущільнення порівняно з неадаптивним методом ущільнення на основі лінійної форми Фібоначчі, який передбачає тільки одну модель джерела даних й одне правило кодування для кожного блоку вхідних даних.

Література

1. Кричевский Р.Е. Сжатие и поиск информации / Р.Е. Кричевский. – М.: Радио и Связь, 1989. – 176 с.
2. Рябко Б.Я. Эффективный метод адаптивного арифметического кодирования для источников с большими алфавитами / Б.Я. Рябко, А.Н. Фионова // Проблемы передачи информации. – 1999. – Т. 35, вып. 4. – С. 34–39.
3. Rissanen J.J. Universal modeling and coding / J.J. Rissanen, G.G. Langdon // IEEE Trans. Inf. Theory. – 1981. – V. IT-27, 1. – P. 12–23.
4. Huffman D.A. A method for the construction of minimum redundancy codes / D.A. Huffman // Proceedings of the Institute of Electrical and Radio Engineers. – 1952. – V. 40, 9. – P. 1098–1101.

5. Cormack G.V. Data compression using dynamic Markov modeling / G.V. Cormack, R.N. Horspool // Comput. J. – 1987. – V. 30, 6. – P. 541–550.
6. Storer J. A. Data Compression: Methods and Theory / J. A. Storer //Computes Science Press. – 1988. – V. 47, 1. – P. 23–29.
7. Ziv J. Compression of individual sequences via variable-rate coding / J. Ziv, A. Lempel // IEEE Trans. Inf. Theory. – 1978. – V. IT-24, 5. – P. 530–536.
8. Welch T.A. A Technique for High Performance Data Compression / T.A. Welch // Computer/ – 1984. – №6. – P. 176--189.
9. Анисимов А.В. Обратное преобразование Фибоначчи / А.В. Анисимов, Я.П. Рындин, С.Е. Редько // Кибернетика. – 1982. – № 3. – С. 9–11.
10. Кшановський О.Д. Арифметичні методи ущільнення цифрової інформації / О.Д. Кшановський, С.В. Тітарчук, В.А. Лужецький //Вісник ВПІ. – 1999. – №5. – С. 83–87.