

УДОСКОНАЛЕНИЙ АДАПТИВНИЙ МЕТОД УЩІЛЬНЕННЯ ДАНИХ НА ОСНОВІ ВІДКИДАННЯ ПОСЛІДОВНОСТЕЙ НУЛІВ ТА ОДИНИЦЬ

Т. М. Чеборака, аспірант
Вінницький національний технічний університет
altamira_90@mail.ru

Серед великої чисельності відомих підходів та методів ущільнення інформації не існує універсального методу ущільнення, який би забезпечував однаково високий коефіцієнт ущільнення. Тому, актуальним є розробка адаптивних методів ущільнення інформації, які б ущільнювали дані з різними властивостями та характеристиками, оскільки саме від цих критеріїв залежить ступінь ущільнення даних.

Цей підхід базується на використанні методів відкидання послідовностей однакових символів у молодших, старших, внутрішніх або молодших і старших розрядах. Відмінність цього підходу полягає у зміні правил кодування.

Правило кодування (тип 0) застосовується, коли жоден із методів не зменшує обсяг вихідного блоку даних. З метою розпізнавання цих послідовностей, до початкового блоку дописується ознака перетворення p із значенням 0, що наведена на рис. 1.

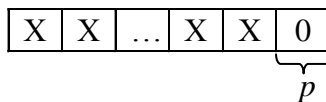


Рисунок 1 – Структура перетвореного блоку за правилом кодування (тип 0)

Правило кодування (тип 1), застосовується для блоків, що містять символи одного типу 0 або 1. Ущільнена послідовність, що наведена на рис. 2 містить такі поля:

- $p = 1$ – ознака перетворення;
- $type = 00$ – тип схеми відкидання;
- c – тип символу, що відкидається.

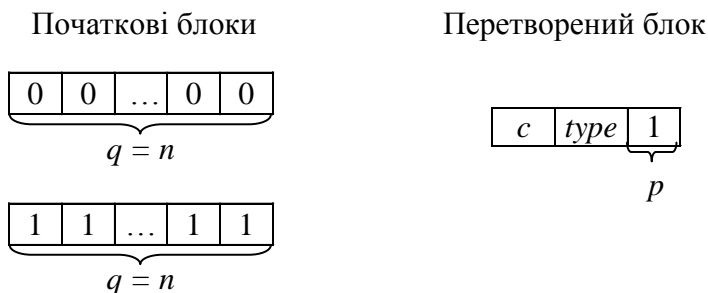


Рисунок 2 – Структури початкового та перетвореного блоку за правилом кодування (тип 1)

Правило кодування (тип 2) застосовується для блоків, що містять послідовність однакових символів на початку або в кінці послідовностей, при виконанні умови $q \geq \log_2 n + 4$. Ущільнена послідовність, що наведена на рис. 3 містить такі поля:

- $p = 1$;
- $type = 10$;
- q – двійковий код кількості однакових символів послідовності, що відкидається (розрядність коду $\log_2 n$);
- h – ознака відкидання послідовностей однакових символів у молодших ($h = 0$) або старших ($h = 1$) розрядах;
- c ;
- залишок – код, що залишається без змін (розрядність коду $n - q - 1$).

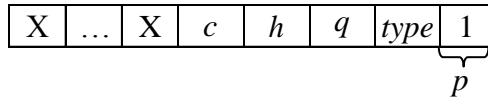


Рисунок 3 – Структура перетвореного блоку за правилом кодування (тип 2)

Правило кодування (тип 3) застосовується для блоків, що містять послідовність однакових символів у внутрішніх розрядах, у разі виконання умови $q \geq 2 \log_2 n + 2$. Ущільнена послідовність, що наведена на рис. 4 містить такі поля:

- $p = 1$;
- $type = 01$;
- q ;
- c ;
- l – двійковий код кількості символів, що залишається без змін у молодших розрядах (розрядність коду $\log_2 n$);
- залишок у молодших розрядах (розрядність l);
- залишок у старших розрядах – код, що залишається без змін (розрядність коду $n - q - l - 2$).

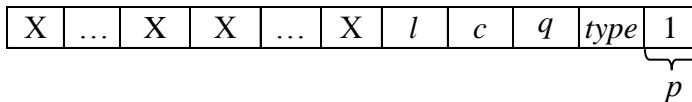


Рисунок 4 – Структура перетвореного блоку за правилом кодування (тип 3)

Правило кодування (тип 4) застосовується для блоків, що містять послідовність однакових символів у старших і молодших розрядах, у разі виконання умови

$$\begin{cases} q \geq \log_2 n + 5, \text{ якщо } q_l = q_h, \\ q \geq 2 \log_2 n + 5, \text{ якщо } q_l \neq q_h, \end{cases}$$

де $q = q_l + q_h$.

Ущільнена послідовність, що наведена на рис. 5 містить такі поля:

- $p = 1$;
- $type = 11$;
- q_{\min} розрядністю $\log_2 n$, в яке записується значення $\min(q_l, q_h)$;
- c_h – тип символу, що відкидається у старших розрядах;
- c_l – тип символу, що відкидається у молодших розрядах;
- t_e – тип нерівності q_l і q_h : $t_e = 01$ при $q_l > q_h$, $t_e = 10$ при $q_l < q_h$, $t_e = 11$ при $q_l = q_h$;
- Δq розрядністю $\log_2 n$, в яке записується значення $|q_l - q_h|$;
- залишок – код, що залишається без змін (розрядність коду $n - q - 2$).

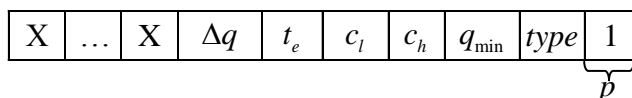


Рисунок 5 – Структура перетвореного блоку за правилом кодування (тип 4)

Перевагами запропонованого підходу є підвищення значення коефіцієнта ущільнення за рахунок аналізу вхідної послідовності даних та вибору найбільш ефективної схеми відкидання послідовностей нулів та одиниць.