

АДАПТИВНЕ УЩІЛЬНЕННЯ ДАНИХ

А. В. Кульчицький, аспірант
Вінницький національний технічний університет
andriyk88@gmail.com

Ущільнення даних – це алгоритмічне перетворення даних, яке проводиться задля зменшення їх об'єму.

Процес ущільнення даних умовно можна розділити на дві частини:

- моделювання;
- кодування.

Моделювання – це процес представлення вхідних даних у найбільш зручній для подальшої обробки формі.

Кодування інформації – це процес перетворення вхідної послідовності символів у вихідну, яка задовольняє певні критерії.

Головною вимогою, що висувається до процесу ущільнення, є те, що вихідна (результуюча) послідовність символів повинна бути меншою за вхідну. Якщо дана вимога не виконується, то це означає що ущільнення є не ефективним.

Процес моделювання методу, що розглядається, передбачає виконання наступних дій:

1. Представлення вхідних M даних як послідовності символів a_i , де $a_i \in \{0,1\}$.
2. Розбиття послідовності на блоки b_j довжиною l .
3. Кожному блоку b_j ставиться у відповідність число N_j :

$$N_j = f(b_j).$$

В результаті вхідні дані матимуть вигляд:

$$M^* = \{N_0, N_1, \dots, N_{k-1}\}.$$

Таке представлення вхідних даних називається числовою моделлю даних.

Процес кодування числової моделі передбачає кодування кожного її елемента. Кодування елемента полягає у розкладанні елемента за формулою:

$$N_i = k^a + R,$$

де k – додатнє дійсне число;

a - ціле число,

R – остача, яка дорівнює $R = N_i - k^a$.

Не всі дані однаково добре піддаються ущільненню. Тому для покращення результатів ущільнення потрібно проводити адаптацію до даних. На етапі моделювання адаптація проводиться шляхом введення додаткових перетворень над бітами в блоках даних, а на етапі кодування адаптація до вхідних даних проводиться такими способами:

1. Змінюючи a визначаємо k і R при загальна довжина a , k і R є найменшою.

2. Змінюючи k обчислюємо a і R при загальна довжина a , k і R є найменшою.

Усі вище зазначені модифікації процесу кодування числової моделі забезпечить значну адаптивність ущільнення.