

## **КЛАСИФІКАЦІЯ МЕТОДІВ УЩІЛЬНЕННЯ, ЩО БАЗУЮТЬСЯ НА ОБЧИСЛЕННІ ВІДХИЛЕНЬ**

**В.А. Лужецький, д.т.н. професор,  
В.А. Каплун, асистент  
Вінницький національний технічний університет  
e-mail: valuka@rambler.ru**

Основні відомі методи ущільнення використовують або статистичні характеристики даних, що ущільнюються, або словниковий принцип. Однак, сучасні мікропроцесори ефективніше здійснюють арифметичні операції над числами. Тому останнім часом здійснюються пошуки підходів щодо ущільнення даних, які базуються на їх представленні як цілих чисел. У доповіді розглядається класифікація методів ущільнення даних, основою яких є обчислення відхилень.

Набір даних, що підлягає ущільненню, можна охарактеризувати діапазоном представлення чисел заданої розрядності, а також такими статистичними параметрами, як мінімальне значення, середнє значення для усього набору або для груп чисел і накопичуване середнє значення для всього набору.

Виходячи з цього, методи ущільнення, що базуються на обчисленні відхилень, можна розділити на дві групи:

- ущільнення з урахуванням статистик;
- ущільнення без урахування статистик.

Серед методів, що здійснюють ущільнення з урахуванням статистичних характеристик, можна виділити такі.

1. Ущільнення, при якому здійснюється розбиття на піддіапазони. У цьому випадку в кожному виділеному піддіапазоні обчислюються відхилення або від центрів

діапазонів, або від середнього значення чисел, що належать даному діапазону. Ущільнена інформація зберігатиметься у такому вигляді: для кожного з відхилень запам'ятовуються знаковий розряд, кількість відкинутих повторюваних старших розрядів, ознака попадання у той або інший піддіапазон. Крім того, для однозначного відновлення первинного вигляду інформації необхідно зберігати самі середні значення чисел у піддіапазоні або числові значення центрів під діапазонів.

2. Ущільнення, при якому вхідний набір чисел розбиваються на групи з  $n$  чисел у кожній ( $n=2,3,\dots$ ), передбачають обрахування і зберігання або відхилень від середніх значень або б) від мінімальних значень у певній групі. При цьому також для кожного з відхилень зберігаються молодші розряди чисел після відкидання певної кількості повторюваних розрядів і кількість цих відкинутих розрядів. Крім того, у першому випадку слід зберігати знаковий розряд і середнє значення чисел у групі.

3. Ущільнення з урахуванням накопичуваного середнього значення, при якому, зафіксувавши перше число з вхідного потоку даних, обчислюються і зберігаються молодші розряди відхилень від накопичуваного середнього значення. При цьому, окрім самих значень відхилень, необхідно зберігати знаковий розряд і кількість відкинутих повторюваних розрядів.

Серед групи методів, що не передбачають знаходження статистичних характеристик вхідних даних, можна виділити такі: а) ущільнення, основане на зберіганні відхилень від сусіднього елемента; б) відхилення від констант; в) відхилення від відповідних значень апроксимуючої функції. В усіх цих випадках, крім самих відхилень, необхідно зберігати і додаткову інформацію для однозначного подальшого відновлення первинної інформації.