УДК 004.627(043.2)

**Соломяний О.Є.**

*Національний авіаційний університет,**Київ*

**ПОРІВНЯЛЬНИЙ АНАЛІЗ АЛГОРИТМІВ СТИСНЕННЯ ДАНИХ**

Стиснення даних – це процедура подання даних у такій формі, яка забезпечить зменшення їх розміру (обсягу). Стиснення використовується при локальному зберіганні інформації (наприклад, архівації), а також при передачі даних по каналах зв’язку, в тому числі і в мережі Інтернет. При архівації, файли в архіві можуть бути як стиснені (без втрат), так і мати початковий розмір та структуру, але першочергове завдання архіву тримати у собі саме стиснуті файли. Метадані можуть містити інформацію про початковий розмір файлів, інформацію про формат файлів, структуру директорій, коментарі до файлів, інформацію для відновлення архіву і т. д.

Стиснення буває без втрат (коли можливе відновлення вихідних даних без спотворень) або з втратами (відновлення можливе з незначними спотвореннями). Стиснення без втрат використовується при обробці та збереженні комп'ютерних програм і даних. Стиснення з втратами зазвичай застосовується для зменшення об'єму звукової, фото, та відеоінформації. І, як показує практика, стиснення з втратами для такого роду інформації є цілком прийнятним.

Існує багато практичних алгоритмів стиснення даних, але всі вони базуються на трьох теоретичних способах зменшення надлишковості даних. Перший спосіб полягає в зміні вмісту даних, другий – у зміні структури даних, а третій – в одночасній зміні як структури, так і вмісту даних. А також існує багато різних практичних методів стиснення без втрати інформації, які, як правило, мають різну ефективність для різних типів даних та різних обсягів. Однак, в основі цих методів лежать три теоретичних алгоритми:

• алгоритм RLE (Run Length Encoding);

• алгоритми групи KWE (KeyWord Encoding);

• алгоритм Хафмана.

В основі алгоритму RLE лежить ідея виявлення послідовностей даних, що повторюються, та заміни цих послідовностей більш простою структурою, в якій вказується код даних та коефіцієнт повторення. В основі алгоритму стиснення за ключовими словами покладено принцип кодування лексичних одиниць групами байт фіксованої довжини. В основі алгоритму Хафмана лежить ідея кодування бітовими групами. Спочатку проводиться частотний аналіз вхідної послідовності даних, тобто встановлюється частота входження кожного символу, що зустрічається у ній. Після цього символи сортуються по спаданню частоти входження. Основна ідея полягає в наступному: чим частіше зустрічається символ, тим меншою кількістю біт він кодується.

*Науковий керівник – Глазок О.М., канд. техн. наук, доцент.*