

Лабораторна робота №4

Архіватори

Мета роботи: ознайомитися з методами архівації і компресії інформації, сучасними програмами – архіваторами та їх ефективністю відносно різних типів файлів.

Загальна інформація. По мірі розвитку комп'ютерних систем, збільшення об'ємів інформації для користування та необхідності передачі її по локальним і глобальним мережам з'явилася й потреба в стисненні різних даних. Стиснення зменшує об'єм дискового простору, необхідного для зберігання файлів в комп'ютері, і часу, необхідного для передачі інформації мережею.

Стиснення файлів здійснюється за допомогою спеціальних програм - архіваторів. Прийнято розрізняти архівацію і компресію даних. При архівації відбувається злиття декількох файлів або каталогів у єдиний файл - архів, а при компресії - скорочення обсягу вихідних файлів шляхом установаження й видалення надмірності інформації. Ступінь надмірності залежить від типу даних. Наприклад, для відео-файлів ступінь надмірності в кілька разів більше, ніж у графічних даних, а ступінь надмірності графічних даних, у кілька разів більше, ніж текстових. Ступінь надмірності залежить також від прийнятої системи кодування. При обробці інформації надмірність також відіграє важливу роль, вона підвищує її якість. Але коли мова заходить не про обробку, а про зберігання готових документів або їхньої передачу, то надмірність можна зменшувати, що дає ефект стиснення даних. Як правило, сучасні програми - архіватори забезпечують і стиснення і компресію даних.

Принцип архівації (алгоритм) полягає у наступному: збережена у файлі інформація представлена у вигляді двійникового коду - послідовності одиниць і нулів. Програма- архіватор переглядає файл і аналізує його на наявність повторюваних кодових комбінацій. При виявленні таких комбінацій кожен повторюваний сегмент вихідного файлу замінюється коротшим кодом архівного файлу з використанням посилання на перший з повторюваних фрагментів (сукупність цих посилань і таких початкових сегментів звичайно називається словником).

Програми - архіватори розрізняються за способом визначення повторюваних комбінацій вихідного коду й формату створюваних архівних файлів. Найбільш ефективним вважається той архіватор, що найбільшою мірою зменшує розмір вихідного файлу. Математична складність цього завдання пов'язана з комбінаторним перебором дуже великої кількості варіантів.

Всі алгоритми (методи) стиснення даних можна розділити на дві групи - стиснення без втрат і з втратами. Перший тип алгоритмів використовується в тих випадках, коли об'єм вихідних даних має повністю співпадати з об'ємом файлів, відновлених після стиснення. На алгоритмах цієї групи ґрунтуються програми - архіватори й протоколи передачі даних у комп'ютерних мережах. Метод архівації без втрат використовується при стисненні системних файлів (*.exe) та файлів, що передаються глобальною мережею (*.html). Другий тип алгоритмів стиснення використовується у випадках, коли користувачеві не потрібне співпадання вхідної й обробленої інформації. Стиснення даних із

втратами частини інформації забезпечує найвищий ступінь і швидкість стиснення даних. Стиснення з втратами допускається при архівації файлів з форматами *.jpg, *.mpeg.

Найбільше широко відомим архіватором для Windows є WinRAR.

Архіватор WinRAR. Для запуску програми WinRAR виконується одна з наступних дій:

1. У меню Пуск виберіть пункт Програми → WinRAR;
2. Двічі натисніть на ярлику WinRAR робочого столу Windows;
3. Натисніть на кнопці WinRAR панелі швидкого запуску панелі завдань Windows.

На екрані розкриється головне вікно програми.

До найбільш зручних можливостей WinRAR можна віднести наступні особливості:

- створення багатотомних архівів, як звичайних, так і таких що саморозпаковуються;
- підтримка великої кількості різних форматів архівних файлів;
- можливість використати в якості основного один з форматів архівних файлів RAR або ZIP;
- інтеграція команд WinRAR в інтерфейс Windows;
- можливість швидкого пошуку файлів на обраному диску, не залишаючи вікна WinRAR, як у звичайній папці, з використанням звичних засобів навігації;
- наявність меню Історія призначена для швидкого повернення до попередніх архівів, з якими працював користувач найближчим часом.

Для перегляду архівів є вбудований модуль візуалізації, але можна підключити й зовнішню утиліту.

У режимі мультимедіа – стиснення WinRAR намагається використати додатковий алгоритм для поліпшення стиснення цифрового звуку або повнокольорового зображення. У цьому випадку може бути досягнуто на 30% більш високий ступінь стиснення, ніж при звичайному методі архівації.

Використання методу безперервної архівації (solid archive) може на 10 - 15% збільшити ступінь стиснення при впакуванні великої кількості однотипних файлів невеликого розміру. Це досягається за рахунок створення загального для всіх файлів словника. Однак даний метод має й свої недоліки - більший час архівації /розархівації, при ушкодженні даного типу архівів відновлення інформації практично неможливе.

Для формату RAR підтримується створення так званого формату recovery record - в архів записується надлишкова інформація, що дозволяє відновити дані з архіву навіть при його частковому ушкодженні (найчастіше внаслідок ушкодження фізичного носія інформації). Однак хоча додавання цієї інформації значно збільшує надійність зберігання/передачі інформації, обсяг архіву при цьому також збільшується.

Цілісність архіву визначається за допомогою CRC-коду циклічного контролю, що розраховується для кожного заархівованого файлу окремо.

Крім того, у форматі RAR реалізована підтримка багатотомних архівів (multi-volume). Багатотомний архів являє собою розбитий на окремі файли архів великого обсягу й необхідний для транспортування великого обсягу інформації

за допомогою носіїв інформації з меншим обсягом або для передачі файлів великого об'єму через інтернет.

У WinRar можна створювати архіви, що самостійно розпаковуються – так звані SFX-архіви. Вони являють собою архів з вбудованою програмою - распаковщиком. Архів може бути багатотомним і безперервним. Має розширення *.exe. Такий метод архівації призначений для перенесення даних на комп'ютери без встановленого архіватора.

ПОРЯДОК ВИКОНАННЯ РОБОТИ:

1. Дати відповіді на контрольні запитання.
2. У папці «User» створити папку Winrar_archive.
3. У дану папку записати вказані викладачем файли:
 - рисунок - формат *.bmp,
 - рисунок - формат *.jpg,
 - текстовий файл *.doc,
 - програмний файл *.exe.
4. Для кожного з вищевказаних типів файлів (*.bmp, *.jpg, *.doc, *.exe) провести архівацію різними методами та занести в таблиці отримані результати (див. табл.1).

Таблиця 1

Метод архівації	Розмір, байт	Ступінь стиснення, %
Початковий розмір		
(Winrar) zip швидкий (fast)		
(Winrar) rar швидкий (fast)		
(Winrar) zip звичайний (normal)		
(Winrar) rar звичайний (normal)		
(Winrar) zip максимальний (best)		
(Winrar) rar максимальний (best)		
(Winrar) zip максимальний SFX (архів, що самостійно розпаковується)		
(Winrar) rar максимальний SFX (архів, що самостійно розпаковується)		
(Winrar) rar максимальний (best) + інформація для відновлення (put recovery record)		

5. За даними таблиць побудувати з допомогою редактора Excel графіки (гістограми) – залежності ступеня стиснення файлу від методу архівації – для кожного з типів файлів.

6. Створити багатотомний архів всієї папки Winrar_archive (split to volumes), задавши розмір тому 500 кБ і метод архівації гаг максимальний (best); записати кількість і назву томів.

7. Зробити висновки про ефективність архівації різних типів файлів з допомогою програми Winrar.

КОНТРОЛЬНІ ПИТАННЯ:

1. Призначення програм-архіваторів.
2. Методика (алгоритм) стиснення інформації.
3. Стиснення без втрат і із втратами. Область застосування.
4. Можливості програми WinRAR.
5. SFX-архів.
6. Мультимедіа-стиснення.
7. Безперервний архів.
8. Багатотомний архів: створення й область застосування.