

СИСТЕМА КОМПРЕСІЇ ДАНИХ ДЛЯ WEB

Проблема передачі великих обсягів даних є актуальною в умовах інтенсивного розвитку Інтернету (зокрема соціальних мереж та сценарно-складних сайтів, кількість яких неухильно зростає). Таким чином, дослідження в області компресії даних різних типів, є актуальним завданням, вирішення якого дозволяє підвищити ефективність використання пропускну здатності мереж без істотних матеріальних витрат.

Завдання дослідження:

1. Дослідити існуючі програми компресії даних, аналізуючи їх структуру, ефективність і швидкість стиснення, вибрати оптимальний метод компресії високонадлишкових даних і графічних файлів, розробити систему стиснення даних для Web на основі власної варіації обраного методу компресії та порівняти її з існуючими аналогами. Вивчити найближчі аналоги системи стиснення даних (архіватори і трафік-компресори).

2. Проаналізувати існуючі універсальні алгоритми компресії без втрат і методи контекстного моделювання, а також їх основні характеристики, що впливають на швидкість і ефективність стиснення.

Результати дослідження:

1. У процесі роботи над проектом були проаналізовані існуючі універсальні алгоритми стиснення без втрат (різні варіанти арифметичного кодування такі як інтервальне кодування і адаптивне арифметичне кодування) і методи контекстного моделювання, а також їх основні характеристики, що впливають на швидкість і ефективність стиснення, зокрема проведено аналіз методів оцінки ймовірностей утечі та способів підвищення точності оцінок в контекстних моделях високих порядків. На підставі проведених досліджень було з'ясовано, що алгоритми передбачення за частковим збігом (окремий випадок контекстного моделювання) є оптимальною по співвідношенню ефективність/швидкість компресії основою для універсальних систем компресії даних.

2. Дослідження виявили, що оптимальними порядками моделі для текстів на природних мовах є порядки 4, 5 і 6. Тому для високонадлишкових даних була створена власна варіація алгоритму передбачення за частковим збігом.

3. Для малонадлишкових даних (графічних файлів) на основі досліджень їх внутрішньої структури був розроблений власний алгоритм

стиснення без втрат, що ґрунтується на таких базових методах стиснення, як символільне подавлення, заміна і відносне кодування.

Висновки:

1. При використанні системи компресії даних, обсяг переданих по мережі даних скорочується в середньому на 80% у разі текстової інформації, на 70% для виконуваних файлів, на більш ніж 90% для баз MS Access і на 60% в разі невеликих зображень, і все це при прийнятній швидкості компресії/декомпресії.

2. Система компресії даних не має жорстких вимог до програмного і апаратного забезпечення, займає невеликий об'єм і проста у використанні. Програма може бути модифікована і вдосконалена шляхом додавання нових методів компресії та оптимізації існуючих.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео / Ватолин Д., Ратушняк А., Смирнов М., Юкин В. – М.: ДИАЛОГ-МИФИ, 2003. – 384 с.

2. Вернер М. Основы кодирования. Учебник для ВУЗов. / М. Вернер – М.: Техносфера, 2004. – 288 с.

3. Сэломон Д. Сжатие данных, изображений и звука / Д. Сэломон – М.: Техносфера, 2004. – 368 с.

4. Пахомов С. Сравнение 64-битных архиваторов / С. Пахомов // Компьютер Пресс, – 2013 – №3 (279) – С. 52–64.

5. Кириченко В.В. Аналитический обзор алгоритмов сжатия цифровой информации / В.В. Кириченко // ПФМТ – 2016 – №2 (27) – С. 77–83.