

РЕФЕРАТ

Актуальність теми.

Кожного дня кількість інформації, що зберігається, передається та обробляється в комп'ютерних системах зростає. Прогрес в галузі технічних засобів для зберігання та передачі інформації не стоїть на місці, проте зі збільшенням розміру даних суттєво збільшується і їх вартість.

Для ефективного використання дискового простору було реалізовано методи для архівування (ущільнення) даних без втрат інформації. Такі методи дозволяють зменшувати розмір інформації для зберігання або передачі по каналах зв'язку. Очевидна економічна вигода оптимізації подання інформації та актуальність підвищення ефективності методів ущільнення даних.

Головне призначення програм-архіваторів – ущільнення файлів з метою економії пам'яті. Оскільки з ущільненими файлами часто неможливо працювати за їх прямим призначенням, їх використовують для зберігання, резервного копіювання або обміну інформацією через мережу Інтернет.

Основний принцип алгоритмів ущільнення базується на тому, що в будь-якому файлі, який містить не випадкові дані, інформація частково повторюється. Використовуючи статистичні математичні моделі можна визначити ймовірність повторення певної комбінації символів. Після чого можна створити коди, що позначають обрані фрази, і фразам, які повторюються найчастіше призначити найкоротші коди. Для цього використовують ентропійне кодування, кодування повторів і ущільнення за допомогою словника. З їх допомогою 8-бітний символ, або цілий рядок можуть бути замінені лише декількома бітами, усуваючи таким чином зайву інформацію.

Об'єктом дослідження є технологія ущільнення даних.

Предметом дослідження є оцінка використання підстановок довільної розрядності, які реалізуються автоматами Мілі, для підвищення ефективності ущільнення даних.

Мета роботи. Перевірка багаторозрядних підстановок на базі автоматів Мілі для підвищення ефективності ущільнення даних та порівняння методів вибору ефективних підстановок для перетворень.

Наукова новизна даної роботи полягає в модифікації методу ущільнення даних за допомогою підстановок довільної розрядності зокрема шляхом реалізації багаторозрядних підстановок на автоматах Мілі

Практична цінність отриманих в роботі результатів полягає в тому, що запропонований метод дозволяє вдосконалити та підвищити ефективність існуючих засобів ущільнення даних.

Апробація роботи.

Результати проміжних досліджень були представлені на: X науковій конференції магістрантів та аспірантів «Прикладна математика та комп'ютеринг» ПМК-2019-2 (Київ, 13-15 листопада 2019р.) та Міжнародній науково-технічній Internet-конференції (Київ, 11 листопада 2019р.).

Структура та обсяг роботи.

Магістерська дисертація складається з вступу, чотирьох розділів, висновків, що містять основні результати, списку літератури та додатків.

У вступі сформульована задача дослідження, обґрунтована її актуальність, визначена мета роботи, зазначені її наукова новизна та практичне значення.

У першому розділі розглянуто існуючі методи ущільнення даних та область їх застосування.

У другому розділі способи реалізації прямих та обернених багаторозрядних підстановок для модифікації файлів.

У третьому розділі розглянуто програмну реалізацію багаторозрядних підстановок на базі автоматів Мілі.

У четвертому розділі розглянуто методи ущільнення даних на базі підстановок довільної розрядності, приведено результати експериментальних досліджень та порівняльну характеристику існуючих і запропонованого методів ущільнення даних без втрат інформації.

У висновках приведено підсумок за результатами проведеної роботи.

Ключові слова: ущільнення даних, автомат Мілі, багаторозрядні підстановки, ущільнення без втрат, бієктивні відображення.

ABSTRACT

Theme urgency. The amount of information stored, transmitted and processed on computer systems is increasing every day. The advances in the storage and transmission of technical equipment are not in place, but with the increase in the size of the data, their value is significantly increasing.

To use disk space efficiently, methods have been implemented for archiving (compacting) data without loss of information. These methods allow you to reduce the size of information for storage or transmission through communication channels. The obvious economic benefit of optimizing the submission of information and the relevance of improving the efficiency of data compaction methods.

The main purpose of the program-archiver is to file files with the help of the memory. Because compacted files are often impossible to handle for their intended purpose, they are used to store, back up, or share information over the Internet.

The basic principle of the algorithm is based on the fact that in any file, there is no need for data, information is often repeated. Using statistical mathematical models, you can determine the likelihood of a certain combination of characters. Whenever you can change the code, you can choose the phrases, phrases, and phrases that are repeated, most often, the most important short codes. For the whole entropy of entropy code, the code of repetition and repetition for an auxiliary vocabulary. With additional 8-bit symbol, because the number of rows can be better than the lower decile of the bits, assuming such an order I will give information.

Object of research are data compaction technology.

Subject of research are is an estimate of the use of arbitrary bit substitutions implemented by Mealy machines to improve data compaction efficiency.

Research objective: checking multi-digit substitutions based on Mealy machines to improve data compaction efficiency and compare methods for choosing effective conversion substitutions.

Scientific novelty of this work lies in the further development of the method of data compaction with the help of arbitrary bit substitutions, in particular, by implementing multi-bit substitutions on the Mealy machine.

Practical value of the results obtained is that the proposed method makes it possible to improve and increase the efficiency of existing data sealers.

Approbation. The intermediate results of the studies were presented at the IX Scientific Conference of PhD students and undergraduates «Applied mathematics and computing» AMC-2019-2 (Kyiv, 13-15 November 2017) and International Scientific and Technical Internet Conference (Kyiv, November 11, 2019).

Structure and content of the thesis.

Master's thesis consists of an introduction, four sections, conclusions containing the main results, a list of references and appendices.

The introduction presents the task of the study, substantiates its relevance, defines the purpose of the work, indicates its scientific novelty and practical significance.

The first section describes the existing data compaction methods and their scope.

The second section describes how to implement direct and inverted multi-bit file modifications.

The third section presents the software implementation of multi-bit substitutes based on Mealy machines.

The fourth section describes methods of data compilation on the basis of substitutions of arbitrary bits, presents the results of experimental studies and comparative characteristics of existing and proposed methods of data compaction without loss of information.

In Conclusion, summed up the results of work.

Keywords: data compression, Mealy machine, multi-bit substitutions, lossless compression, bi-active reflections.