

ВИКОРИСТАННЯ АЛГОРИТМІВ НЕЛІНІЙНОГО СТИСНЕННЯ В СИСТЕМАХ ЗБЕРІГАННЯ ДАНИХ

Вінницький національний технічний університет

Анотація

Запропоновано використання нелінійного алгоритму стиснення даних, а саме адаптивного алгоритму Хаффмана, який надає можливість проводити процес стиснення даних, стиснення повідомлення в runtime-режимі, що підвищить швидкість роботи веб-серверу.

Ключові слова: нелінійне стиснення, адаптивний алгоритм Хаффмана, система зберігання даних, компресор, декомпресор.

Abstract

The use of the nonlinear data compression algorithm, namely the adaptive Huffman algorithm, which provides the ability to perform data compression, compression of the message in the runtime mode, which will increase the speed of the operation of the web server, is proposed.

Keywords: nonlinear compression, adaptive Huffman algorithm, data storage system, compressor, decompressor.

Більшість алгоритмів стиснення, що використовуються в системах зберігання даних на сьогоднішній день, базуються на моделі послідовної обробки. Системи зберігання даних, щоб мати можливість розпакувати блоки поза чергою або паралельно, повинні скинути стан стиснення перед кожним блоком, зменшуючи адаптивність та обмежуючи коефіцієнти стиснення. Означений недолік можна усунути за допомогою використання нелінійного стиснення даних, що дозволить системам зберігання даних накладати довільний частковий порядок на міжблочні залежності. Взаємно неупорядковані блоки можуть бути стиснуті і розпаковані поза порядком або паралельно, і компресор буде мати можливість адаптивно стискати кожен блок на основі попередніх блоків. Ця структура повністю фіксує залежності системи даних, що дозволяє компресору адаптуватися за допомогою довгострокового стану без обмеження послідовної обробки. Попередній досвід роботи з алгоритмом стиснення Хаффмана показує, що нелінійне стиснення підходить для багатьох систем зберігання даних [1].

Для усунення означених недоліків доцільно використати нелінійні алгоритми стиснення даних (NLC). Оскільки основним недоліком більшості алгоритмів даних є поточний вибір «все-чи-нічого» між отриманням стану стиснення даного блоку даних з «усіх» попередніх даних у лінійній послідовності, або без «жодної» попередньої інформації, програма вказує на NLC компресор в явному вигляді через будь-яку пряму ациклічного графу (DAG), який був раніше стиснутий. Декомпресор NLC також припускає, що перед тим як розпакувати даний блок, програма вже повинна розпакувати "попередні" блоки, які визначаються як залежності. В результаті, декомпресії нерівних блоків в DAG є незалежними процесами, а тому можуть бути виконаними паралельно. Таким чином, DAG накладає часткове обмеження на блоки даних як на DAG-вузли. Невпорядковані вузли і зв'язані дані з ними можуть оброблятися повністю й незалежно, під час стиснення або декомпресії. Представлення DAG забезпечує дві основні переваги: нелінійний компресор пропонує одну абстракцію для додатків з декількома логічними потоками, незалежні гілки можуть оброблятися одночасно. На рисунку 1 представлено порівняння методів стиснення даних.

Якість нелінійного стиснення даних, може бути покращена за допомогою використання адаптивного алгоритму Хаффмана [2]. У алгоритмі адаптивного кодування Хаффмана найбільші складності виникають при розробці процедури поновлення моделі чергових символів. Усунути такий недолік може модифікація H-дерева, що відобразить обробку нового символу. Адаптивне стиснення дозволить не передавати модель повідомлення разом з ним самим і обмежитися одним проходом по

повідомленню як при кодуванні, так і при декодуванні. Схема алгоритму адаптивного стиснення Хаффмана наведено на рисунку 2.

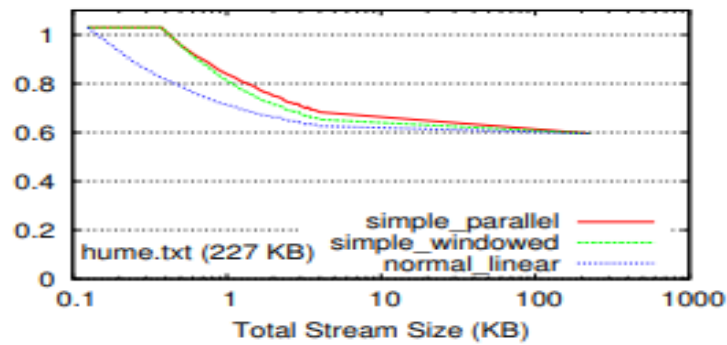


Рисунок 1 – Порівняння методів стиснення даних

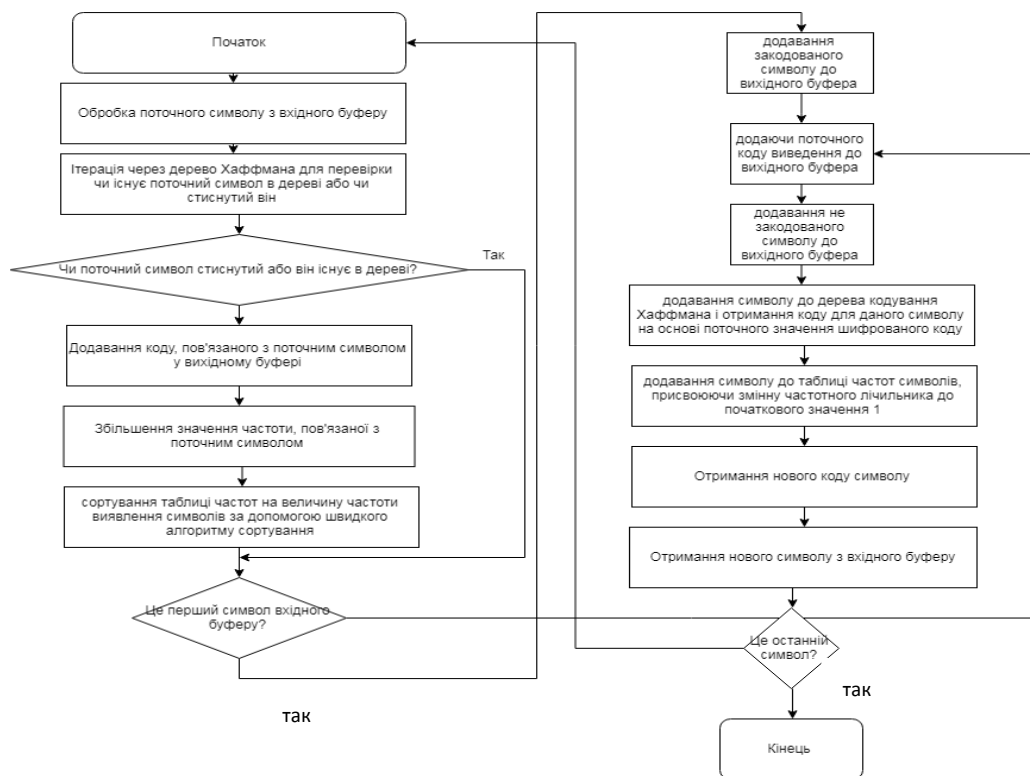


Рисунок 2 – Схема алгоритму адаптивного стиснення Хаффмана

Отже, за рахунок використання нелінійного алгоритму стиснення даних, а саме адаптивного алгоритму Хаффмана, який надає можливість проводити процес стиснення даних, стиснення повідомлення в runtime-режимі, може бути досягнуто підвищення швидкодії роботи веб-сервера.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. D. A. Lelewer and D. S. Hirschberg. Data compression. ACM Computing Surveys, 19:261–296, September 1987.
2. D. Huffman. A method for the construction of minimum-redundancy codes. Proceedings of the IRE, 40(9):1098–1101, Sept. 1952.

Савчук Тамара Олександрівна – PhD, професор, заступник завідувача кафедри комп'ютерних наук, Вінницький національний технічний університет, м. Вінниця.

Горобець Юрій Володимирович — студент кафедри комп'ютерних наук ВНТУ, Вінницький національний технічний університет, м. Вінниця, e-mail: yuriy.sparrow@gmail.com

Savchuk Tamara Oleksandrivna – PhD, Professor, Deputy Head of the Department of Computer Science, Vinnytsia National Technical University, Vinnytsia

Horobets Yurii Volodymyrovich — student of the Computer Sciences Chair, Vinnytsia National Technical University, Vinnytsia, e-mail: yuriy.sparrow@gmail.com