SnappyR: A New High-Speed Lossless Data Compression Algorithm

Rui Chen*, Lihao Xu†

*Mathworks Inc.
Natick, MA, 01760.
ruichen@mathworks.com

†Wayne State University
Detroit, MI, 48202
lihao@wayne.edu

We propose a high-speed lossless data compression algorithm, named SnappyR. Improved upon Snappy, we design new structures of the *literal* and the *match* tokens to achieve better compression ratio than Snappy. Numerous benchmarks are conducted on different sets of data corpus. The evaluations consistently show that SnappyR provides a better compression ratio comparing to Snappy, as well as LZ4, and better than LZO in most cases. Although a little slower than Snappy, the compression and decompression speeds of SnappyR are still much higher than entropy encoding based compression algorithms, such as ZSTD, deflate or Zlib. Thus, SnappyR can become another viable replacement or alternative to Snappy, LZ4 or LZO for computing and storage systems and applications, where high-speed lossless data compression is needed. More details about SnappyR algorithm design and performance evaluation can be found at [1].
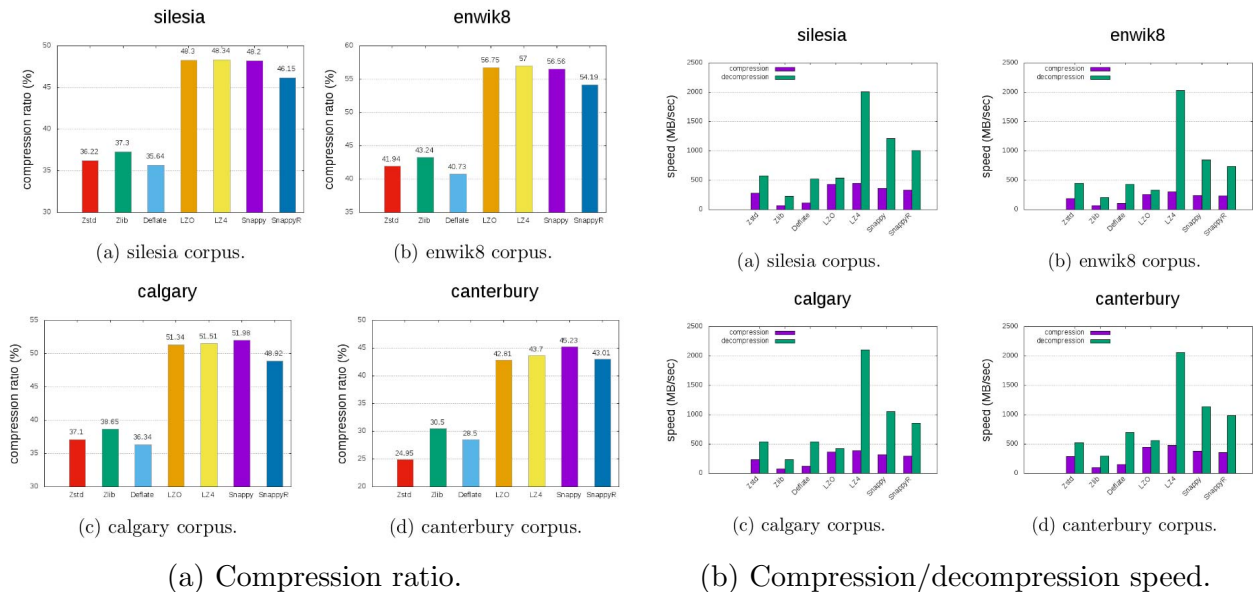
(a) silesia corpus.

(b) enwik8 corpus.

(c) calgary corpus.

(d) canterbury corpus.

(a) Compression ratio.

(a) silesia corpus.

(b) enwik8 corpus.

(c) calgary corpus.

(d) canterbury corpus.

(b) Compression/decompression speed.

Figure 1: SnappyR performance comparison with *blocksize=64KB*.

[1] Rui Chen Lihao Xu. Snappyr: A new high-speed lossless data compression algorithm. *Available: https://lihaoxu.eng.wayne.edu/NISL/Papers/Tech/SnappyR-DCC.pdf*, 2022.