Корпорацією Google був заупущений в 2013 році алгоритм Zopfli який використовував звичайний формат стискання даних Deflate.В той час як Zopfli є загальноприйнятним алгоритмом та форматом стискання даних відбулася спроба покращити алгоритм який працював з більшим вдношенням стискання та швидкістю стискання на файлах достатньо великого розміру.

В порівнянні були проведені тести на 22 бітному розсувному вікні для brotli.Алгоритми LZMA та LZHAM так використовували 22 бітне вікно,та 15 бітне для Deflate та Zopfli.Більші відношення дозволяють отримувати краще відношення стискання при втраті швидкості декодування.Версії програм наведені в додатку .

Тестовим комп’ютером мА процесор Intel Xeon CPU-1650(3.5 Ghz)-шестиядерний з шістьма додатковими віртуальними потоками.ОС:Linux 3.13.0.

Тип тестових даних – Кантерберійський корпус стискання даних та власноручно створений корпус з веб контенту з 1 285 файлів .Середній розмір файлу – 55 кілобайт.Було оцінено відношення стискання ,швидкість стискання та відновлення .Відмінністю від інших алгоритмів є те ,що Brotli використовує статичний словник.Він містит 13 504 слів та складів які мають найвищі частоти в англійській ,іспанській,китайській,ініді,російській та арабській.Також були використанні слова які найчастіше зустрічають в веб-документах на кшталт HTML ,XML,JS,CSS файлів.Для цих слів були застосовані перетворення аби охопити більшу частину можливих стрічок в документах. Загалом наявно понад 1.6 млн послідовностей.В експерименті застосовували корпус де були наявні документи з 93 мовами.Для оцінки використовувалося геометричне середнє значення в якому одне вимірювання ке було б надзвичайно далеко від всіх інших результатів не впливало би так сильно на загальний результат.

Результати

Таблиці 1,2 вказуюьь на результати в різних корпусах.Рисунок 1 показує що швидкість відновлення та відношення стискання на Кантерберійському корпусі.Ми можемо побачити що алгоритм Brotli при виставленому параметру якості(quality) 1 стискає та відновлює приблизно з тією ж швидкістю що й deflate.Але,він надає краще відношення стискання приблизно на 12-16 %.За параметру 9 Brotli подібний до Deflate 9.Але мА швидший на 28% в відновленні enwik8 ,з збільшенням відношення стискання на 13-21 %.Brotli 11 значно швидший в стисканні за Zopfli та має на 20 -26 % вище відношення стискання .

Brotli швидше відновлює дані аніж Deflate для тестових корпусів.В той час як як алгоритми LZMA та LZHAM значно повільніші за deflate.Геометричне середнє для швидкостей відновлення дорівнює 342.МБ\с для Brotli,та 323.6 МБ\с для deflate .

Стосовно стикання ,то brotli 1 швидший на 6 % за deflate 1.Brotli 9 повільніший на 32.3 % за deflate 9.Слід також враховувати не тільки параметри а й відношення стискання .При цьому ,відношення стискання для brotli 1 ближче до deflate 9 і навіть перевищує його .Нааприклад,приставанні Кантерберійського корпусу при відношенні 3.3 broltli 1 стискає при швидкості 98.3 МБ\с., а deflate -15.5МБ\с.

LZMA стискає третій корпус (enwik8) з відношенням на 2.5 %. Вище,але в втричі довше відновлює документ.

Висновки

Даний експеримент був проведений з 22 бітним розсувним вікном (за замовчуванням в Brotli). Алгоритми,такі як LZMA та LZHAM як правило вживають з більшим розсувним вікном.

Brotli використовує статичний словник для швидшого стискання коротших файлів . Інші алгоритми також можуть бути модифіковані таким чином,аби отримати дещо кращі відношення стискання .

Зокрема,слід взяти до уваги що Кантерберійський корпус містить в собі в основному короткі документи англійською,і відповідно значна частина слів були одразу закодовані через їхіні відповідники в словнику .

Згідно цих результатів лише Brotli може бути використаний в якості заміни для алгоритму Deflate через кращі показники такі як відношення стискання ,швидкість стискання ,швидкість відновлення .