

Abstraksi

Information Retrieval (IR) merupakan bagian dari *computer science* yang berhubungan dengan pengambilan informasi dari dokumen-dokumen yang didasarkan pada isi dan konteks dari dokumen-dokumen itu sendiri. Proses dalam *Information Retrieval* dapat digambarkan sebagai sebuah proses untuk mendapatkan *relevant documents* dari *collection documents* melalui pencarian *query* yang diinputkan *user*. Berbagai pendekatan untuk meningkatkan performansi *Information Retrieval (IR)* telah dilakukan. Salah satu cara untuk meningkatkan performansi tersebut adalah dengan kompresi indeks (*index compression*). Dua jenis teknik kompresi pada *Information Retrieval System* yaitu *lossy compression* dan *lossless compression*. Pada *lossless compression* semua informasi tetap terjaga, contohnya *posting file compression*.

Indeks merupakan bagian paling utama dalam *Information Retrieval System*. Tujuan menyimpan indeks adalah untuk mengoptimalkan kecepatan dan kinerja dalam menemukan dokumen yang relevan untuk permintaan pencarian. Tanpa indeks, mesin pencari akan memindai setiap dokumen, yang akan memerlukan banyak waktu dan daya komputasi. Namun, semakin banyaknya dokumen yang masuk semakin bertambah juga kapasitas indeks. *Index Compression* adalah teknik yang digunakan untuk lebih mengefisienkan indeks, baik dari kapasitas maupun performansi dari *Information Retrieval System*. Dengan mengkompresi indeks, dapat mengurangi kapasitas yang digunakan hingga 75%. *Index Compression* juga dapat meningkatkan kecepatan transfer dari disk ke memori.

Dalam tugas akhir ini, akan dilakukan implementasi *Index Compression* menggunakan *Variable Byte Code*. *Variable byte code* merupakan salah satu teknik dalam kompresi indeks yang diterapkan pada *Information Retrieval* guna mengurangi kapasitas disk yang terpakai dan pemakaian waktu pencarian yang lebih cepat. Oleh karena itu Diharapkan setelah menggunakan *Index Compression* menggunakan *Variable Byte Code*, kapasitas indeks akan berkurang dan performansi dari *Information Retrieval System* meningkat.

Kata Kunci: *Information Retrieval, Information Retrieval System, Index Compression, lossless compression, Variable Byte Code.*