

Bab I

Pendahuluan

1.1 Latar Belakang

Mesin pencari atau *Search Engine* merupakan alat yang sangat berguna seiring dengan perkembangan *World Wide Web* (WWW) sebagai salah satu sumber informasi yang mudah diakses. Dokumen akan di telusuri dan di proses oleh komponen yang dimiliki *search engine* dengan tujuan membentuk *database index* yang mempermudah pengguna *search engine* untuk mencari informasi yang diinginkan. Komponen utama yang dimiliki oleh *search engine* untuk mendukung layanan yang dimilikinya, yaitu: *web crawler*, *indexing system*, dan *searching system*[3].

Web Crawler bertugas untuk menjelajahi dan mengumpulkan semua informasi yang ada di dalam halaman web. *Mobile crawler* adalah *crawler* yang memanfaatkan peran *mobile agent* dalam pelaksanaan tugasnya. *Mobile agent* akan bermigrasi ke sumber informasi, membawa keseluruhan kode program dan melakukan operasi penjelajahan halaman web dan pengunduhan terhadap informasi secara lokal pada *remote server* [4].

Selanjutnya, halaman yang telah didapatkan oleh *crawler* dianalisa oleh *indexing system* dengan cara mengindeks setiap kemungkinan *term* yang terdapat didalamnya dan disimpan ke dalam *database index*. Sejauh ini, kita mengasumsikan bahwa setiap dokumen halaman web adalah dokumen yang bersifat statis, dalam artian bahwa tidak terdapat perubahan terhadap isinya. Tetapi dalam kenyataannya, terkadang suatu halaman web juga terjadi penambahan dokumen, penghapusan dokumen, ataupun pembaharuan. Sehingga akan terjadi juga penambahan, penghapusan ataupun pembaharuan istilah atau *term* yang ada pada daftar list indeks. Cara yang paling mudah untuk menyelesaikan masalah ini adalah dengan merekonstruksi indeks secara berkala.

Dalam tugas akhir ini, akan dikaji metode pengindeksan pada mesin pencari yang memanfaatkan peran dari *mobile crawler* dalam pengambilan datanya. Metode pengindeksan yang akan dikaji adalah metode pengindeksan statis dan metode pengindeksan dinamis :

1. *Static indexing* merupakan metode pengindeksan secara *one-pass* proses, dimana indeks harus dibangun kembali dari awal jika terdapat perubahan.
2. *Dynamic indexing* merupakan metode pengindeksan yang memerlukan penggabungan dari dua buah *index file*[10].

Kedua metode tersebut akan dilihat perbandingan performansinya, dalam kasus ini adalah waktu dan memory yang digunakan berdasarkan jumlah data yang telah diambil oleh *mobile crawler*. Perbandingan ini dilakukan untuk mendapatkan solusi terbaik, dalam hal performansi, terhadap proses pengindeksan dokumen yang sering terjadi perubahan pada isinya.

1.2 Perumusan Masalah

Perumusan masalah dari tugas akhir ini berdasarkan latar belakang masalah yang dipaparkan di atas adalah:

- a. Bagaimana mengimplementasikan *web crawler* dengan menggunakan *mobile agent*
- b. Bagaimana mengimplementasikan metode *dynamic indexing* dan *Static Indexing* pada system yang akan dibangun.
- c. Bagaimana performansi *dynamic indexing* dan *static indexing* dengan parameter waktu dan penggunaan memory jika dibandingkan berdasarkan jumlah dokumen.

Adapun batasan masalah dalam tugas akhir ini adalah sebagai berikut:

1. Kemampuan *mobile crawler* dalam melakukan pengunduhan dokumen hanya mencapai 500 dokumen.
2. Dokumen yang diunduh merupakan dokumen *web* dengan tipe HTML.
3. Data yang digunakan adalah dokumen berita berbahasa Inggris.
4. *Term Indexing* yang dilakukan hanya terdiri dari *Tokenization*, *Stoplist* dan *Stemming*.

1.3 Tujuan

Hal-hal yang ingin dicapai dalam tugas akhir ini yaitu :

1. Merancang dan membangun *web crawler* yang menggunakan *mobile agent*.
2. Mengimplementasikan metode *static indexing* dan *dynamic indexing* untuk mengolah data yang didapatkan oleh *mobile crawler*
3. Menganalisis dan membandingkan hasil dari pengimplementasian metode pengindeksan dengan mengukur performansi dari masing-masing metode konstruksi pengindeksan, dalam hal ini parameter waktu dan penggunaan memory, berdasarkan jumlah dokumen yang terambil oleh *mobile crawler*.

1.4 Metodologi Penyelesaian Masalah

Metode yang digunakan dalam penyelesaian tugas akhir ini adalah:

1. Studi Literatur dengan mempelajari literatur-literatur yang relevan dengan permasalahan yang meliputi studi pustaka dan referensi tentang:
 - a. *Web crawler*
 - b. *Mobile agent*
 - c. *Static Indexing* dan *Dynamic Indexing*
 - d. Dokumen *web*
2. Melakukan perancangan perangkat lunak dengan menggunakan analisis kebutuhan sistem.
3. Melakukan implementasi berdasarkan analisis dan desain sistem yang telah dibuat.
4. Melakukan pengujian perangkat lunak yang telah dibangun
5. Melakukan analisis performansi *mobile crawler* yang mengimplementasikan *static indexing* dan *dynamic indexing*.
6. Penyusunan laporan tugas akhir dan kesimpulan akhir.

1.5 Sistematika Penulisan

Struktur Pembahasan Tugas Akhir ini disusun sebagai berikut :

BAB I PENDAHULUAN

Bab ini berisi latar belakang masalah, perumusan masalah, batasan masalah, tujuan pembahasan, metodologi penyelesaian masalah dan sistematika penulisan.

BAB II LANDASAN TEORI

Membahas dasar teori yang berhubungan dengan pengertian mengenai *World Wide Web (WWW)*, mesin pencari, *web crawler*, *indexing*, *dynamic indexing*, *mobile agent*.

BAB III PERANCANGAN PERANGKAT LUNAK

Bab ini akan membahas proses perancangan *web crawler* dengan memanfaatkan *mobile agent* yang mengimplementasikan *dynamic indexing* dan *static indexing*

BAB IV IMPLEMENTASI DAN ANALISIS HASIL UJI COBA

Membahas tentang analisis dari hasil pengujian ataupun percobaan pada implementasi *web crawler* dengan memanfaatkan *mobile agent* yang mengimplementasikan *dynamic indexing* dan *static indexing*

BAB V KESIMPULAN & SARAN

Pada bab ini akan menjelaskan kesimpulan dan saran sebagai hasil dari analisa dan implementasi Tugas Akhir.