

Penerapan Algoritma Breadth-first Search dan Depth-first Search Pada FTP Search Engine for ITB Network

Cynthia Kustanto¹, Ratna Mutia S², Pocut Viqarunnisa³

Laboratorium Ilmu dan Rekayasa Komputasi
Departemen Teknik Informatika, Institut Teknologi Bandung
Jl. Ganesha 10, Bandung

E-mail : if13066@students.if.itb.ac.id¹, if13086@students.if.itb.ac.id²,
if13106@students.if.itb.ac.id³

Abstrak

Mahasiswa Institut Teknologi Bandung pada umumnya seringkali tidak cukup memiliki referensi dari buku teks ataupun diktat kuliah. Mereka juga seringkali membutuhkan kakas-kakas pemrograman yang *freeware* untuk membantu studi mereka. Sementara itu sudah banyak tersedia referensi-referensi buku juga kakas-kakas pembantu perkuliahan yang menyebar di berbagai server di lingkungan di Institut Teknologi Bandung. Sayangnya mahasiswa sulit untuk mendapatkan referensi tersebut karena kesulitan mengakses dan mengetahui alamat yang tepat dari suatu kakas yang mereka cari. Karena itu dibutuhkan suatu algoritma pencari yang efektif yang dapat mencari kakas yang dibutuhkan yang dapat diterapkan pada FTPsearch di ITB.

Menurut penelitian terdapat tiga pola strategi yang digunakan untuk menemukan file yang dicari pada suatu mesin pencari yaitu *strictly depth-first strategy*, *extreme breath-first strategy* dan *partially breath-first pattern*. Terdapat dua algoritma yang sering digunakan pada mesin pencari, yaitu Breadth-first Search (BFS) yang dimodifikasi dan metode pencarian Depth-first Search (DFS)

Pada makalah ini akan dibahas dan dianalisis kemampuan kedua algoritma tersebut (BFS yang dimodifikasi dan DFS) disesuaikan dengan ketiga pola strategi yang digunakan orang untuk menemukan file (*strictly depth-first strategy*, *extreme breath-first strategy* dan *partially breath-first pattern*) untuk diterapkan pada mesin pencari FTP di ITB (FTP search).

Kata kunci: *strictly depth-first strategy*, *extreme breath-first strategy* dan *partially breath-first pattern*, Breadth-first Search dan Depth-first Search

1. Pendahuluan

Kebutuhan mahasiswa Institut Teknologi Bandung untuk mendapatkan kakas-kakas pemrograman yang *freeware* atau referensi tambahan untuk membantu studi semakin dibutuhkan. Sementara itu sudah banyak referensi-referensi buku juga kakas-kakas pembantu perkuliahan yang menyebar di berbagai server di lingkungan di Institut Teknologi Bandung. Karena itu dibutuhkan suatu algoritma pencari yang efektif yang dapat diterapkan pada FTP Search di ITB sebagai mesin pencari yang umum digunakan di lingkungan ITB. Algoritma ini disesuaikan dengan tiga pola strategi yang digunakan orang untuk menemukan file yang dicari pada suatu mesin pencari.

2. Breadth-first Search

2.1 Pengertian Breadth First Search

Breadth-first search adalah algoritma yang melakukan pencarian secara melebar yang mengunjungi simpul secara *preorder* yaitu mengunjungi suatu simpul kemudian mengunjungi semua simpul yang bertetangga dengan simpul tersebut terlebih dahulu. Selanjutnya, simpul yang belum dikunjungi dan bertetangga dengan simpul-simpul yang tadi dikunjungi, demikian seterusnya. Jika graf berbentuk pohon berakar, maka semua simpul pada aras d dikunjungi lebih dahulu sebelum simpul-simpul pada aras $d+1$.

Algoritma ini memerlukan sebuah antrian q untuk menyimpan simpul yang telah dikunjungi. Simpul-simpul ini diperlukan sebagai acuan untuk mengunjungi simpul-simpul yang bertetangganya. Tiap simpul yang telah dikunjungi masuk ke dalam antrian hanya satu kali.

Algoritma ini juga membutuhkan table Boolean untuk menyimpan simpul yang telah dikunjungi sehingga tidak ada simpul yang dikunjungi lebih dari satu kali.

2.2 Metode Pencarian

Salah satu contoh kakas pencarian yang menggunakan metode BFS adalah *WebCrawler*.

WebCrawler adalah suatu kakas yang membuat indeks isi dari suatu dokumen di Web yang selanjutnya akan dimanfaatkan oleh mesin pencari.

Terdapat tiga langkah yang dilakukan oleh *WebCrawler* ini ketika mengunjungi dokumen, yaitu menandai bahwa suatu dokumen telah dikunjungi, mengenali link yang terdapat pada dokumen tersebut, kemudian isinya didaftarkan pada daftar indeks.

Pada akhirnya, *WebCrawler* akan menampilkan file yang paling banyak berkaitan dengan kata kunci.

Contoh hasil pencarian dengan *WebCrawler*

Kata kunci : "molecular biology human genome project"

Hasil pencarian :

1000 [Guide to Molecular Biology Databases](#)
0754 [Guide to Molecular Biology Databases](#)
0562 [Biology Links](#)
0469 [bionet newsgroups](#)
0412 [Motif Bioinformatics Server](#)
0405 [LBL Catalog of Research Abstracts \(1993\)](#)
0329 [Molecular Biology Databases](#)
0324 [GenomeNet WWW server](#)
0317 [DOE White Paper on Bio-Informatics](#)
0292 [Molecular biology](#)
0277 [Oxford University Molecular Biology Data Centre Home Page](#)
0262 [Biology Internet Connections](#)
0244 [Harvard Biological Laboratories - Genome Research](#)
0223 [About the GenomeNet](#)
0207 [Biology Index](#)

Algoritma BFS yang diterapkan pada *WebCrawler* ini adalah mendaftarkan setiap link yang ada dan menyimpannya dalam data base kemudian mengunjunginya satu persatu sambil mencatat link yang berhubungan

3. Depth-first Search

3.1 Pengertian Depth First Search

Depth-first search (DFS) melakukan pencarian secara preorder. Mengunjungi anak suatu simpul sebelum simpul tetangganya. Berkaitan dengan mesin pencari, DFS ini cenderung mengindeks dokumen berdasarkan suatu link.

3.2 Metode Pencarian

Algoritma DFS yang diterapkan pada mesin pencari dalam melakukan pengindeksan adalah mengunjungi suatu server kemudian menyimpan

semua link yang berhubungan dengan server tersebut baru kemudian mengunjungi server lain.

Salah satu yang menerapkan algoritma DFS pada mesin pencarian adalah FTPSearch.

FTPSearch adalah suatu mesin pencari dokumen yang tersimpan di jaringan ITB. Dapat diakses pada <http://ftpsearch.itb.ac.id>.

FTPSearch akan menampilkan daftar hasil pencarian berdasarkan server. File-file yang tersimpan pada suatu server akan ditampilkan terlebih dahulu kemudian baru berpindah pada server lain. FTPSearch tidak memperhatikan file mana yang lebih berkaitan dengan kata kunci karena FTPSearch tidak melakukan observasi sampai pada isi dokumen tapi hanya melihat judul dokumen.

4. Pola strategi menemukan file pada suatu mesin pencari

Secara umum terdapat 3 pola strategi yang dilakukan oleh user mesin pencari untuk menemukan file yang sesuai dengan kata kunci yaitu *strictly depth-first strategy*, *extreme breath-first strategy* dan *partially breath-first pattern*

4.1 Pengertian

Metode *strictly depth-first strategy* berarti pengguna mengamati tiap link hasil dari mesin pencari mulai dari atas, dan memutuskan segera untuk membuka dokumen atau tidak.

Metode *extreme breath-first strategy* berarti pengguna melihat keseluruhan link daftar hasil pencarian kemudian memilih link yang mengacu ke dokumen yang paling sesuai lalu mengunjungi link tersebut. Sekali-kali pengguna juga melihat kembali daftar hasil pencarian lalu mengunjungi ulang beberapa link yang menurut dia berkaitan.

Metode *partially breath-first pattern* merupakan metode campuran dimana pengguna melihat beberapa link baru memutuskan link mana yang akan dibuka.

4.2 Hasil Eksperimen

Pada eksperimen yang dilakukan oleh penelitian yang dilakukan oleh Kerstin Klockner, Nadine Wirshum, Anthony Jameson, empat puluh satu subjek diberi waktu sepuluh menit untuk mendapatkan informasi tentang "Assessment centers" dengan cara membuka dokumen yang relevan yang dikembalikan oleh Google sebagai respon pencarian. Sebuah daftar hasil pencarian terdiri dari 25 hasil telah disiapkan sebelumnya dan disajikan dalam sebuah halaman web. Subjek harus menggulung layer untuk dapat melihat keseluruhan halaman. Gerakan mata para subjek dan aksi klik tetikus direkam dengan bantuan pendeteksi ASL

504. Berdasarkan rekaman video yang dibuat melalui pendeteksi itu, pemrosesan hasil pencarian yang dilakukan para subjek dianalisis. Tiga kategori diidentifikasi : Sebagian besar subjek (65%) mengaplikasikan strategi DFS. Kebalikannya, minoritas subjek mengaplikasikan pola strategi BFS yang ekstrim, melihat ke seluruh daftar hasil pencarian sebelum membuka sebuah dokumen. Pola BFS yang parsial ditunjukkan oleh sisa 20% subjek, yang terkadang melihat ke beberapa entri berikutnya sebelum memilih dokumen yang akan dibuka.

Pada eksperimen dua, dua puluh tujuh subjek diminta untuk melakukan dua aktivitas yang sama dengan eksperimen satu, dengan waktu 5 menit untuk masing-masing aktivitas. Untuk menciptakan situasi yang membuat BFS terlihat relative menarik, kita memperbolehkan subjek untuk membuka maksimal sepuluh dari 25 dokumen yang di daftar, memberi penghargaan kepada mereka untuk setiap penemuan dokumen yang relevan (hamper setengah dari dokumen relevan). Di sini, strategi yang berlawanan ditemukan lagi : 52% subjek tidak menunjukkan niat untuk melihat keseluruhan daftar. Minoritas subjek sebanyak 11% menggunakan strategi BFS yang ekstrim, melihat ke seluruh daftar sebelum membuka sebuah dokumen; sisanya sebanyak 37% mengaplikasikan strategi campuran, melihat dulu ke dua sampai enam dokumen berikutnya dalam daftar.

4. Analisis

Berdasarkan penelitian yang dilakukan oleh Kerstin Klockner, Nadine Wirshum, Anthony Jameson pada makalahnya yang berjudul "Depth- and Breadth-First Processing of Search Result" kita dapat menyimpulkan bahwa pengguna cenderung melakukan pencarian secara *strictly depth-first strategy*. Yaitu melihat suatu link yang paling atas dari hasil pencarian kemudian mengaksesnya dan terus menelusuri link yang terdapat pada document tersebut yang berkaitan dengan kata kunci yang dicari.

Sehingga agar pencarian oleh FTP search lebih efektif perlu ada penyesuaian dengan algoritma mesin pencari. Algoritma yang menurut kami paling sesuai adalah algoritma BFS pada mesin pencari yang akan menampilkan daftar file yang paling dekat relevansinya dengan kata kunci.

Metode FTP search melakukan hal yang sama dengan *WebCrawler* yaitu mengunjungi setiap server, mencatat link file dan memasukkannya ke dalam basis data. Sehingga file yang paling relevan ditempatkan di bagian paling atas daftar hasil pencarian. Perbedaan dengan web crawler ftp search akan mengelompokkan file-file tersebut berdasarkan server.

Kelebihan metode baru ini bagi pengguna FTP search adalah pengguna akan dapat langsung melihat file yang paling relevan pada bagian atas

daftar hasil pencarian dan meminimalisasi pengaksesan lintas server.

5. Kesimpulan

Berdasarkan analisis yang dilakukan maka algoritma BFS pada mesin pencarian dengan memperhatikan kebiasaan pengguna dalam membuka file daftar hasil pencarian akan mengefektifkan pencarian file di ftp search. Sehingga daftar hasil pencarian FTP search dapat lebih memudahkan pengguna dalam melakukan pencarian.

6. Daftar Pustaka

- [1] Pinkerton Brian.1994."Finding What People Want: Experiences with the WebCrawler"
- [2] Kl'ockner Kerstin. 2005 "Depth- and Breadth-First Processing of Search Result Lists: An Example Paper in the SIGCHI Style"
- [3] Salter Richard
<http://www.cs.oberlin.edu/classes/drag/labs/graphs/graphs5.html>. Diakses tanggal 17 Mei 2005
- [4] Munir Rinaldi. 2005. "Diktat Kuliah Strategi Algoritmik"