

“Kombinasi Algoritma Pattern Matching dan BFS-DFS pada aplikasi Music Discovery”



Disusun Oleh :

Levanji Prahyudy / 13513052

Kombinasi Algoritma Pattern-Matching dan BFS-DFS dalam aplikasi “Music Discovery”

Levanji Prahudy / 13513052
Program Studi Teknik Informatika
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia
13513052@std.stei.itb.ac.id

Abstraksi — Pada makalah ini akan dibahas penerapan Algoritma Pattern Matching dan Algoritma BFS-DFS dalam aplikasi “Music Discovery”. Music Discovery adalah aplikasi untuk melakukan pencarian musik dengan suara asli dari lingkungan sekitar aplikasi tersebut. Jadi dengan makalah ini diharapkan pembaca memahami konsep dari aplikasi tersebut.

Kata kunci — Pattern, Matching, BFS, DFS, Algoritma, Music, Discovery

I. PENDAHULUAN

Akhir – akhir ini, aplikasi – aplikasi pada platform khusus seperti *mobile* berkembang dengan pesat. Setiap harinya, sedikitnya ada sekitar 100 aplikasi baru yang dikeluarkan, mulai dari aplikasi kecil sampai yang besar, bahkan dari aplikasi yang sangat berguna sampai yang tidak berguna sama sekali.

Tetapi karena pada umumnya kumpulan aplikasi – aplikasi tersebut tersedia penilaian pengguna terhadap aplikasi yang bersangkutan, dan orang dapat menilai sebuah aplikasi secara bebas, maka aplikasi yang tidak berguna biasanya mendapat penilaian buruk sehingga jarang atau bahkan tidak ada lagi yang mengunduh aplikasi tersebut. Jadi dalam makalah ini kita akan membahas sebuah aplikasi yang mendapatkan penilaian bagus dari jutaan penggunanya, yaitu aplikasi untuk mencari lagu dari potongan lagu yang dimasukkan dari suara asli, biasa dikenal dengan sebutan aplikasi “Music Discovery”.

Mengingat banyaknya algoritma penyelesaian masalah yang telah dikenal secara internasional, kita dapat mengetahui penerapan algoritma – algoritma tersebut pada berbagai aplikasi, mungkin hampir semua aplikasi berskala menengah atau besar. Pada konteks ini algoritma yang diterapkan adalah algoritma pencocokan string (*String Matching*) dan algoritma pencarian BFS-DFS.

Algoritma *Pattern-Matching* seringkali diterapkan dalam pencocokan String. Algoritma ini melakukan pencarian semua kemunculan string pendek yang disebut “*pattern*” di string yang lebih panjang yang disebut “*text*”. Dalam makalah ini akan *pattern* dan *text* akan dituliskan dalam bahasa Indonesia, menjadi “pola” dan “teks”.

Pencocokan string merupakan permasalahan yang paling

sederhana dari semua permasalahan string lainnya, dan seringkali dianggap sebagai bagian dari pemrosesan data dan temu balik informasi.

Sedangkan Algoritma BFS-DFS adalah algoritma pencarian data dalam struktur “pohon”. Algoritma ini mulai dari “akar” sebagai kunci, dan mengeksplorasi simpul yang bertetangga dengan “akar” tersebut. BFS adalah algoritma yang mencari dengan cara melebar, sedangkan DFS adalah algoritma pencarian dengan cara mendalam.

Permasalahan – permasalahan di dunia nyata biasanya cukup kompleks, sehingga Algoritma baik pencocokan string maupun BFS-DFS dapat digunakan untuk menyelesaikan permasalahan bentuk kompleks tersebut menjadi sederhana. Sebagai contoh, BFS-DFS digunakan untuk menyelesaikan persoalan pedagang keliling (*Travelling Salesman Problem*)^[3]. Jadi dalam makalah ini akan dibahas aplikasi kedua algoritma tersebut dalam permasalahan yang lebih dikenal oleh para pecinta teknologi, yaitu aplikasi “Music Discovery”.

II. DASAR TEORI

A. Algoritma Pencocokan String

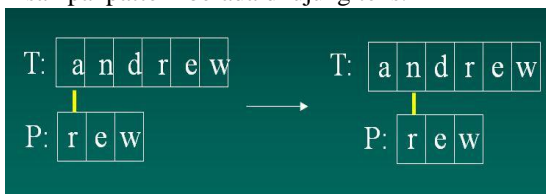
Algoritma-algoritma pencocokkan string dapat diklasifikasikan menjadi tiga bagian menurut arah pencariannya^[1]. Tiga kategori itu adalah :

1. Dari arah yang paling alami, dari kiri ke kanan, yang merupakan arah untuk membaca, algoritma yang termasuk kategori ini adalah:
 - a. Algoritma Brute Force
 - b. Algoritma dari Morris dan Pratt, yang kemudian dikembangkan oleh Knuth, Morris, dan Pratt
2. Dari kanan ke kiri, arah yang biasanya menghasilkan hasil terbaik secara praktikal, contohnya adalah:
 - a. Algoritma dari Boyer dan Moore, yang kemudian banyak dikembangkan, menjadi Algoritma turbo Boyer-Moore, Algoritma tuned Boyer-Moore, dan Algoritma Zhu-Takaoka;

Contoh sistematika pencarian yang akan diambil adalah Algoritma Brute Force, karena Algoritma Brute Force merupakan algoritma pencocokan string yang ditulis tanpa memikirkan peningkatan performa. Algoritma ini sangat jarang dipakai dalam praktik, namun berguna dalam studi pembandingan dan studi-studi lainnya.

Secara sistematis, langkah-langkah yang dilakukan algoritma brute force pada saat mencocokkan string adalah:

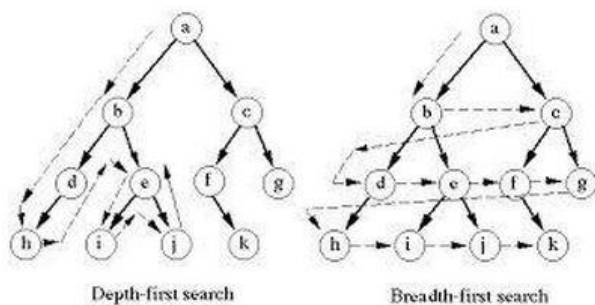
1. Algoritma brute force mulai mencocokkan pattern pada awal teks.
2. Dari kiri ke kanan, algoritma ini akan mencocokkan karakter per karakter pattern dengan karakter di teks yang bersesuaian, sampai salah satu kondisi berikut dipenuhi:
 - 2.1 Karakter di pattern dan di teks yang dibandingkan tidak cocok (mismatch).
 - 2.2 Semua karakter di pattern cocok. Kemudian algoritma akan memberitahukan penemuan di posisi ini.
3. Algoritma kemudian terus menggeser pattern sebesar satu ke kanan, dan mengulangi langkah ke-2 sampai pattern berada di ujung teks.



Gambar 1. Pattern Matching

B. Algoritma BFS dan DFS

Algoritma BFS (Breadth-First Search) dan DFS (Depth-First Search) adalah algoritma pencarian yang dipakai apabila tidak ada informasi tambahan. Seperti yang telah disinggung sebelumnya, pada “pohon” data, BFS akan mengunjungi semua simpul yang bertetangga dengan akar terlebih dahulu, sedangkan DFS akan mengunjungi salah satu simpul yang bertetangga dengan akar terlebih dahulu^[2].



Gambar 2. BFS dan DFS

Tahap – tahap pencarian dengan menggunakan algoritma BFS secara garis besar adalah sebagai berikut :

1. Kunjungi simpul v
2. Kunjungi semua simpul yang bertetangga dengan simpul v terlebih dahulu.
3. Kunjungi simpul yang belum dikunjungi dan bertetangga dengan simpul-simpul yang tadi dikunjungi, demikian seterusnya.

Sedangkan pencarian dengan menggunakan DFS secara garis besar adalah sebagai berikut :

1. Kunjungi simpul v
2. Kunjungi simpul w yang bertetangga dengan simpul v.
3. Ulangi DFS mulai dari simpul w.
4. Ketika mencapai simpul u sedemikian sehingga semua simpul yang bertetangga dengannya telah dikunjungi, pencarian dirunut-balik (backtrack) ke simpul terakhir yang dikunjungi sebelumnya dan mempunyai simpul w yang belum dikunjungi.
5. Pencarian berakhir bila tidak ada lagi simpul yang belum dikunjungi yang dapat dicapai dari simpul yang telah dikunjungi.

C. Kombinasi Algoritma Pattern-Matching dengan BFS-DFS

Kedua algoritma yang telah dibahas memiliki sebuah keterkaitan, yakni keduanya sama – sama merupakan algoritma pencarian. Jadi kedua algoritma tersebut dapat diterapkan bersamaan atau dikombinasikan dalam sebuah pencarian.

Dalam proses pencarian pada pohon dengan BFS maupun DFS, apabila data yang terkandung dalam pohon tersebut berupa string atau pola, maka algoritma pencocokan pola dapat diterapkan juga secara bersamaan. Kegunaan dari kombinasi kedua algoritma tersebut adalah untuk membuat pencarian menjadi lebih mangkus, karena biasanya pada penerapan algoritma BFS atau DFS, pencarian string yang dilakukan adalah Brute Force, yang memiliki kompleksitas algoritma yang relatif buruk jika dibandingkan dengan algoritma pencocokan string yang lain.

III. PEMBAHASAN

A. Konsep Music Discovery

Music Discovery adalah aplikasi untuk mencari musik atau lagu yang tersedia di internet dengan memanfaatkan suara dari lingkungan sekitar. Biasanya ketika kita mengetahui nada atau potongan musik tetapi tidak mengetahui judul lagu tersebut, kita menggunakan aplikasi ini untuk mencari lagu tersebut dari internet.

Pada aplikasi Music Discovery, potongan lagu diinterpretasikan sebagai *keyword* atau pola, dan lagu secara keseluruhan diinterpretasikan sebagai teks yang akan dicari dengan algoritma pencocokan string.

Sedangkan algoritma BFS-DFS diterapkan saat pencarian di internet.



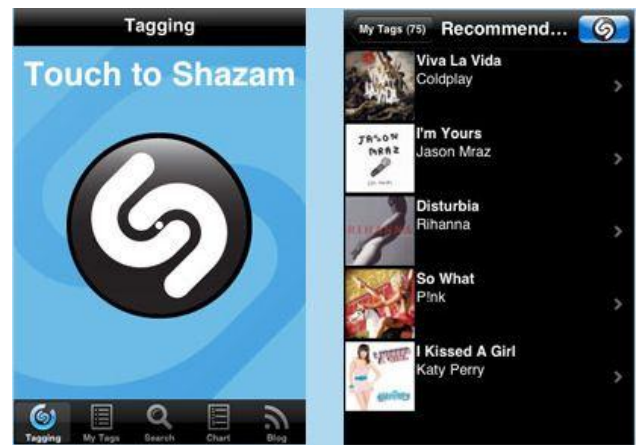
Gambar 3. Gelombang suara sebelum diinterpretasikan ke string

Karena aplikasi ini memanfaatkan suara asli, aplikasi ini menerjemahkan pola suara asli tersebut menjadi string yang melambungkan gelombang suara music dengan menerapkan hukum Fisika. Jadi secara garis besar, aplikasi ini mencari keyword berupa gelombang suara di internet menggunakan algoritma BFS-DFS, sambil mencocokkan gelombang suara pada lagu asli dengan menggunakan algoritma pencocokan string.

Sama halnya dengan Search Engine, Aplikasi Music Discovery tidak hanya mencari satu lagu yang memiliki pola yang sesuai dengan potongan lagu yang dicari, tetapi mencari beberapa lagu dengan batas jumlah pencarian yang ditetapkan oleh pembuat aplikasinya sendiri. Aplikasi ini juga akan berhenti mencari ketika melewati batas jumlah pencarian yang ditentukan.

Akan tetapi, prioritas yang dicari tentunya akan berbeda dengan aplikasi search engine lainnya. Ketika pencocokan string yang lain mencari dengan prioritas artikel, aplikasi Music Discovery mencocokkan potongan lagu yang telah diterjemahkan ke dalam bentuk string dengan memprioritaskan pencarian pada website yang memiliki kumpulan video musik, sebagai contoh website yang sudah sangat terkenal di jaman sekarang yaitu Youtube, tetapi tentu dengan tampilan yang berbeda, bukan seperti mencari di Youtube. Sebagai contoh aplikasi Music Discovery yang diambil adalah “Shazam”.

Tampilan pada aplikasi “Shazam”, sebagai salah satu Music Discovery adalah desain yang dibuat sendiri oleh developer dari aplikasi tersebut. Namun, ketika hasil pencarian diklik atau disentuh (pada platform *Touch-Screen Mobile*), aplikasi ini akan mengarahkan pengguna ke Youtube. Contoh tampilan terdapat pada gambar di bawah ini :



Gambar 4. Tampilan Shazam

B. Penerapan Pattern Matching

Pada aplikasi Music Discovery, Pattern Matching digunakan untuk mencocokkan pola yang berupa potongan lagu dengan teks yang berupa lagu secara keseluruhan. Secara umum aplikasi ini membatasi input pola yang dimasukkan, misalnya 10 detik. Jadi aplikasi ini hanya mencocokkan lagu yang dimiliki oleh website berbasis video musik dengan 10 detik potongan lagu tersebut.

Karena pola gelombang suara cukup rumit untuk diinterpretasikan sebagai string, maka akan diberikan *prototype* interpretasi gelombang suara sebagai string saja. Penerapan algoritma pencocokan pola yang dipakai bisa berbeda – beda tergantung aplikasi Music Discovery yang dibuat. Gambaran interpretasi gelombang suara sebagai string dapat dilihat pada gambar di bawah ini :

Pola : 11010110
 Teks :
 01101001010010101010
 01101011001010101001
 01010010101100101

Gambar 5. Pencocokan pola pada aplikasi Music Discovery

Gambar di atas merupakan gambaran interpretasi potongan lagu 10 detik tersebut sebagai pola yang ingin dicari, yaitu 11010110, sedangkan lagu keseluruhan berupa teks yang mengandung potongan lagu 10 detik tersebut. Walaupun tidak sesederhana yang digambarkan (misalnya pada aplikasi tersebut menetapkan toleransi kesalahan pada pola, seperti mengutamakan pencarian suara yang paling dominan pada lagu tersebut), tetapi secara garis besar gambaran tersebut mewakili aplikasi

Algoritma pencarian pola pada aplikasi Music Discovery.

Pola : 11010111
Teks :
01101001010010101010
01101011001010101001
01010010101100101

Toleransi Kesalahan

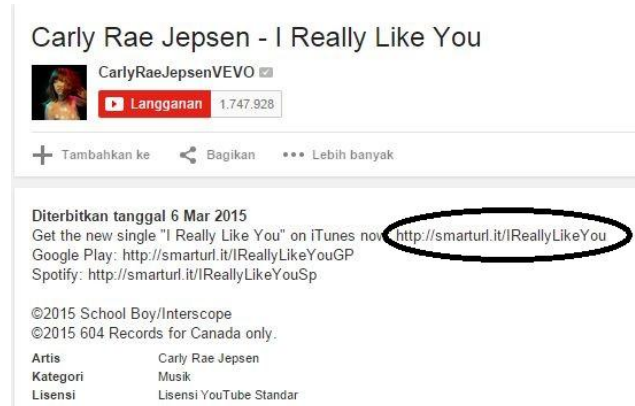
Gambar 6. Contoh Toleransi Kesalahan pada Aplikasi

Dapat dipastikan pembuat aplikasi ini menggunakan algoritma pencocokan pola dengan menyesuaikan tipe pola yang akan dicari, sehingga algoritma yang digunakan memiliki kompleksitas algoritma semangkus mungkin. Misalnya apabila pola gelombang suara ketika diinterpretasikan sebagai string yang satu dengan yang lainnya mirip, tentu pembuat aplikasi lebih memilih menggunakan algoritma Knuth-Morris-Pratt dibandingkan dengan Boyer-Moore, atau algoritma pencocokan pola yang lain yang lebih mangkus.

C. Penerapan Algoritma BFS-DFS

Selain menerapkan algoritma pencocokan pola, aplikasi Music Discovery ini secara bersamaan menerapkan algoritma BFS atau DFS untuk mencari musik video di website yang telah ditetapkan oleh pembuat aplikasi tersebut.

Jika pencarian yang digunakan adalah BFS, maka aplikasi akan mulai mencari video dengan cara melebar, yaitu mencari dari satu video lagu, pindah ke video lagu yang lain, dan seterusnya. Kemudian mulai mencari dari video lagu pertama jika deskripsi video tersebut menampilkan *hyperlink* ke video lain, kemudian *hyperlink* pada lagu kedua, dan seterusnya. Contoh *hyperlink* yang ditampilkan pada sebuah video musik dapat dilihat pada gambar di bawah ini :



Gambar 7. Hyperlink pada Musik Video yang dicari

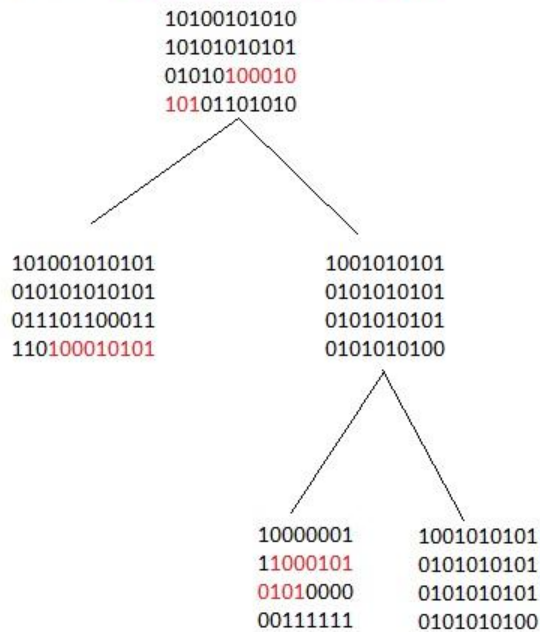
Selain itu jika pencarian yang digunakan adalah DFS, maka aplikasi akan mulai mencari video dengan cara mendalam yaitu mencari dari satu video lagu, pindah ke *hyperlink* yang ditampilkan oleh video lagu tersebut jika ada, dan ke *hyperlink* lainnya. Jika sudah tidak ada *hyperlink* ke video lain (atau *hyperlink* yang ditampilkan sama dengan video yang sudah pernah dicari sebelumnya), aplikasi tersebut melakukan *backtrack* ke video sebelumnya, dan seterusnya.

Namun pada penerapannya, ketika aplikasi ini masuk ke sebuah URL, pertama – tama aplikasi ini memvalidasi apakah URL itu berisi video atau bukan. Jika bukan, otomatis berpindah ke URL lain, tergantung algoritma apa yang digunakan, BFS atau DFS.

D. Penerapan Kombinasi Algoritma Pattern-Matching dengan BFS-DFS

Aplikasi Music Discovery ini menerapkan kombinasi algoritma pencocokan string dan algoritma BFS-DFS, seperti yang sudah dibahas di bab Dasar Teori. Pencarian di internet yang diterapkan menggunakan algoritma BFS-DFS, sedangkan pencocokan string yang dilakukan menerapkan algoritma pencocokan pola. Gambaran kombinasi kedua algoritma dapat dilihat pada gambar di bawah ini :

Pola : 10001010101



Gambar 8. Kombinasi Algoritma Pattern Matching dan BFS-DFS

IV. KESALAHAN UMUM

Pada penerapannya, seringkali pembaca mengasumsikan bahwa gelombang suara yang dihasilkan berupa string biner. Padahal, aplikasi ini juga harus memanfaatkan algoritma lain yaitu mengubah gelombang suara menjadi string. Selain itu, pembaca seringkali menganggap pencarian string dilakukan sekali per lagu, padahal sebenarnya beberapa kali, masing – masing untuk gelombang suara yang berbeda – beda. Sebuah lagu memiliki beberapa gelombang suara, misalnya gelombang suara gitar, keyboard, dan lain – lain, jadi pada praktiknya, pencarian 1 lagu dilakukan beberapa kali sesuai jumlah gelombang suara yang dimiliki oleh lagu tersebut.

V. KESIMPULAN

Berdasarkan pembahasan di atas, dapat ditarik kesimpulan bahwa aplikasi pencari lagu atau Music Discovery menerapkan kombinasi algoritma pencocokan string dan algoritma BFS-DFS.

REFERENSI

- [1] (Inggris)Lecroq, Thierry Charras, Christian. 2001. Handbook of Exact String Matching Algorithm. ISBN 0-9543006-4-5.
- [2] Munir, Rinaldi. "Strategi Algoritma", Informatika, Bandung : 2010
- [3] Leiserson, Charles E.; Schardl, Tao B. (2010). A Work-Efficient Parallel Breadth-First Search Algorithm (or How to Cope with the

- Nondeterminism of Reducers) (PDF). ACM Symp. on Parallelism in Algorithms and Architectures.
- [4] <http://www.mobiletor.com/38049/shazam-updates-its-music-discovery-apps-for-iphone-and-ipod-touch/>

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 1 Mei 2014

Levanji Prahudy / 13513052