

Penerapan Algoritma Boyer Moore Horspool Pada Pencarian Katalog Buku

Dian wahyuni Kartini, Rizky Yanur Tanjung, Indra Lasmana

Prodi Teknik Informatika, STMIK Budi Darma, Medan, Indonesia

Email: dianwahyuninababan2104@gmail.com

Abstrak

Katalog adalah presentasi ciri-ciri dari sebuah bahan pustaka atau dokumen (misalnya: judul, pengarang, deskripsi fisik, subyek, dll.) koleksi perpustakaan yang merupakan wakil ringkas bahan pustaka tersebut yang disusun secara sistematis. Algoritma *Horspool* menerapkan informasi pencarian yang di simpan untuk melakukan pergeseran yang lebih jauh karena pencocokan string dilakukan dari kanan ke kiri, sehingga waktu pencarian string lebih singkat. *Field* yang digunakan dalam pencarian ini adalah istilah buku secara umum. Dengan memasukkan nama istilah buku sebagai inputannya dalam pencarian, kemudian akan didapatkan hasil pencarian berupa arti dari istilah tersebut. hasil dari aplikasi ini menampilkan keseluruhan *pattern* yang cocok dengan teks.

Kata Kunci: Penerapan Algoritma *Boyer Moore Horspool* Pada Pencarian Katalog Buku

1. PENDAHULUAN

Teknologi informasi dan pertumbuhan telekomunikasi saat ini sangat berkembang, sehingga penggunaan dan pemanfaatan teknologi informasi dan teknologi jaringan merupakan suatu indikator pertumbuhan kemajuan teknologi dan pertumbuhan disegala bidang, khususnya dunia pendidikan. Perkembangan teknologi informasi dan telekomunikasi menuntut cara kerja yang cepat dan tepat untuk memberikan informasi yang serba akurat. Salah satu bentuk perkembangan teknologi adalah jaringan internet yaitu suatu bentuk penyampaian informasi teknologi dalam dunia maya yang didukung oleh infrastruktur teknologi seperti hardware, software dan brainware. Dengan jaringan internet kita dapat menerima atau mengakses informasi dalam berbagai format. Komputer merupakan media teknologi informasi yang dapat menyajikan informasi yang diperlukan, juga dapat membantu mengurangi kesalahan atau kekeliruan yang sering terjadi pada sistem manual, sehingga hasil yang diperoleh dapat dimanfaatkan dengan tepat dan cepat. Oleh karena itu penggunaan teknologi 341omputer dan perangkat lunak (software) yang baik akan mengurangi kesalahan dalam penyajian data yang diperlukan.

Katalog merupakan daftar koleksi sebuah pusat dokumentasi atau beberapa pusat dokumentasi yang disusun menurut sistem tertentu.^[1] Daftar tersebut dapat berbentuk kartu, lembaran, buku atau bentuk lain, yang memuat informasi mengenai pustaka atau perpustakaan yang terdapat di perpustakaan atau unit informasi^[2].

Sesuai dengan perkembangan jaman bentuk-bentuk sumber informasi itu tidak hanya berupa barang cetak semata, seperti buku, Koran, surat kabar dan sejenisnya. Jenis-jenis sekarang sudah banyak yang di ahlimediasi, seperti CD, Flash Disk, Film dan sejenisnya. Mengingat beragamnya dari sumber informasi akan mempengaruhi kecepatan dan ketepatan penggunaannya untuk menemukan kembali sumber informasi yang dibutuhkannya.

Algoritma Horspool merupakan penyederhanaan dari algoritma Boyer-Moore. Perbedaan antara keduanya adalah pada metode penggeseran *pattern*nya. Jika Boyer-Moore menggunakan dua metode praproses *bad character shift* dan *good shufix shift*, akan tetapi Horspool hanya menggunakan satu metode praproses yaitu *bad character shift*. Kompleksitas rata-rata algoritma ini sama dengan Boyer-Moore $O(n)$, seangkan untuk metode praproses nya adalah $O(m+\sigma)$.

2. TEORITIS

2.1 String Matching

String matching adalah proses pencarian semua kemunculan queri yang selanjutnya disebut *pattern* kedalam string yang lebih panjang atau teks. Pencocokan string (*string matching*) merupakan bagian penting dari sebuah proses pencarian string dalam dokumen tergantung dari teknik dan cara pencocokan string yang sering digunakan. Prinsip kerja algoritma *string matching* adalah sebagai berikut:

1. Memindahkan teks dengan bantuan sebuah *window* yang ukurannya sama dengan *pattern*.
2. Menempakan *window* pada awal.
3. Membandingkan karakter pada *window* dengan karakter dari *pattern*. Setelah pencocokan (baik hasilnya cocok atau tidak cocok), dilakukan *shift* kekanan pada *window*. Prosedur ini dilakukan berulang-ulang sampai *window* berada pada akhir teks. Mekanisme ini disebut mekanisme suding *window*.

2.2 Algoritma Horspool

Algoritma *Horspool* adalah penyederhanaan dari algoritma *Boyer-Moore* yang dibuat oleh R. Nigel Horspool . Menurut *Horspool*, R.N. 1980, masalah dalam pencarian teks ini adalah mencari dalam teks yang besar untuk menemukan *pattern*

pertama. [1] Karena teks yang dicari bisa sangat besar memungkinkan ratusan ribu karakter maka penting untuk menggunakan teknik yang lebih efisien. Algoritma *Horspool* bekerja dengan metode yang hampir sama dengan algoritma *Boyer-Moore* namun tidak melakukan lompatan berdasarkan karakter pada *pattern* yang ditemukan tidak cocok pada teks. Algoritma *Horspool* mempunyai nilai pergeseran karakter yang paling kanan dari *window*. Pada tahap observasi awal preprocessing, nilai *shift* akan dihitung untuk semua karakter. Pada tahap ini, dibandingkan *pattern* dari kanan ke kiri hingga kecocokan atau ketidakcocokan *pattern* terjadi. Karakter yang paling kanan pada *window* digunakan sebagai indeks dalam melakukan nilai *shift*. Dalam kasus ketidakcocokan karakter tidak terdapat pada *pattern* terjadi, *window* digeser oleh panjang dari sebuah *pattern*. Jika tidak, *window* digeser menurut karakter yang paling kanan pada *pattern* Baeza-Yates, R.A. Regnier, M. 1992.

Terdapat dua tahap pada pencocokan string menggunakan algoritma *Horspool* Singh, R. Verma, H.N. 2011, yaitu:

1. Praproses Pada tahap ini, dilakukan observasi *pattern* terhadap teks untuk membangun sebuah tabel *bad-match* yang berisi nilai *shift* ketika ketidakcocokan antara *pattern* Secara sistematis, langkah-langkah yang dilakukan algoritma *Horspool* pada tahap praproses adalah:
 - a. Algoritma *Horspool* melakukan pencocokan karakter ter-kanan pada *pattern*.
 - b. Setiap karakter pada *pattern* ditambah ke dalam tabel *bad-match* dan dihitung nilai *shift*-nya.
 - c. Karakter yang berada pada ujung *pattern* tidak dihitung dan tidak dijadikan karakter ter-kanan dari karakter yang sama dengannya.
 - d. Apabila terdapat dua karakter yang sama dan salah satunya bukan karakter ter-kanan, maka karakter dengan indeks terbesar yang dihitung nilai *shift*-nya.
 - e. Algoritma *Horspool* menyimpan panjang dari *pattern* sebagai panjang nilai *shift* secara default apabila karakter pada teks tidak ditemukan dalam *pattern*.
2. Tahap pencarian Secara sistematis, langkah-langkah yang dilakukan algoritma *Horspool* pada tahap praproses adalah :
 - a. Dilakukan perbandingan karakter paling kanan *pattern* terhadap *window*.
 - b. Tabel *bad-match* digunakan untuk melewati karakter ketika ketidakcocokan terjadi.
 - c. Ketika ada ketidakcocokan, maka karakter paling kanan pada *window* berfungsi sebagai landasan untuk menentukan jarak *shift* yang akan dilakukan.
 - d. Setelah melakukan pencocokan baik hasilnya cocok atau tidak cocok dilakukan pergeseran ke kanan pada *w indow*.
 - e. Prosedur ini dilakukan berulang-ulang sampai *window* berada pada akhir teks atau ketika *pattern* cocok dengan teks.

2.3 Katalog Buku

Katalog buku adalah daftar koleksi sebuah pusat dokumentasi atau beberapa pusat dokumentasi yang disusun menurut sistem tertentu [2]. Daftar tersebut dapat berbentuk kartu, lembaran, buku atau bentuk lain, yang memuat informasi mengenai pustaka atau kepustakaan yang terdapat di perpustakaan atau unit informasi [3], Adapun keterangan-keterangan yang dituliskan pada katalog adalah sebagai berikut:

- a. Nama Lengkap (dengan gelar-gelar kebangsawanannya) dari si pengarang, atau yang dianggap pengarang
- b. Judul buku, jika ada judul tambahannya serta edisi.
- c. Tempat penerbit dan nama penerbitnya
- d. Tahun terbit atau *copy-right* datenya.
- e. Jumlah halaman pendahuluan (angka romawi) dan jumlah halaman isi (angka arab). Jumlah halaman bisa diganti dengan jumlah jilid yang ada, jika buku-buku itu berjilid.
- f. Keterangan-keterangan (*notes*) seperti: app;digr;graf;illus.,tab., dan lain-lain
- g. Keterangan tentang halaman bibliografi.
- h. *call-number* buku (biasanya ditentukan pada awal proses buku).
- i. Nomor stambuk buku (pada pinggir kiri bawah kartu).

2.4 Bentuk-Bentuk Katalog

Berikut adalah beberapa bentuk fisik katalog:

1. Katalog berbentuk kartu
Katalog ini berukuran 7x12 cm. bentuk ini lah yang paling banyak digunakan. Katalog yang berbentuk kartu yang telah tersusun secara sistematis dalam laci-laci katalog dapat menerima entri-entri baru tanpa merubah yang ada.

2. Katalog berbentuk lembaran-lembaran lepas, kemudian dibendel(dijilid) menjadi satu atau bebrapa berkas setelah disusun menurut system tertentu.
3. Katalog berbentuk tercetak setelah uraian-uraian katalog disusun menurut system tertentu, kemudian dicetak menjadi semacam bibliografi sebanyak yang diperlukan, kelebihan bentuk ini ialah katalog dapat diperbanyak dan dibawa kemanamana. Tetapi kelemahannya tidak dapat menerima entri-entri baru. Ini berarti entri baru harus disusun dan dicetak sebagai suplemen.
4. Katalog elektronik, bentuk katalog ini muncul berkat kemajuan dibidang teknologi informasi seperti computer. Dalam hal ini katalog berada dalam suatu basis data dikomputer, sehingga tidak perlu lagi diadakan penyusunan dengan sistematika tertentu seperti bentuk lainnya. Kelebihan katalog bentuk ini adalah lebih cepat dan mudah diakses, menghemat tenaga dan biaya dalam pembuatannya, dan entri-entri baru dapat dimaksud setiap saat. Kelemahannya ialah jikalau listrik padam, maka tidak bisa dipergunakan. Selanjutnya, dari katalog computer ini kemudian bisa diformat dalam bentuk CD-ROM. Kelebihannya ia bisa dibawa kemana-mana, tetapi untuk mengaksesnya tetap saja diperlukan perangkat computer.

3. ANALISA DAN PEMBAHASAN

Algoritma Horspool hampir mirip dengan algoritma Boyer-Moore, pergeseran pada Algoritma Horspool adalah:

1. Menggunakan pencocokan pattern indeks terakhir dengan text.
2. Jika pattern indek terakhir terjadi kesamaan dengan text maka akan dilakukan pengecekan untuk indek pattern sebelumnya.

Hal itu dilakukan secara terus menerus sampai ditemukanya pattern pada teks.

Analisis Hasil Pencarian Katalog Buku:

Pada running time text yang diinputkan adalah “Jogiyanto. 2005. Analisis Perancangan dan Desain. Yogyakarta: Andi.”. Dengan pattern sebagai berikut:

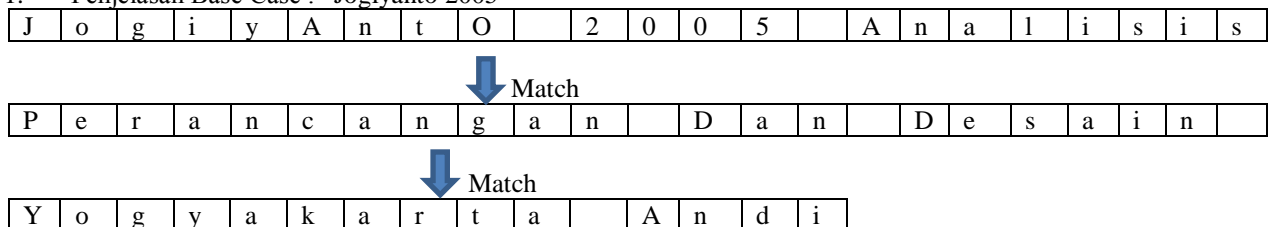
Untuk best case = “Jogiyanto 2005”

Untuk average case = “Analisis Perancangan Dan Desain”

Untuk worst case = “Yogyakarta Andi”

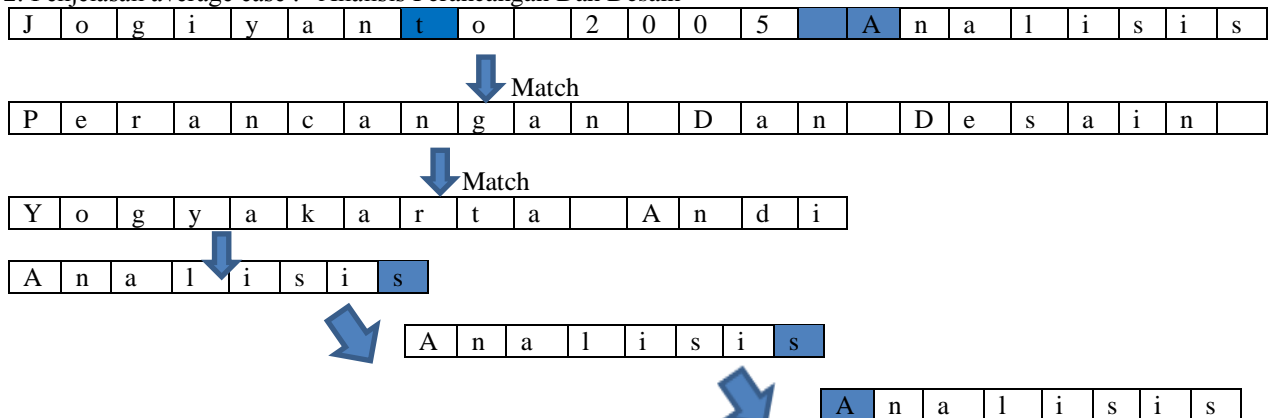
Setiap case akan dirunning sebanyak 10 kali, sehingga didapatkan 10 data running time. Pada akhir data akan disediakan rata-rata dari 10 running time, berikut penjelasan dari setiap case:

1. Penjelasan Base Case : “Jogiyanto 2005”



Pada kasus ini tidak ada pergeseran

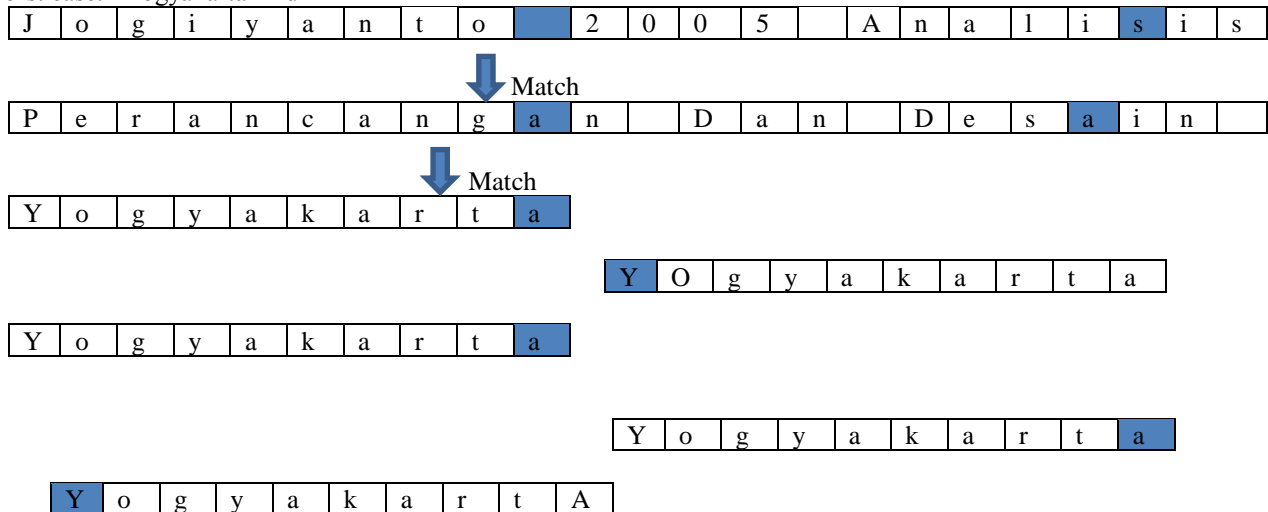
2. Penjelasan average case : “Analisis Perancangan Dan Desain”



Terjadi 3 kali pergeseran pattern pada kasus di atas.

3. Penjelasan W

orst case: “Yogyakarta Andi”



Pada kasus ini terjadi 12 pergeseran pattern, yang mana apabila tidak ada kecocokan pada pattern indeks terakhir maka langsung di lanjutkan pencocokan ke indeks teks selanjutnya sampai terjadi matching.

4. KESIMPULAN

Berdasarkan pembahasan diatas dapat disimpulkan bahwa Penerapan Algoritma horspool sangatlah mudah dan sama halnya seperti Algoritma Boyer Moore hanya berbeda dipergeseran patternnya saja. Jika Boyer-Moore menggunakan dua metode praproses *bad character shift* dan *good shufix shift*, akan tetapi Horspool hanya menggunakan satu metode praproses yaitu *bad character shift*. Sehingga algoritma ini sangat gampang dikerjakan jika sudah memahaminya.

Algoritma string matching pada penelitian ini menggunakan satu algoritma Horspool untuk membandingkan kinerja dengan algoritma string matching sehingga harus menambahkan fasilitas dalam pencarian kalimat.

REFERENCES

- [1] Fazar Ikhwan Guntara, S.Kom (2018). Implementasi Algoritma Horspool pada Aplikasi Taksonomi Berbasis Android p-ISSN: 2527-4856
- [2] Ir. Abdul Rahman Sale, Dip.Lib.,M.Sc. dan Ir. Janti G. Sujana, MA (2009). Pengantar Kepustakaan. Jakarta: Sagung Seto. hlm. 55-59. ISBN 9789793288680.
- [3] Sulistyio-Basuki (2004). Pengantar Dokumentasi. Bandung: Rekayasa Sains. hlm. 139. ISBN 9799747813.