

Penerapan Pencocokan String pada Pencarian Akun di Media Sosial Instagram

Radiyya Dwisaputra/13515023
Program Studi Teknik Informatika
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia
13515023@std.stei.itb.ac.id
radiyyasaputra@gmail.com

Abstract—Ada banyak media sosial yang terdapat di internet. Salah satunya adalah Instagram, media sosial yang menyediakan fitur berbagi foto dan video. Pengguna Instagram dapat membuat akun Instagram yang memiliki banyak atribut, diantaranya adalah username dan nama. Username adalah identitas utama dari sebuah akun. Untuk kemudahan penggunaannya, Instagram menyediakan fitur *explore*, yaitu fitur dimana pengguna Instagram dapat mencari akun lain dengan hanya menuliskan nama atau username dari akun lain tersebut. Algoritma pencarian nama atau username ini menggunakan algoritma pencocokan string.

Keywords—string, KMP, algoritma, Instagram, username

I. PENDAHULUAN

Seiring dengan berkembangnya teknologi, berbagai sektor turut berkembang. Salah satu sektor tersebut ialah kegiatan sosial. Kegiatan sosial saat ini bukanlah suatu hal yang harus dilakukan secara tatap muka atau secara langsung. Dengan adanya internet memudahkan antar individu untuk melakukan interaksi sosial walaupun tidak bertemu secara langsung. Interaksi sosial ini banyak dilakukan orang karena biayanya yang murah, serta kemudahan yang ditawarkan. Selain itu, interaksi sosial dengan internet dapat mendekatkan yang jauh, selama orang-orang tersebut terhubung dengan internet.

Internet adalah sebuah jaringan yang menghubungkan jutaan, atau bahkan milyaran komputer, maupun *device* (seperti *smartphone*) yang ada di dunia. Penggunaan internet sendiri makin banyak digunakan, seiring dengan teknologi yang berkembang. Begitu pula dengan kecanggihan, maupun fitur yang ada di internet. Pada awal diciptakan, internet hanya digunakan untuk kepentingan militer dan pendidikan. Hingga saat ini, internet digunakan untuk banyak hal, seperti *e-commerce* (belanja *online*), interaksi sosial, media hiburan, dan lain-lain.

Interaksi sosial *via* internet ini dilakukan dengan media sosial, atau juga yang sering disebut dengan jejaring sosial. Media sosial ini, menghubungkan orang-orang di dunia dengan berbagai fitur. Beberapa fitur tersebut diantaranya adalah *chatting* (interaksi teks), *photo sharing* (berbagi foto), *blogging* (interaksi blog), *calling* (interaksi suara), *video call* (interaksi

video), *video sharing* (berbagi video), dan lain-lain. Dalam menggunakan media sosial ini, seseorang harus memiliki *device* yang *online* (terhubung dengan jaringan).

Banyak situs saat ini yang merupakan media sosial, bahkan terdapat situs media sosial yang sudah memiliki anggota aktif sebanyak 2 miliar. Yang mana merupakan jumlah yang sangat besar dibandingkan jumlah penduduk Indonesia. Sosial media memiliki perkembangan yang sangat luar biasa. Hal ini terlihat dari situs media sosial pertama yakni *sixDegrees.com* yang memiliki fitur profil dan menambahkan teman dan belum begitu dikenal masyarakat. Media sosial mulai banyak dikenal masyarakat luas ialah saat media sosial Facebook yang didirikan oleh Mark Zuckerberg muncul. Dengan suksesnya media sosial Facebook yang ada menyebabkan timbulnya berbagai media sosial hingga saat ini dengan memiliki fitur yang sangat banyak. [1]

Salah satu media sosial *photo sharing* yang banyak digunakan adalah Instagram. Instagram adalah media sosial untuk berbagi foto, dan kini memiliki pengguna yang jumlahnya hingga ratusan juta. Instagram memiliki banyak sekali fitur yang kian ditambahkan oleh pengembangnya. Fitur-fitur baru Instagram dinilai mematikan media sosial sejenis lainnya, karena beberapa fitur Instagram ini memiliki banyak kemiripan dengan fitur utama media sosial sejenis.

Dalam beberapa media sosial, *username* yang digunakan tidaklah boleh untuk sama karena username disini berguna sebagai id pengguna dari media sosial tersebut. Oleh karena itu setiap pembuatan akun media sosial perlu dilakukan pengecekan username terlebih dahulu agar tidak ada username sama yang terdaftar. Tetapi dalam pembuatan media sosial bisa saja terdapat tampilan nama yang sama sedangkan username sebuah media sosial harus bersifat unik.

II. TEORI DASAR

A. Pencocokan String

Pencocokan string adalah cara untuk mencari sebuah teks

besar, dengan parameter sebuah string. Cara ini sangat bermanfaat dan banyak digunakan untuk berbagai bidang yang ada di dunia. Pencocokan string digunakan terutama untuk pencarian pada sebuah search engine seperti google atau semacamnya.

Pencocokan string dapat digunakan untuk mencari sebuah pola dari teks yang sudah ada. Pola tersebut akan dicari apakah ditemukan dalam teks yang sudah ada atau tidak. Pencocokan string sendiri tidaklah hanya menggunakan satu algoritma saja melainkan ada berbagai algoritma yang dapat digunakan yakni algoritma Brute force, algoritma Knuth Morris Path, dan algoritma Boyer Moore. Ketiga algoritma memiliki langkah-langkah yang berbeda dalam pencarian sebuah pola pada teks yang ada.

Pada bagian ini akan dijelaskan secara singkat mengenai langkah-langkah yang digunakan pada setiap algoritma

1. Algoritma Brute Force

Pada algoritma brute force pola diperiksa pada teks dengan cara memeriksa tiap karakter yang ada mulai dari karakter pertama pola hingga terakhir. Algoritma Brute Force merupakan algoritma yang paling sederhana dalam pencocokan string tetapi dapat menimbulkan kompleksitas yang sangat besar.

Langkah-langkah Algoritma Brute Force :

- Pencocokan karakter pertama pada pola dan teks
- Apabila karakter sama, maka karakter berikutnya pada pola akan diperiksa dengan karakter berikutnya pada teks.
- Apabila karakter tidak sama, maka karakter pertama akan dimajukan satu karakter untuk diperiksa dengan karakter berikutnya pada teks.
- langkah b dan c dilakukan secara terus menerus hingga ditemukan pola pada teks.
- bila hingga akhir teks tidak ditemukan pola maka pencarian berhenti

//Contoh

```
ANDREW
REW
REW
REW
REW
```

```
//Bruteforce untuk mendapatkan indeks
ditemukannya pola
//P[] : string pola yang akan dicari
//T[] : string teks
//m : panjang pola
//n : panjang teks
For i←0 to n-m do
  j←0
  while j<m and P[j]=T[i+j] do
    j←j+1
  if j = m return i
```

```
return -1
```

2. Algoritma Knuth Morris Path

Pada algoritma Knuth Morris Path atau biasa disingkat menjadi algoritma KMP memiliki prinsip yang mirip dibandingkan dengan algoritma brute force yakni pengecekan dilakukan secara terurut dari kiri ke kanan. Yang membedakan ialah KMP menggunakan prefix dan suffix pada pola untuk melakukan pencocokan string pola pada teks. Panjang prefix dan suffix yang sama pada sebuah pola akan dicatat yang biasa disebut dengan fungsi pinggiran. Fungsi pinggiran ini akan digunakan dalam pencocokan string dengan KMP.

Langkah-langkah algoritma KMP :

- Pencocokan karakter pertama pada pola dan teks
- Apabila karakter sama, pengecekan maju ke karakter berikutnya untuk diperiksa karakter berikutnya pada pola dengan karakter berikutnya pada teks.
- Apabila karakter tidak sama, ambil fungsi pinggiran dari karakter sebelumnya lalu tambahkan dengan 1 itulah karakter berikutnya yang akan diperiksa.
- langkah b dan c diulangi terus menerus hingga diperoleh pola pada teks.
- bila hingga akhir teks tidak ditemukan pola maka pencarian berhenti

//Contoh

```
abacaabaccabacabaabb
abacab
  abacab
    abacab
      abacab
```

//Pseudocode KMP

```
//Fungsi kmpMatch
Function kmpMatch (
  Input P : string pola
  Input T : string teks
  Output X : integer
)
```

```
i←0
```

```
j←0
```

```
fail[] ← ComputeFail(pola);
```

```
//m = panjang pola
```

```
//n = panjang teks
```

```
while (i<n)
```

```
  if(P[j]=T[i])
```

```
    if(j=m-1)
```

```
      return i-m+1
```

```
    i←i+1
```

```

    j←j+1
else if(j>0)
    j←fail[j-1]
else
    i←i+1

return -1

//fungsi computeFail

Function computeFail(
    Input P : String pola
    Output fail : array of integer
)
int fail[0] ← 0
j←0
i←0

while (i<m)
    if(P[j]=P[i])
        fail[i]←j+1
        i←i+1
        j←j+1
    else if (j>0)
        j ← fail[j-1]
    else
        fail[i] ← 0
        i←i+1

return fail

```

Berdasarkan langkah-langkah pengecekan diatas dapat dilihat bahwa algoritma KMP dapat mempercepat waktu pencarian karena tidak harus memeriksa seluruh karakter melainkan hanya dilakukan pengecekan pada karakter yang masih mungkin saja.

3. Algoritma Boyer moore

Algoritma Boyer moore memiliki langkah yang lebih kompleks dibandingkan dengan algoritma sebelumnya karena terdapat 3 kasus dalam setiap langkah yang dilakukan. Pencocokan string pada algoritma Boyer moore dilakukan dari pola kanan ke kiri. Algoritma Boyer moore banyak digunakan untuk mencocokkan string pola tertentu pada teks yang panjang. Algoritma Boyer moore juga memiliki fungsi seperti halnya fungsi pinggiran pada KMP hanya saja fungsi pada boyer moore ialah fungsi untuk menunjukkan indeks terakhir setiap karakter muncul pada pola. Fungsi ini akan digunakan saat pencocokan string.

Langkah – langkah algoritma Boyer moore :

- Pencocokan karakter terakhir pada pola dan teks
- Apabila karakter sama, pengecekan mundur ke karakter sebelumnya untuk diperiksa karakter sebelumnya.
- Apabila karakter tidak sama terdapat 3 kasus yakni :

- o Jika karakter yang dibandingkan pada teks ada pada pola di sebelah kiri karakter yang sedang diuji maka pola akan dimajukan sesuai karakter tersebut.
 - o Jika karakter yang dibandingkan pada teks ada pada pola di sebelah kanan karakter yang sedang diuji maka pola yang diujikan maju sebanyak 1 karakter.
 - o Jika karakter yang dibandingkan pada teks tidak ada pada pola maka pola yang diujikan maju sesuai jumlah karakter
- langkah b dan c diulangi terus menerus hingga diperoleh pola pada teks.
 - bila hingga akhir teks tidak ditemukan pola maka pencarian berhenti

```

//Contoh
abacaabadcabacabaabb
abacab
abacab
abacab
abacab
abacab
abacab

```

```

//Pseudocode Boyer moore
//fungsi bmMatch

```

```

Function bmMatch (
    Input P : string pola
    Input T : string teks
    Output X : integer
)

//n adalah panjang teks
//m adalah panjang pola

Last[] ← buildLast(pola)
i ← m-1

if (i > n-1)
    return -1

j ← m-1
do
    if (P[j] = T[i])
        if (j = 0)
            return i
        else
            i←i-1
            j←j-1
    else
        lo ← last(T[i])

```

```

i ← i+m-min(j,1+lo)
j ← m-1
while (i<= n-1)

return -1

//fungsi buildLast
Function buildLast(
  Input P : String pola
  Output array : integer[]
)

Last[] ← int[128]

For i=0 to i<128 do
  Last[i] ← -1
For i=0 to i<m
  Last[P(i)]=i
Return last;

```

[4]

B. Instagram

Dikutip dari kamus yang dikembangkan oleh Cambridge University, instagram adalah “the name of a social networking service for taking, changing, and sharing photographs and video”. Instagram adalah media sosial yang digunakan untuk mengunggah foto dan video. Foto dan video tersebut mempunyai atribut *likes* dan *comments*. Pada instagram, seorang pengguna memiliki linimasa pribadi. Akun pada instagram memiliki beberapa atribut diantaranya ialah username, nama, foto, bio, *website*, *email*, jenis kelamin, *followers*, *following*, nomor *handphone*, *display picture*, serta status privasi dari sebuah akun. Semua atribut tersebut disimpan dalam sebuah basis data. *Followers* adalah akun-akun lain yang dapat melihat foto yang diunggah oleh sebuah akun, yang mana foto-foto tersebut dapat dilihat pada linimasa pribadi akun-akun lain yang menjadi *followers* dari akun tersebut. *Following* adalah akun-akun yang di-*follow* oleh sebuah akun. Status privasi sebuah akun adalah sebuah penentu, dimana akun yang *private*, adalah akun yang unggahannya tidak dapat dilihat oleh akun yang tidak menjadi *followers*-nya. Sedangkan akun yang memiliki status *public* adalah akun yang unggahannya dapat dilihat oleh semua akun lain.

Username pada instagram haruslah merupakan sesuatu yang unik sehingga tidak ada lebih dari satu akun yang memiliki username yang sama. Username hanya dapat terdiri dari kombinasi angka, huruf, *underscore* dan titik. Sedangkan nama pada sebuah akun instagram tidak memiliki batasan dan tidak harus unik sehingga bisa lebih dari satu akun memiliki nama yang sama.

Instagram memiliki fitur *explore* untuk kemudahan penggunaannya. Fitur *explore* ini menampilkan foto dan video berdasarkan jenis foto dan video yang sering dilihat oleh



Gambar 1 Contoh tampilan profile di Instagram

pengguna. Foto dan video ini juga ditampilkan berdasarkan akun lain yang di-*follow* oleh pengguna. Selain menampilkan foto dan video, fitur *explore* ini juga dapat mencari foto berdasarkan *hashtags* (simbol tagar yang digunakan untuk mencari foto), dan akun lain berdasarkan nama atau *username*.



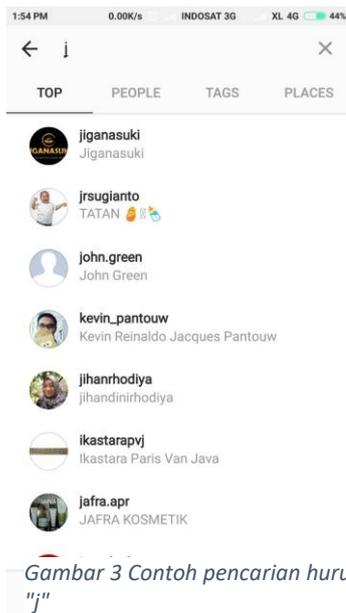
Gambar 2 Contoh tampilan explore di instagram

Pendaftaran akun di instagram tidak membutuhkan biaya sedikitpun atau gratis, sehingga instagram dipakai oleh sangat banyak orang. Hingga saat ini, terdapat lebih dari 600 juta pengguna dari seluruh dunia. Untuk di Indonesia sendiri, terdapat lebih dari 22 juta pengguna aktif instagram. [2] & [3]

III. CARA KERJA SEARCHING INSTAGRAM

Untuk pencarian akun di instagram terdapat sebuah kolom pencarian yang terdapat pada bagian atas pencarian dan pencarian bersifat *real time*. Pencarian diawali pada following/follower kita terlebih dahulu lalu dilanjutkan pada akun-akun *verified* dan selanjutnya akun-akun lain yang sesuai dengan kata pada pencarian.

Instagram akan menampilkan hasil pencarian setiap karakter yang ditambahkan. Instagram menyimpan akun yang sering dikunjungi sehingga membuat hasil pencarian diharapkan sesuai dengan keinginan pengguna. Instagram mempelajari aktivitas yang kita lakukan dalam instagram saat melihat akun-akun instagram tertentu.



Gambar 3 Contoh pencarian huruf "j"

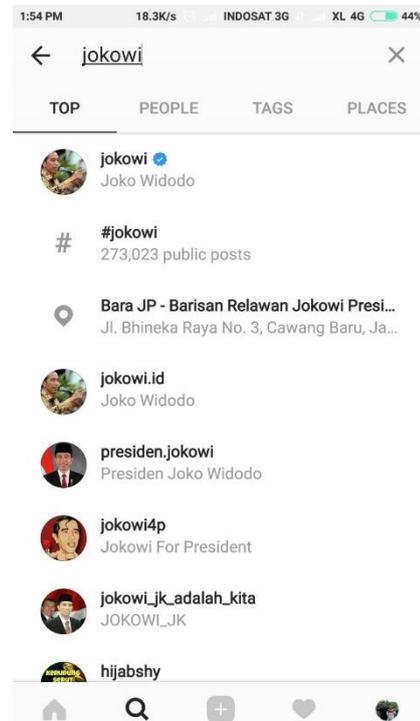
Melalui gambar di atas terlihat bahwa pencarian pada instagram menampilkan akun dengan username atau nama yang sesuai.

Algoritma pencocokan string pada *searching* akun di Instagram dengan menggunakan KMP akan dijelaskan dengan pseudocode dibawah ini.

```
// pseudocode Searching Instagram
k ← 0
While(not Empty(searchbox))
  //lakukan pencarian
  For i=0 to i<jumlahuser do
    j ← kmpMatch(masukanuser,username[i])
    if(j!=-1)
      Display(Profile[i])
      k ← k+1

if(k=0)
```

Display("No results found")



Gambar 4 Contoh pencarian string "jokowi"

IV. ANALISIS SEARCHING INSTAGRAM

Algoritma dasar *searching* yang digunakan pada instagram ialah dengan algoritma KMP hal ini dikarenakan jumlah karakter yang dimaksudkan dalam *search box* tidak terlalu banyak yakni maksimal 30 karakter. Selain itu hal ini dipilih karena KMP memiliki kompleksitas yang kecil. Media sosial yang memiliki nama setinggi instagram tentunya harus memberikan fasilitas terbaik untuk penggunaanya, dan tentunya hal ini menuntut instagram untuk mempersingkat waktu *searchingnya*. Alasan-alasan tersebut menunjukkan bahwa algoritma dasar *searching* yang paling tepat untuk pencarian username di instagram adalah algoritma KMP.

Algoritma KMP hanyalah algoritma dasar hal ini ditunjukkan dengan hasil pencarian yang mendasar pada instagram yaitu akun-akun yang sering dikunjungi oleh pengguna akan muncul terlebih dahulu dibandingkan akun-akun yang jarang dikunjungi. Hal ini dilakukan instagram untuk memudahkan pengguna memperoleh hasil pencarian yang diinginkan.

Hal-hal yang membuat Instagram dapat melakukan hal seperti itu ialah dengan adanya *machine learning* yang membuat Instagram dapat mempelajari perilaku yang dilakukan oleh pengguna.

V. KESIMPULAN

Instagram sebagai media sosial ternama, menggunakan algoritma dasar searchng KMP untuk fitur pencarian akun dengan parameter *username* atau nama di halaman *explore*-nya. Hal ini menunjukkan bahwa penggunaan pencocokan string sangat luas dalam kehidupan nyata.

VI. REFERENSI

- [1] <https://kumparan.com/aditya-panji/jumlah-pengguna-aktif-facebook-hampir-2-miliar> diakses pada 17 Mei 2017 10.00
- [2] <http://tekno.kompas.com/read/2016/12/16/09060027/pengguna.instagram.m.tebus.600.juta> diakses pada 18 Mei 2017 20.00
- [3] <http://www.cnnindonesia.com/teknologi/20160623112758-185-140353/ada-22-juta-pengguna-aktif-instagram-dari-indonesia/> diakses pada 18 Mei 2017 20.30
- [4] Munir, Rinaldi. 2003. *Matematika Diskrit Edisi Kedua*. Bandung: Penerbit Informatika.

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, 18 Mei 2017



Radiyya Dwisaputra - 13515023