Membandingkan Pengaruh Panjang dan Besar Varian Karakter terhadap Entropi Password Menggunakan Algoritma Brute Force

Husnulzaki Wibisono Haryadi / 13515005
Program Studi Teknik Informatika
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia
13515005@std.stei.itb.ac.id

***Abstrak*—Password merupakan aspek penting dari kehipan sehari – hari manusia di era digital. Password mengamankan hak akses eksklusif atas informasi – informasi penting yang disimpan oleh pemiliknya secara digital. Oleh karena itu password selain bersifat rahasia, harus pula kuat agar tidak rentan terkena serangan brute force. Namun manakah yang lebih utama, panjang atau jumlah variasi kata dalam menentukan kekuatan sebuah sebuah password? Makalah ini akan membahas hubungan antara tingkat keacakan (entropi) password, relasinya dengan kekuatan password, dan bagaimana panjang dan banyak kemungkinan karakter mempengaruhinya.**

***Kata Kunci*—entropi password, panjang string, karakter potensial string, keamanan digital.**

I. PENDAHULUAN

1.1 Kemudahan Akses Informasi di Era Digital dan Dampaknya dalam Kehidupan Manusia

Dengan semakin pesatnya perkembangan zaman, kebutuhan akan akses informasi yang lebih cepat juga meningkat. Kebutuhan inilah yang mendorong teknologi informasi untuk berkembang, demi meningkatkan kemampuan manusia untuk saling berbagi informasi dan berkomunikasi. Teknologi informasi memungkinkan manusia untuk menjangkau kemanapun di seluruh belahan dunia dalam waktu yang singkat. Kemudahan akses atas informasi digital mendorong semakin banyak pihak dari berbagai latar belakang untuk melakukan digitalisasi atas data – data penting yang mereka butuhkan sehari – hari, mulai dari data yang terkait pekerjaan hingga informasi pribadi. Dengan adanya teknologi informasi, manusia dapat bekerja secara lebih efisien dan hidup lebih nyaman.

Namun, seiring dengan semakin berkembangnya teknologi untuk menyimpan dan mengakses berbagai informasi secara digital, terdapat pula dampak buruk yang mengiringinya. Kemudahan akses membuat informasi – informasi penting, seperti informasi yang bersifat privat, semakin mudah untuk diakses oleh pihak – pihak yang tidak bertanggung jawab. Informasi yang telah didapatkan oleh pihak – pihak ini selanjutnya dapat disalahgunakan, seperti dimanipulasi atau dimanfaatkan untuk melaksanakan penipuan. Hal ini tentu saja sangat berbahaya dan merugikan bagi pemilik informasi. Penyelesaian terbaik atas permasalahan ini adalah dengan melakukan penyandian atau enkripsi terhadap informasi yang akan disimpan secara digital. Melalui proses enkripsi, informasi digital yang awalnya bermakna dapat diubah menjadi data tersandikan yang tidak memiliki makna. Data ini akan tetap dalam bentuk demikian kecuali orang yang melakukan penyandian mendekripsi data tersebut menjadi informasi yang berarti kembali. Dengan demikian, apabila data tersebut jatuh ke tangan pihak yang salah, mereka tidak serta merta dapat mengeksploitasi data tersebut. Dengan adanya prosedur enkripsi data, privasi dan keamanan atas informasi digital tiap individu menjadi terjamin kembali.

1.2 Password sebagai Unsur Penting dalam Enkripsi Data

Dengan adanya konsep enkripsi sebagai langkah pencegahan terhadap pencurian dan penyalahgunaan informasi digital, dibutuhkan sebuah metode yang dapat menjamin agar pada saat yang dibutuhkan, pemilik informasi dapat mengakses dan melakukan dekripsi untuk mendapatkan informasi yang ia butuhkan. Dengan kata lain, setiap pemilik informasi yang berniat menyandikan informasinya harus memiliki kredensial yang menandakan bahwa ia benar – benar pemilik informasi yang sah, sedangkan tiap prosedur enkripsi harus memiliki metode untuk mendekteksi dan mengenali kredensial sebelum menjalankan fungsinya.

Disinilah password muncul. Password berperan sebagai salah satu wujud kredensial yang berupa *string* atau serangkaian karakter yang memungkinkan sebuah entitas

(baik itu sebuah individu maupun instansi tertentu) untuk dapat dikenali sebagai pemilik yang sah atas informasi tertentu. Dengan demikian, bagi pemiliknya, password merupakan hal yang sangat penting dan seringkali menjadi satu – satunya jembatan antara pemilik informasi dan informasinya.

1.3 Sifat dan Atribut dalam Sebuah Password

Layaknya gembok yang mencegah siapapun selain pemilik kunci untuk membuka pintu, enkripsi mencegah siapapun selain pemilik password untuk memanfaatkan informasi yang telah ia sandikan. Seperti layaknya gembok dan kunci pula, konsep password pada penyandian data bukan ada tanpa cela. Kemampuan sebuah password untuk menjaga hak akses atas informasi terbatas pada fakta bahwa pemilik informasi harus menjadi satu – satunya entitas yang mengetahui password tersebut. Dengan kata lain sebuah password harus bersifat rahasia agar dapat berfungsi secara efektif.

Aturan ini pun masih memiliki celah. Walaupun pemilik password sendiri telah merahasiakan passwordnya, seseorang yang berniat mencuri informasi yang terkunci oleh password dapat mencari tahu informasi – informasi terkait password itu sendiri, seperti panjang karakter dan karakter apa saja yang mungkin terkandung di dalamnya. Berbekal informasi ini, seseorang tersebut dapat menebak satu persatu kemungkinan susunan karakter pada password. Metode ini disebut dengan istilah *brute force*. Apabila pemilik password memilih password yang memiliki komposisi yang sederhana, dengan waktu dan sedikit keberuntungan, pelaku brute force dapat menebak string dari password. Dari ilustrasi ini dapat disimpulkan bahwa selain bersifat rahasia, password juga harus bersifat rumit agar sulit ditebak. Ada dua atribut utama yang menentukan seberapa rumitnya sebuah password, yaitu banyaknya kemungkinan karakter dan panjang password itu sendiri. Ukuran seberapa sulitnya sebuah password untuk ditebak itu sendiri disebut dengan istilah password entropy.

1.4 Hubungan antara Panjang Karakter, Banyak Kemungkinan Karakter, dan Entropi Sebuah Password

Dari pemaparan inilah muncul pertanyaan hipotetis yang akan menjadi pokok bahawan makalah ini : "Manakah yang lebih kuat, sebuah password sepanjang 8 karakter yang memiliki kemungkinan untuk mengandung seluruh karakter ASCII atau sebuah password 10 karakter yang hanya mengandung kemungkinan huruf besar dan huruf kecil saja." Tentunya akan nada argumen yang menyatakan bahwa password yang mengandung lebih banyak karakter, walaupun lebih pendek, akan lebih kuat dari pada password yang panjang, dan ada pula argumen yang akan menyatakan sebaliknya.

"Kekuatan password" yang dipersoalkan disini berkaitan sangat erat dengan entropi password tersebut. Panjang dan besar variasi karakter adalah atribut – atribut dari sebuah password yang secara langsung mempengaruhi entropinya. Dalam makalah ini, hubungan antara kedua atribut tersebut dan nilai entropi sebuah password akan diperbandingkan dan di visualisasikan. Makalah ini bermaksud untuk memberikan gambaran terkait seberapa besar pengaruh kedua atribut tersebut terhadap kekuatan password secara keseluruhan.

II. DASAR TEORI

2.1 Brute Force

Brute Force Attack adalah metode untuk meretas password (password cracking) dengan cara mencoba semua kemungkinan kombinasi yang ada pada "wordlist". Metode ini dijamin akan berhasil menemukan password yang ingin diretas. Namun, proses untuk meretas password dengan menggunakan metode ini akan memakan banyak waktu. Lamanya waktu akan ditentukan oleh panjang dan kombinasi karakter password yang akan diretas.

Brute Force Attack menggunakan formula sebagai berikut:

$$KS = L(m) + L(m+1) + L(m+2) + \ldots + L(M)$$

Keterangan:

L = Jumlah karakter yang kita ingin definisikan
M = Panjang maksimum kata kunci
m = Panjang minimum kata kunci

Istilah Brute Force sendiri di populerkan oleh Kenneth Thomson, dengan mottonya "When in doubt, use brute-force" (jika rugu, gunakan brute-force).

Brute Force dapat digunakan untuk meretas password secara offline maupun online, namun kombinasi karakter password yang panjang terkadang membuat lama waktu pemecahan menjadi terlalu lama sehingga brute force dinyatakan tidak efisien.

2.2 Password

Password merupakan sederet karakter yang membuat informasi penting untuk melakukan proses autentikasi, yaitu proses sistem untuk memastikan bahwa orang yang mengakses sistem tersebut adalah orang yang sebenarnya (bukan orang lain). Password biasanya bersifat statis, artinya password tidak akan berubah sampai seseorang mengubahnnya sendiri. Password sifatnya amat rahasia dan tidak boleh diberitahukan kepada orang lain, karena

jika orang lain sampai mengetahuinya, maka dia akan mendapatkan akses ke sebuah informasi. Misalnya, jika informasi itu adalah dokumen-dokumen penting negara atau account di sebuah bank, tentunya akan terjadi hal yang tidak diinginkan dan akan terjadi hal yang amat sangat fatal jika orang lain mengetahuinnya.

Sistem keamanan akan membandingkan kode-kode yang dimasukkan oleh pengguna (yang terdiri atas nama pengguna/user name dan password) dengan daftar atau basis data yang disimpan oleh sistem keamanan sistem atau jaringan tersebut (dengan menggunakan metode autentikasi tertentu, seperti halnya kriptografi, hash atau lainnya). Jika kode yang dibandingkan cocok, maka sistem keamanan akan mengizinkan akses kepada pengguna tersebut terhadap layanan dan sumber daya yang terdapat di dalam jaringan atau sistem tersebut, sesuai dengan level keamanan yang dimiliki oleh pengguna tersebut. Idealnya, kata kunci merupakan gabungan dari karakter teks alfabet (A-Z, a-z), angka (0-9), tanda baca (!?,.=-) atau karakter lainnya yang tidak dapat (atau susah) ditebak oleh para intruder sistem atau jaringan. Meskipun begitu, banyak pengguna yang menggunakan kata sandi yang berupa kata-kata yang mudah diingat, seperti halnya yang terdapat dalam kamus, ensiklopedia (seperti nama tokoh, dan lainnya), atau yang mudah ditebak oleh intruder sistem.

Pada makalah ini, banyaknya karakter yang terkandung dalam suatu password akan disebut dengan istilah "panjang password", sedangkan banyaknya kemungkinan karakter yang dapat terkandung di dalam password disebut dengan istilah "banyak varian karakter" pada password.

2.3 Password Entropy

Password entropy adalah ukuran yang dapat dipergunakan untuk menghitung kekuatan sebuah password secara kuantitatif.

Secara umum, kekuatan sebuah password terukur dari seberapa rumitnya password tersebut. Kerumitan ini berdampak pada banyak rata – rata percobaan yang harus dilakukan untuk menebak string password tersebut. Kerumitan password, menurut Claude E. Shannon (1984), dapat ditentuktan dengan menghitung entropi dari password tersebut, dimana entropi memiliki rumus :

$$H = -\sum_{i=1}^{n} p_i \bullet \log_2(p_i)$$

dimana H adalah entropi password, n adalah panjang password, dan pi adalah probabilitas terpilihnya karakter ke-i dalam string password.

Entropi akan menentukan banyaknya kemungkinan permutasi sebuat password. Banyak kemungkinan ini dapat dihitung dengan rumus :

$$N = 2^H$$

Dimana N merupakan jumlah kemungkinan dan H adalah besar entropi. Dapat ditarik kesimpulan bahwa meningkatnya H akan berdampak terhadap meningkatnya jumlah kemungkinan secara eksponensial.

III. PENGAMBILAN DATA

Dalam makalah ini, penulis bermaksud menguji lama proses pemecahan password dan membandingkan lama waktu pemecahan pada password dengan panjang dan banyak varian berbeda. Proses pengambilan data waktu pemecahan diambil sebanyak tiga kali dan diambil rata – ratanya.

Panjang karakter password yang diuji berada pada rentang 1 – 16 karakter. Rentang ini diambil untuk membandingkan password standard yang paling umum dipakai (8 karakter) dengan password yang memiliki panjang dua kali lipatnya.

Banyak varian karakter dalam password yang diuji berada pada rentang 1 – 95 karakter. Angka ini diambil untuk agar dapat menjaring sampel yang mengandung seluruh karakter ASCII.

Kedua atribut yang diuji memiliki jarak antar sampel sebesar 1 karakter. Angka ini diambil dengan harapan data yang diambil memberi hasil visualisasi yang detil.

3.1 Spesifikasi Program Penguji

Program yang dipergunakan dalam pengujian dan pengambilan data pada makalah ini adalah program sederhana yang dapat meng-*generate* password acak dan memecahkan password tersebut menggunakan metode brute force. Program dapat menghitung dan memberikan keluaran berupa lamanya waktu yang dibutuhkan untuk memecahkan tiap password.

Kode sumber program dapat diakses pada laman https://github.com/ayamberkakienam/SimplePasswordCracker.

A. Bahasa Pemrograman
   Program yang dipergunakan untuk memperoleh data percobaan ditulis dalam bahasa Java versi 1.5,

B. Struktur Program
   Program terdiri atas empat kelas yang semuanya tergabung di dalam satu *package*. Keempat kelas tersebut adalah:

   - Kelas Lock
     Kelas ini mengandung atribut password yang akan di-*generate* secara acak saat diinstantiasi. Panjang dan besar varian karakter dalam password ditentukan dalam parameter *constructornya*.

- Kelas LockException
  Kelas ini berfungsi menangani *error* yang mungkin terjadi apabila parameter dalam *constructor* diisi diluar nilai yang mampu ditangani oleh kelas.

- Kelas PasswordCracker
  Kelas ini bertujuan untuk memecahkan password yang terdapat pada kelas Lock. Objek Lock dapat dipecahkan oleh setiap objek PasswordCracker akan ditentukan saat instantiasi objek.

- Kelas Driver
  Kelas ini mengandung program utama dimana semua objek diinstantiasi. Apa saja jenis keluaran yang akan dihasilkan dan bagaimana format keluaran ditentukan dalam kelas ini.

C. Algoritma Program Utama
Program bekerja dengan alur sebagai berikut:

1. Program melakukan *assignment* objek Lock dan PasswordCracker.

2. Program membuat loop bersarang *(nested loop)*. Pada loop pertama, program mengiterasi variabel panjang karakter. Pada loop kedua, program mengiterasi variabel besar varian karakter.

3. Pada loop kedua, program menginstantiasi isi dari objek Lock sesuai dengan variabel yang telah ditetapkan sebelumnya. Program juga menginstanstiasi objek PasswordCracker sesuai dengan objek Lock yang ada.

4. Masih pada loop yang sama, program memanggil fungsi pemecahan password pada objek PasswordCracker. Program mencatat waktu mulai dan waktu selesai tiap percobaan pemecahan.

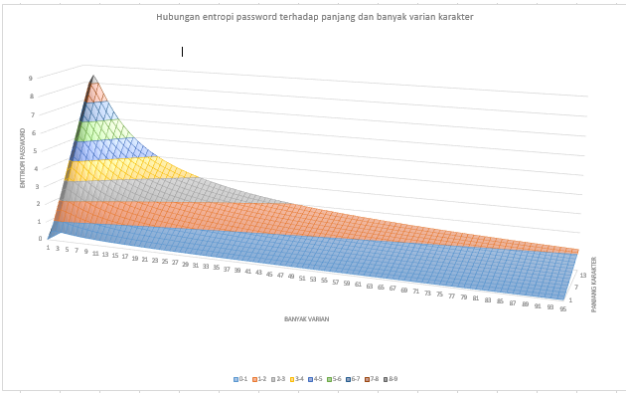
5. Program menampilkan lama waktu pemecahan password untuk tiap iterasi.

D. Keluaran
Program memberikan keluaran berupa matriks durasi pemecahan password dengan metode *brute force*. Kolom matriks berhubungan dengan banyaknya varian karakter dalam password sedangkan baris matriks berhubungan dengan jumlah karakter dalam password.

3.2 Data Hubungan antara Panjang dan Banyak Varian Karakter Terhadap Entropi Password

Varian \ Panjang karakter	4	6	8	10	12	15	16
5	1.86	2.79	3.72	4.64	5.57	6.97	7.43
10	1.33	1.99	2.66	3.32	3.99	4.98	5.32
15	1.04	1.56	2.08	2.60	3.13	3.91	4.17
20	0.86	1.30	1.73	2.16	2.59	3.24	3.46
30	0.65	0.98	1.31	1.64	1.96	2.45	2.62
40	0.53	0.80	1.06	1.33	1.60	2.00	2.13
50	0.45	0.68	0.90	1.13	1.35	1.69	1.81
60	0.39	0.59	0.79	0.98	1.18	1.48	1.58
70	0.35	0.53	0.70	0.88	1.05	1.31	1.40
80	0.32	0.47	0.63	0.79	0.95	1.19	1.26
95	0.28	0.41	0.55	0.69	0.83	1.04	1.11

Gambar 3.1 Tabel hasil perhitungan entropi berdasarkan panjang dan banyak varian karakter



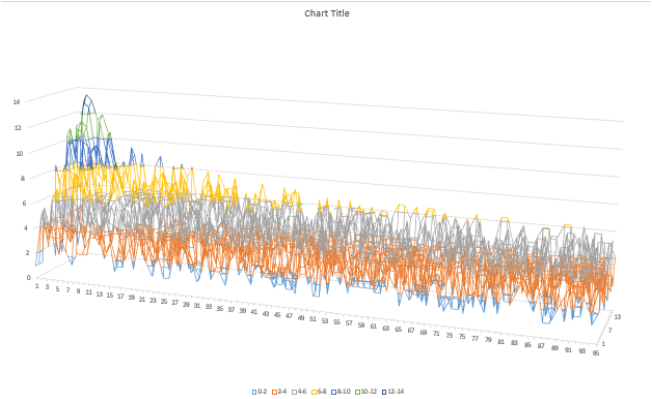
Gambar 3.2 Grafik hubungan antara panjang dan banyak varian karakter terhadap entropi password, dimana sumbu x merupakan banyak varian karakter, sumbu z merupakan panjang karakter, dan sumbu z merupakan besar entropi.

3.3 Data Hasil Pengujian Lama Waktu Pemecahan Password

	Panjang karakter					
	4	6	8	10	12	
5	28.99119	160	110.3784	152.2923	525.3056	6.9
10	422.6467	89.02037	36	194.1129	174.4436	4.9

15	12.16111	8.029185	169.6367	336	184.8657	3.91	4.17
20	90.50967	14.67206	76.10926	49.35075		160	3.24	3.46
30	42.48423	271.1594	103.842	44.18538	225.6137	2.45	2.62
40	40.19018	252.9822	191.0914	40.09498	302.8595	2.00	2.13
50	191.3279	285.5167	59.17692	294.3638	91.51588	1.69	1.81
60	18.31543	225.2936	110.8513	68.1777	10.48207	1.48	1.58
70	281.8135	64.36518	13.06736	509.3609	38.77873	1.31	1.40
80	17.00416	164.2523	198.3251	89.79978	90.35669	1.19	1.26
95	41.17696	19.7297	23.63342	141.5476	135.6434	1.04	1.11

Gambar 3.3 Tabel durasi pemecahan password berdasarkan panjang dan banyak varian karakter (dalam detik)



Gambar 3.4 Grafik hubungan antara panjang dan banyak varian karakter terhadap durasi pemecahan password

IV. PEMBAHASAN

4.1 Korelasi antara Hasil Pengujian dan Entropi Password

Dengan membandingkan grafik pada Gambar 3.2 dan Gambar 3.4 dapat terlihat bahwa hasil pengujian memiliki grafik yang identik dengan nilai entropi tiap – tiap sampel. Ini membuktikan bahwa metode pengujian yang digunakan relevan terhadap ruang pengamatan, yaitu untuk membandingkan lama waktu pemecahan password yang berbanding lurus dengan entropi password tersebut.

4.2 Korelasi antara Panjang dan Banyak Varian Karakter terhadap Waktu Pemecahan Password

Walaupun grafik membuktikan bahwa benar terdapat korelasi antara entropi dan lama waktu pemecahan, tetapi berdasarkan hasil yang ditunjukkan oleh Gambar 3.3, masih sulit untuk menarik kesimpulan berdasarkan sampel karena data yang didapat masih cukup acak. Hal ini kemungkinan diakibatkan oleh kurangnya jumlah pengambilan sampel, sehingga simpangan data masih cukup besar.

Hal yang menurut penulis menarik dari hasil pengujian adalah fakta bahwa pada panjang karakter yang tinggi, password dengan banyak varian rendah memiliki lama waktu pemecahan yang lebih tinggi daripada password dengan banyak varian yang lebih tinggi. Hal ini juga dapat terlihat dari Gambar 3.1, dimana entropi dari password dengan panjang karakter 16 dan besar varian 95 justru jauh lebih kecil dari password dengan varian 5 dan panjang yang sama.

Hal ini dapat dijelaskan oleh rumus entropi yang dikemukakan oleh Claude E. Shannon yang telah dijabarkan pada bab sebelumnya, dimana

V. KESIMPULAN

Dari hasil percobaan dan perbandingannya terhadap entropi password, dapat ditarik kesimpulan bahwa panjang sebuah password memiliki dampak lebih besar terhadap kekuatan password dibanding banyak variasi karakter yang terdapat di dalam password tersebut. Hal ini dikarenakan berdasarkan rumus entropi password, pertambahan panjang password akan menyebabkan meningkatnya jumlah kemungkinan permutasi password secara eksponensial.

Akan tetapi, dalam kenyataannya, akan sulit sekali mengaplikasikan password dengan panjang 16 karakter namun hanya memiliki variasi 2 – 4 karakter dalam kehidupan sehari – hari karena sulit bagi manusia untuk menghafalnya. Sulitnya menghafal password yang panjang menyebabkan manusia cenderung memilih pemakaian password dengan panjang sedang namun varian tinggi (contohnya password yang berisi nama panggilan, namun dengan tiap huruf vokal yang diganti dengan angka). Awamnya, pemilik password akan mengira jenis password seperti ini sebagai password yang kuat karena memang akan membingungkan bagi otak manusia untuk membacanya. Akan tetapi bagi komputer, jenis – jenis password yang hanya mensubstitusikan huruf dengan angka seperti ini justru mempermudah proses pemecahan dengan brute force. Seringkali pencuri password memanfaatkan heuristic dalam penulisan algoritma pemecah passwordnya sehingga password – password seperti ini justru semakin cepat terpecahkan.

Dengan demikian, penulis menghimbau bagi semua pengguna layanan digital untuk meningkatkan kualitas passwordnya, bukan dengan mengubah variasi huruf dan kata yang terdapat dalam password, tetapi dengan memperbesar ukuran password.

VI. SARAN

Dalam pengerjaan makalah ini, masih banyak kesulitan yang didapatkan dalam proses analisa dan penarikan kesimpulan terhadap data – data hasil uji. Hal ini dikarenakan proses pemecahan password menggunakan metode brute force memakan waktu yang sangat lama sehingga jumlah pengambilan sampel tidak dapat dilakukan sebanyak yang penulis targetkan pada awalnya. Hal ini berdampak besar pada kualitas data, sebab waktu

pemecahan password yang dihasilkan secara acak menggunakan metode brute force memiliki deviasi yang sangat besar, sehingga apabila tidak diimbangi dengan jumlah pengambilan sampel yang besar akan menghasilkan kumpulan data yang sulit terukur dan teranalisa. Atas alasan ini, penulis berpesan untuk mengalokasikan waktu yang besar apabila akan mempraktekkan metode pengujian seperti yang penulis lakukan.

DAFTAR PUSTAKA

[1] Munir Rinaldi, "Matematika Diskrit Rivisi Keenam", Informatika Bandung, September 2016

[2] Shannon, Claude E. "A Mathematical Theory of Communication", Bell System Technical Journal, Vol. 27, pp. 379–423, 623–656, 1948

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, 4 Mei 2017

Husnulzaki Wibisono Haryadi / 13515005