Aisyah Journal of Informatics and Electrical Engineering
Universitas Aisyah Pringsewu

Journal Homepage
http://jti.aisyahuniversity.ac.id/index.php/AJIEE

Optimasi Pencarian Alamat IP Menggunakan Counting Bloom Filter Pada Web Application Firewall

Dimaz Arno Prasetio[1], Ahmad Fauzi[2], Ema Utami[3], Kusnawi[4]
Magister Teknik Informatika Universitas Amikom Yogyakarta
Jl. Ring Road Utara, Condong Catur,Sleman, Yogyakarta, Indonesia
[1]dimaz.1234@students.amikom.ac.id, [2]ahmad.1238@students.amikom.ac.id
[3]eama.u@amikom.ac.id, [4]khusnawi@amikom.ac.id

Abstract

Proses pencarian IP pada Web Application Firewall menjadi sebuah tantangan ketika database IP yang dikumpulkan mulai membesar. WAF harus dapat dengan cepat memberikan keputusan apakah sebuah IP masuk kedalam kategori buruk (blacklisted ip) atau tidak. Berdasarkan penelitian-penelitian sebelumnya algoritma bloom filter dapat dengan mengurangi waktu pencarian secara signifikan, namun hal itu tidak cukup karena bloom filter memiliki kelemahan pada struktur datanya yang tidak dapat dilakukan penghapusan, berdasarkan masalah tersebut kami mengajukan pemodelan baru pencarian IP pada WAF dengan menggunakan algoritma Counting Bloom Filter.

Keyword - Pencarian IP, Counting Bloom Filter, WAF

1. PENDAHULUAN

Dengan semakin banyaknya user yang menggunakan sistem informasi untuk berbagai kebutuhan bisnis atau layanan publik, Untuk itu keamanan Informasi telah menjadi kebutuhan yang wajib dipenuhi oleh setiap perusahaan atau instansi yang mempunyai sumberdaya sistem informasi. Sistem yang terhubung secara global haruslah dijaga keamanannya dari para pengguna yang tidak bertanggung jawab.Banyak cara atau teknik serta metode untuk mengamankan sistem informasi, Salah satunya yaitu dengan penerapan Web Application Firewall (WAF).

Cara kerja WAF pada dasarnya adalah mendeteksi setiap serangan yang datang ke dalam website, serangan dapat melalui banyak celah namun pada umumnya melalui inputan yang tersedia. Serangan-serangan pada aplikasi web cukup banyak seperti SQL Injection, XSS, CSRF, IDOR, DDOS, Broken authentication hingga eksploitasi miskonfigurasi pada web server. Untuk mendeteksinya, WAF akan mempelajari setiap inputan dari pengunjung jika sesuai dengan rule yang ada maka akan dilakukan pemblokiran request, dan IP penyerang tersebut akan dimasukkan ke dalam listing blacklisted IP. Peraturan Blacklist IP pada setiap WAF akan berbeda-beda sesuai policy dari perusahaan, namun pada dasarnya blacklist ip tidaklah selalu permanen karena bisa saja setelah melewati waktu *banned,* ip tersebut menjadi ip yang bersih kembali.

Sudah menjadi hal biasa dan menjadi kebiasaan masyarakat dunia maya pasti mengakses website dengan menggunakan macam-macam device untuk mendapatkan informasi baik smartphone, laptop, PC dll. Pada dasarnya terhubungnya device seperti smartphone, laptop, PC adalah karena memiliki perangkat keras yang disebut Network Interface Card (NIC). NIC tersebut dikonfigurasi dengan ip address supaya terkoneksi ke jaringan internet dan bisa mengakses informasi ke setiap server. Untuk itu device yang terkoneksi ke internet maka pasti memiliki ip address masing-masing.

Dalam keamanan sistem informasi ip address menjadi hal yang sangat penting, karena merupakan jalur utama terkoneksinya antara device satu dengan device lainnya di jaringan internet. Karena ip address menjadi alamat tiap-tiap device, maka seorang hacker jika melakukan serangan pasti akan mencari alamat ip terlebih dahulu supaya bisa mencari informasi detail target sehingga bisa melancarkan serangan nya. Demikian juga administrator antisipasi dan mengamankan sistem informasi dari serangan pun harus mengumpulkan atau mengelola setiap ip address dari device user yang pernah mengakses web server milik perusahaan nya, karena ip address tersebut menjadi identitas atau pengenal device yang digunakan oleh user. Dalam pengelolaan ip address user yang mengakses layanan server itu menjadi sebuah masalah para administrator karena banyaknya user yang mengakses server. Dalam WAF ip address akan dikumpulkan dalam

sebuah basis data supaya bisa dikelola dengan baik demi keamanan server, tapi semakin banyak ip address dalam basis data akan membutuhkan waktu dalam pencarian ip address suatu device yang dicurigai melakukan serangan pada server.

Sudah seharusnya bahwa sistem informasi membutuhkan basis data, kecepatan pemrosesan query pada basis data menjadi sebuah faktor yang tidak kalah pentingnya pada peningkatan kinerja sistem informasi. Dengan melakukan pengolahan query yang cepat, maka akan mendapatkan hasil pencarian yang maksimal, serta jeda waktu yang berlebih pada sisi pengguna akan bisa diminimalisir. Dari sekian banyak cara untuk mempercepat pemrosesan query ada salah satu cara yaitu dengan menerapkan metode Bloom Filter. Karena metoda Bloom Filter salah satu metoda yang memanfaatkan struktur data probabilistic space efficient.

Bloom Filter merupakan struktur data yang terdiri dari blok-blok yang berisikan angka biner atau pernyataan true/false. Cara kerja bloom filter adalah mengumpulkan hasil nilai suatu data ke dalam blok-blok yang sudah disediakan, jika sesuai maka bloom filter akan memberikan hasil true, jika terdapat satu atau lebih blok yang tidak sesuai dengan blok-blok positif maka bloom filter akan memberikan hasil false. Hal ini dapat mengurangi waktu pemrosesan pencarian. Bloom filter berguna untuk pekerjaan yang berhubungan dengan list ataupun dataset. Akurasi bloom filter bergantung pada lebar blok, jumlah hash yang digunakan pada saat melakukan filter, dan jumlah data yang dimasukkan kedalam set. Jumlah data yang banyak akan menurunkan akurasi yang dapat diukur dari kemunculan false positif.

Untuk mengoptimasi metode bloom filter, maka dalam penelitian ini digunakan metode counting bloom filter (CBF) dalam mencari ip address pada WAF. Standar bloom filter hanya mendukung penyisipan dan pencarian sedangkan dengan metode Counting bloom filter akan memecahkan masalah, karena CBF menggantikan unit bit tunggal dengan pencacah multi-bit, ketika memasukkan elemen ke k yang sesuai (k adalah jumlah fungsi hash) nilai penghitung ditambah 1. Demikian pula, saat menghapus elemen, nilai penghitung dikurangi 1 [3].

2. KAJIAN PUSTAKA

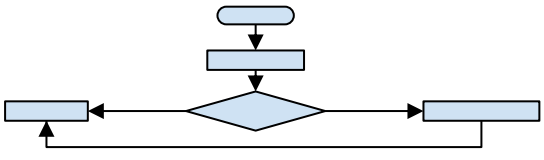
No	Judul	Author, Penerbit, Tahun	Kesimpulan	Metode
1	Improving counting Bloom filter performance with fingerprints [1]	Pontarelli, S., Reviriego, P., & Maestro, J. A. (2016). Information Processing Letters, 116(4), 304–309.	Menurunkan tingkat false positif pada CBF dengan pemodelan fp-cbf	Counting Bloom Filter
2	New Approach for Efficient IP Address Lookup Using a Bloom Filter in Trie-Based Algorithms [2]	Mun, Ju & Lim, Hyesook. (2015). IEEE Transactions on Computers. 65. 1-1.	Memfilter false positif dengan bloom pada pencarian alamat IP berbasis trie algoritma	Bloom Filter
3	Efficient IP address lookup with a counting Bloom filter in Trie-based algorithms [3]	Shu-he Wang (), Bi-hua Tang (), Dong-ming Yuan (), He-fei Hu (), and Jing Ran () Wireless Communication and Sensor Network. August 2016, 743-750	pencarian alamat IP dari pencocokan prefiks terpanjang yang dapat melakukan beberapa operasi dengan node Trie.	Counting Bloom FIlter
4	An Improved Construction for Counting Bloom Filters [4]	Bonomi F., Mitzenmacher M., Panigrahy R., Singh S., Varghese G. (2006). In: Azar Y., Erlebach T. (eds) Algorithms – ESA 2006. ESA 2006. Lecture Notes in Computer Science, vol 4168. Springer, Berlin, Heidelberg.	menyediakan alternatif berbasis hashing sederhana berdasarkan hashing kiri-d yang disebut CBF kiri-d (dlCBF). DlCBF menawarkan fungsionalitas yang sama dengan CBF, tetapi menggunakan lebih sedikit ruang, umumnya menghemat dua faktor atau lebih	Counting Bloom Filter
5	Towards practical use of Bloom Filter based IP lookup in operational network[5]	T. Yang, G. Xie, X. Sun, R. Duan and K. Salamatian, 2014 IEEE Network Operations and Management Symposium (NOMS), Krakow, 2014, pp. 1-4	Pada penelitian ini menggunakan pemodelan modifikasi counting bloom filter dengan menggunakan Withdrawal To Announcement untuk diimplementasikan pada router.	Counting Bloom Filter

6	A modified algorithm based on the Bloom Filter[6]	Jiang, M., Zhao, C., & Gao, X. (2013). 2013 6th International Congress on Image and Signal Processing (CISP).	Penelitian ini memodifikasi fungsi eigenvalues dengan karakteristik yang sesuai dengan algoritma bloom filter dengan memecah berdasarkan filter hash	Bloom Filter
7	PassDB: A password database withstrict privacy protocol using 3D Bloom filter[7]	Patgiri, R., Nayak, S., & Borgohain, S. K. (2020). Information Sciences, 539, 157–176.	Pemodelan 3D bloom filter yang digunakan untuk pencarian pada enkripsi password, algoritma ini tidak bergantung pada hash.	Bloom Filter

3. METODE

Kecepatan dalam menentukan IP pengunjung termasuk IP buruk atau tidak menjadi hal yang penting pada mekanisme WAF. Menurut sumber penyedia informasi IP buruk / *blacklisted* IP dalam sehari jumlah IP unik yang terdaftar dapat mencapai 500.000 IP. Hal ini tentu menjadi tantangan ketika jumlah database IP membesar sedangkan di satu sisi WAF harus dapat menentukan dengan cepat IP pengunjung suatu website.

Salah satu solusi ekstrem untuk mengurangi waktu pencarian pada dataset yang besar adalah menggunakan algoritma bloom filter, namun karakteristik dari bloom filter adalah sebuah struktur data yang dapat bertambah namun tidak dapat dikurangi. Sedangkan pada WAF, IP dapat dibedakan menjadi tiga yaitu White IP, Grey IP, dan Black IP, status IP dapat berubah dari White IP ke Black IP dan sebaliknya, maka diperlukan suatu metode yang lebih tepat yang dapat dilakukan update pada struktur datanya. Untuk itu kami mengajukan sebuah pemodelan pencarian IP menggunakan Counting Bloom Filter dengan alur sistem sebagai berikut:



Gambar 1. Usulan pemodelan pencarian IP berdasarkan counting bloom filter

Setiap pengecekan IP harus melewati Counting Bloom Filter terlebih dahulu, dengan tujuan untuk mengurangi pengecekan data langsung ke database, hal ini sangat efektif mengingat jumlah *blacklisted* IP pasti selalu lebih sedikit dibandingkan non-blacklisted IP. Jika pada counting bloom filter tidak ditemukan IP tersebut maka pengunjung akan dapat terus melakukan akses pada website, sebaliknya jika ditemukan maka ada dua kemungkinan yang pertama adalah IP benar termasuk dalam blacklisted IP di database atau yang kedua merupakan informasi false positive dari kesalahan penghitungan pada counting bloom filter.

4. HASIL DAN PEMBAHASAN

Skenario pengujian menggunakan database MySql dengan cara melakukan analisis performa pencarian IP berdasarkan dua metode, yang pertama adalah metode Sequence, dan yang kedua menggunakan pemodelan yang kami ajukan yaitu pencarian menggunakan CBF + Sequence.

No	N-Search	Total Execution Time Using Sequence (ms)	Average Execution Time Using Sequence (ms)	Total Execution Time Using CBF + Sequence (Our method) (ms)	Average Execution Time Using CBF + Sequence (Our method) (ms)
1	1.000	6149.40	6.1432	5.5355	0.00553
2	10.000	539.02	0.05389	76.8477	0.00768
3	100.000	3183.38	0.03183	491.143	0.00491
4	500.000	5757.69	0.01151	2462.4315	0.00492
5	1.000.000	6363.17	0.00636	4796.2357	0.00479

Dengan melihat hasil uji diatas, terlihat pemodelan yang diajukan dapat menurunkan total waktu pencarian IP hingga 99.9%.

Untuk pengujian akurasinya adalah sebagai berikut:

No	N-Search	True Negatif (TN)	True Positif (TP)	False Positive Rate (FPR)
1	1.000	1	0	0.001
2	10.000	3	17	0.0003
3	100.000	28	100	0.00028
4	500.000	124	517	0.000248
5	1.000.000	229	1010	0.000229

False positive rate dihitung dengan rumus:

fpr = TN / N-Search

Untuk nilai false positive-nya sangat kecil dan berangsur stabil di angka 0.0002 atau dapat dikatakan bahwa error hanya sebesar 0.02%.

5. KESIMPULAN

Berdasarkan penelitian pemodelan CBF yang diajukan untuk pencarian IP pada WAF memiliki peningkatan yang cukup signifikan yaitu menurunkan waktu pencarian hingga 99,9% dan hanya memiliki tingkat error sebesar 0.02%. Untuk IP yang salah pengklasifikasian dapat dilabeli sebagai whitelist IP. Untuk kedepannya akan dicoba untuk melakukan pemodelan dengan jenis algoritma bloom filter lainnya.

REFERENSI

[1] Pontarelli, S., Reviriego, P., & Maestro, J. A. (2016). Improving counting Bloom filter performance with fingerprints. Information Processing Letters, 116(4), 304–309. doi:10.1016/j.ipl.2015.11.002

[2] Mun, Ju & Lim, Hyesook. (2015). New Approach for Efficient IP Address Lookup Using a Bloom Filter in Trie-Based Algorithms. IEEE Transactions on Computers. 65. 1-1. 10.1109/TC.2015.2444850.

[3] Efficient IP address lookup with a counting Bloom filter in Trie-based algorithms
Shu-he Wang (), Bi-hua Tang (), Dong-ming Yuan (), He-fei Hu (), and Jing Ran () Wireless Communication and Sensor Network. August 2016, 743-750

[4] Bonomi F., Mitzenmacher M., Panigrahy R., Singh S., Varghese G. (2006) An Improved Construction for Counting Bloom Filters. In: Azar Y., Erlebach T. (eds) Algorithms – ESA 2006. ESA 2006. Lecture Notes in Computer Science, vol 4168. Springer, Berlin, Heidelberg. https://doi.org/10.1007/11841036_61

[5] T. Yang, G. Xie, X. Sun, R. Duan and K. Salamatian, "Towards practical use of Bloom Filter based IP lookup in operational network," 2014 IEEE Network Operations and Management Symposium (NOMS), Krakow, 2014, pp. 1-4, doi: 10.1109/NOMS.2014.6838341.

[6] Jiang, M., Zhao, C., & Gao, X. (2013). A modified algorithm based on the Bloom Filter. 2013 6th International Congress on Image and Signal Processing (CISP). doi:10.1109/cisp.2013.6745220

[7] Patgiri, R., Nayak, S., & Borgohain, S. K. (2020). PassDB: A password database with strict privacy protocol using 3D Bloom filter. Information Sciences, 539, 157–176. doi:10.1016/j.ins.2020.05.135