



Analysis of Malware Dns Attack on the Network Using Domain Name System Indicators

Analisis Serangan Dns Malware Di Jaringan Menggunakan *Domain Name System* Indikator (Studi Kasus Universitas Bina Darma)

Beni Brahara¹ Dedy Syamsuar² Yessi Novaria Kunang³

Program Studi Magister Teknik Informatika, Universitas Bina Darma
Jl. Jenderal Ahmad Yani No.3, 9/10 Ulu, Seberang Ulu I, Kota Palembang

¹Beni.brahara@gmail.com ²Dedy_Syamsuar@binadarma.ac.id,

³yesinovariakunang@binadarma.ac.id

Abstract

University of Bina Darma Palembang has its own DNS server and in this study using log data from the Bina Darma University DNS server as data in the study, DNS log server data is analyzed by network traffic, using Network Analyzer tools to see the activity of a normal traffic or anomaly traffic, or even contains DGA Malware (Generating Algorithm Domain). DGA malware produces a number of random domain names that are used to infiltrate DNS servers. To detect DGA using DNS traffic, NXDomain. The result is that each domain name in a group domain is generated by one domain that is often used at short times and simultaneously has a similar life time and query style. Next look for this pattern in NXDomain DNS traffic to filter domains generated algorithmically that the domain contains DGA. In analyzing DNS traffic whether it contains Malware and whether network traffic is normal or anomaly, in this study it detects Malware DNS. From the results of the stages of the suspected domain indicated by malware, a suspected domain list table is also created and also a suspected list of IP addresses. To support the suspected domain analysis results, info graphic is displayed using rapidminer tools to test decisions that have been made using the previous tools using the Decision Tree method.

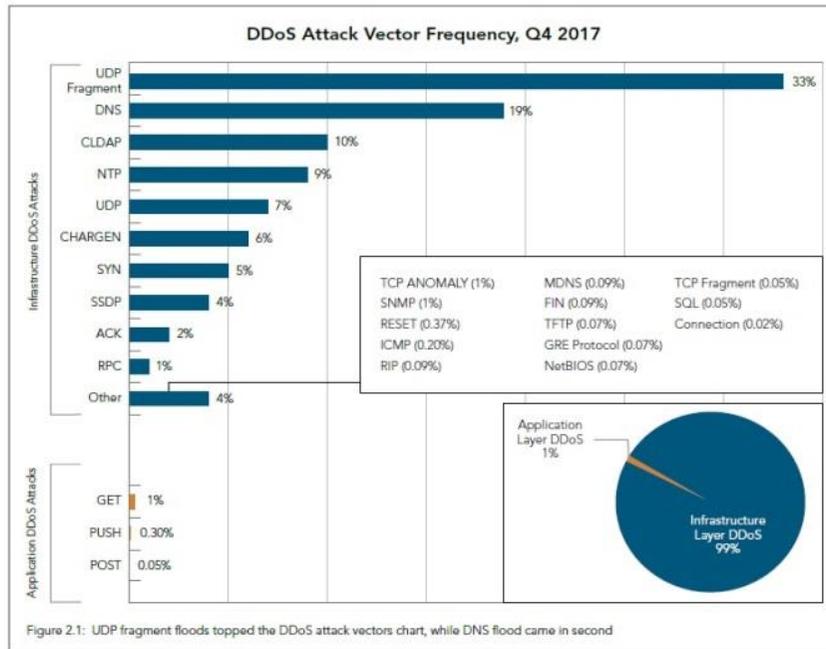
Keyword : *Log, DNS Malware, DGA, Malicious Traffic, Normal traffic, Anomaly*



1. PENDAHULUAN

Pada saat ini jaringan Internet merupakan jaringan yang paling banyak digunakan oleh semua orang di seluruh dunia, Hal ini terlihat dari hasil survey-survey yang dilakukan oleh Asosiasi Penyelenggara Jaringan Internet Indonesia (APJII), yang mengungkap bahwa lebih dari setengah penduduk Indonesia kini telah terhubung ke Internet. Survei yang dilakukan sepanjang 2016 itu menemukan bahwa 132,7 juta orang Indonesia telah terhubung ke Internet. Pada tahun 2017 Akamai melaporkan Indonesia menjadi nomor 1 sebagai sumber serangan di Internet (*malicious traffic*). *Traffic* serangan dari IP Indonesia berkisar 38% dari seluruh serangan di Internet dibandingkan *traffic* dari sekitar 175 negara yang diteliti. *Traffic* serangan ini meningkat hampir 2 kali lipat dibandingkan data sebelumnya yaitu sekitar 21%. Akamai dalam laporan tersebut menyatakan bahwa IP yang terdeteksi sebagai sumber serangan bisa jadi tidak mencerminkan lokasi penyerang. Karena bisa saja seorang penyerang dari Amerika Serikat melancarkan serangan dari IP Indonesia melalui jaringan *botnet* atau komputer yang terinfeksi *Malware*. Selain itu ESSET Indonesia pada bulan Mei melaporkan tingkat prevelansi *Malware* di ASEAN yang cukup tinggi, yaitu sebesar 16,88%. Dari laporan tersebut, *Malware* yang banyak beredar di Indonesia di antaranya adalah *Ramnit* dan *Sality*. Sedangkan berdasarkan hasil survei *malware* yang dilakukan ID-CERT, 52% *Malware* yang dilaporkan adalah *Adware* dan 35%-nya adalah *Trojan*, sisanya merupakan *Virus*, *Worm*, *Keylogger*, *Spyware* dan *Backdoor*, Salah satu cara mengantisipasi sebuah *attack* yaitu menganalisis *traffic* DNS (*Domain Name System*) karena dengan menganalisis lalu lintas jaringan yang kita, kita dapat mengetahui lalu lintas yang aneh yang terhubung didalam jaringan kita (*anomaly traffic*).

Traffic anomaly juga tercatat pada kuartal 2 tahun 2016, 45,2% komputer di Indonesia terserang *Malware* lebih tinggi dengan angka rata-rata global pada kuartal yang sama sebesar 20,8 Kategori perangkat lunak berbahaya yang paling sering ditemui di Indonesia pada kuartal 2 2016 adalah *Trojan* dengan 41,5 persen angka serangan pada seluruh komputer naik 37,8% dibandingkan angka pada kuartal sebelumnya. *Worms* menempati posisi kedua, dengan 24,5% serangan pada seluruh komputer, turun 26,3% dibandingkan kuartal sebelumnya, Berdasarkan data diatas tanpa disadari jaringan lokal perusahaan atau instansi di Indonesia bisa saja terserang *Malware* dikomputer *client* yang merequest paket *server* DNS.



Gambar 1. Akamai State of The Internet Security Q4 2017 report

Berdasarkan data serangan diatas DNS menjadi salah satu yang paling banyak diserang , *Domain Name System* atau biasa disebut sebagai DNS adalah suatu sistem yang memungkinkan nama suatu *host* pada jaringan komputer atau Internet ditranslasikan menjadi *IP address* atau sebaliknya, sehingga *client* dapat terhubung ke web *server* atau ke mail *server* menggunakan domain bukan *IP address* (Sugeng, 2010). Untuk pencegahan mengantisipasi serangan pada DNS dapat dilakukan dengan Monitoring *traffic* DNS dan menganalisis *traffic* jaringan yang memiliki pola yang sama dengan mendeteksi group *host* yang memiliki perilaku yang sama dan pola komunikasi dengan mengamati *traffic* jaringan, Ada 4 set kategori yang bisa digunakan sebagai ciri untuk mendeteksi adanya “*malicious domain*” yang diakses lewat DNS. Kelompok fitur lainnya yang juga bisa dipakai untuk mendeteksi aktifitas yang jahat (*malicious activity*) termasuk fitur waktu (rasio akses domain, umur dari domain, pola pengulangan akses, dan kemiripan per hari), hasil jawaban DNS (jumlah negara unik, jumlah alamat IP unik, dan jumlah IP dalam satu domain), nilai TTL (nilai rata-rata TTL, standard deviasi TTL, dan jumlah nilai TTL yang unik), dan fitur nama domain (jumlah angka yang ada pada nama domain). Salah satu penelitian mengenai DNS *Malware* pernah dilakukan oleh A.Karima (2011).

Network Traffic adalah sumber segala kejahatan siber, menurut *Identity Theft Resource Center (ITRC)* dalam studi mereka sampai bulan Juli 2018 diketahui bahwa telah terjadi 668 kasus kejahatan siber dengan total data hilang mencapai 22.408.258 dari seluruh kategori. Besarnya data yang hilang menunjukkan rentannya pertahanan banyak korporasi dunia terhadap serangan melalui jaringan. Realita ini juga merupakan parameter bagaimana deteksi dan perlindungan dari serangan siber tidak dapat hanya bergantung pada pertahanan perimeter semata. Sangat penting untuk menganalisis sumber data lain untuk menguatkan indikator yang ditemukan di titik akhir dan menciptakan peluang untuk mendeteksi serangan siluman. Teknologi yang mampu melakukan ini disebut analisis lalu lintas jaringan (*Network analysis*).

Cara kerja analisis *traffic* jaringan pada prinsipnya yaitu , perangkat akan memonitor seluruh transaksi data pada sebuah jaringan yang terhubung. Selanjutnya, seorang administrator jaringan dapat dengan mudah membuka port yang dibutuhkan, sehingga seluruh data dapat dianalisa secara *real time*. Dengan menyediakan data yang jauh lebih kaya, analisis lalu lintas jaringan paket yang mendalam dapat mengidentifikasi masalah secara dini, termasuk bilamana terjadi sebaran data yang mencurigakan, termasuk *Malware* dan ancaman serangan digital. Selain itu, dapat pula mengukur kemacetan dalam jaringan, melihat aplikasi apa yang memonopoli sumber daya dan *bandwidth*, dan memperingatkan administrator untuk tren perubahan nama *file* yang merupakan indikator khas serangan *Ransomware*. Lansiran dapat diatur untuk memberi tahu administrator tentang aktivitas yang tidak biasa atau *anomaly* pada jaringan, mengurangi risiko terburuk yang mungkin terjadi maka dari itu pentingnya sistem analisis pada jaringan komputer. Pentingnya sistem analisis jaringan dalam infrastruktur keamanan perusahaan juga ditekankan oleh Gartner dalam pernyataannya yang mengatakan "*Network Traffic Analysis (NTA)* solusi memonitor lalu lintas jaringan, arus, koneksi, dan objek untuk perilaku yang menunjukkan niat jahat. Perusahaan yang mencari pendekatan berbasis jaringan untuk mengidentifikasi serangan lanjutan yang telah melewati keamanan parameter harus mempertimbangkan NTA sebagai cara untuk membantu mengidentifikasi, mengelola, dan menentukan tingkat bahaya atau urgensi.

Berdasarkan uraian latar belakang diatas dan ada 4 kategori yang menjadi point pada penelitian ini maka peneliti mencoba menganalisis lalu lintas jaringan atau *traffic* jaringan DNS di Universitas Bina Darma Palembang , Universitas Bina Darma Palembang memiliki *server* DNS tersendiri dan dari data *log server* DNS yang akan dijadikan data dalam penelitian ini.

2. LANDASAN TEORI

2.1 Traffic

Secara umum *traffic* dapat diartikan sebagai perpindahan informasi dari satu tempat ke tempat lain melalui jaringan telekomunikasi. Besaran dari suatu *traffic* telekomunikasi diukur dengan satuan waktu, sedangkan nilai *traffic* dari suatu kanal adalah lamanya waktu pendudukan pada kanal tersebut (Kristalina, 2016). Besaran *Traffic* terbagi menjadi 2 yaitu ;

- Volume *Traffic*, didefinisikan sebagai jumlah total waktu pendudukan dari sebuah panggilan.

$$V = \int_{t=0}^T J(t)dt$$

T = jumlah periode pengamatan

$J(t)$ = jumlah kanal yang diduduki saat t

- Intensitas *Traffic* didefinisikan sebagai jumlah total waktu pendudukan dalam suatu selang pengamatan tertentu (per satuan waktu)

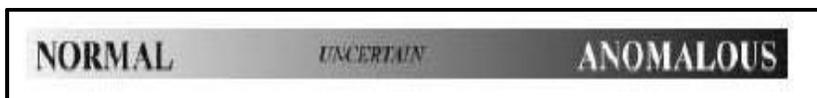
$$A = \frac{\text{volume traffic}}{T} = \frac{V}{T}$$

2.2 Anomaly

Adalah suatu keganjilan, keanehan atau penyimpangan dari keadaan biasa atau normal yang berbeda dari kondisi umum dalam suatu lingkungan (C.P Chaplin, 1989)

2.3 Anomaly detection

Pola dipelajari dari data normal. Data yang tidak terlihat dicek dan dicari penyimpangan dari pola yang telah dipelajari. Metoda ini tidak mampu mengidentifikasi tipe serangan (Wijaya,2009)



Gambar 2.1 Gambaran Mengenai Kegiatan Anomaly dan Normal Berdasarkan waktu kapan audit data dianalisis terdapat 2 kemungkinan:

1. On line IDS

*On line*IDS dapat menangkap usaha penyerangan sebelum status sistem disepakati, tetapi *On line* IDS harus dijalankan bersamaan dengan sistem aplikasi lain yang akan berpengaruh buruk terhadap *throughput*.

2. Off line IDS

Off line IDS hanya dapat mendeteksi serangan setelah terjadi penyerangan. Algoritma *Data Mining* diterapkan untuk menganalisis log data *off line mode*, sehingga *anomaly* dapat ditelusuri, dapat dianalisis oleh orang yang ahli, dan kemudian pola untuk menelusuri serangan yang baru dapat dihitung, dan dapat diinstallkan ke dalam *On line / real time* IDS.

2.4 Faktor Yang Mempengaruhi Traffic

2.4.1 Flashcrowd

Flashcrowd merupakan suatu keadaan dimana terjadi lonjakan akses yang tiba-tiba secara alamiah, tidak disengaja dan tidak terduga ke suatu server yang datang dari berbagai user yang sah dan berhak yang tersebar tidak terbatas secara acak di Internet pada suatu periode waktu tertentu secara gradual dalam hitungan menit atau jam dan jarang sekali terjadi dalam hitungan detik. Hal ini menyebabkan peningkatan dramatis beban *server* dan tekanan berat pada *link* jaringan yang mengarah ke server, sehingga menghasilkan peningkatan substansial dalam *packet loss* dan kepadatan *traffic* (Irsyad, 2015)

2.4.2 Mahalanobis Distance

Mahalanobis Distance (MD) dapat mengukur jarak diantara dua objek data multivariat atau antara suatu titik A dengan suatu distribusi B sehingga dapat mengetahui korelasi antara dua variabel dan menghapus ketergantungan pada skala pengukuran selama penghitungan. Dapat juga menghitung seberapa jauh standar deviasi A terhadap rata-rata B atau seberapa besar kesamaan antara sekumpulan kondisi terhadap sekumpulan kondisi yang ideal dengan memikirkan korelasi antar objek dalam bentuk variabel vektor dan matrik *covariance* dari kedua objek tersebut (Irsyad, 2015).

2.4.3 Cosine Distance

Cosine Distance (CD) merupakan salah satu teknik untuk menghitung jarak antar dua vektor yang pada dasarnya berasal dari sisa nilai dalam skala 1 sampai 0 dari *Cosine Similarity*. *Cosine Similarity* sendiri mengukur kesamaan atau kemiripan antara dua vektor dot product dengan menghitung sudut kosinus diantara kedua vektor tersebut. Biasanya *Cosine Similarity* digunakan dalam ruang positif dengan hasil berada diantara 0 dengan 1 berikut merupakan formula *Cosine Distance* (Irsyad, 2015).

2.5 Domain Generation Algorithm

) *Domain Generation Algorithm* (DGA) adalah sebuah program atau subrutin yang menyediakan perangkat lunak perusak dengan domain baru sesuai permintaan atau dengan cepat. Teknik DGA sedang digunakan karena malware yang bergantung pada domain tetap atau alamat IP dengan cepat diblokir, yang kemudian menghambat operasi. Jadi, daripada mengeluarkan versi baru malware atau menyiapkan semuanya kembali di server baru, *Malware* tersebut beralih ke domain baru secara berkala. Beberapa karakteristik DGA (enrico, 2011):

1. *NXDOMAIN* response
2. Menggunakan nama domain 2LD/ 3LD secara acak. (2LD=2nd level domain)
3. Peningkatan rikues yang banyak dari IP yang sama
4. Ada yang membuat domain dengan kata-kata yg tidak bisa dibaca, ada juga yang mengandung kata-kata yang bisa dibaca

Contoh *Malware* yang menggunakan teknik DGA

1. *Kraken*
2. *Gameover Zeus*
3. *Pykspa*
4. *Cryptolocker*
5. *Dyre, Virus dll.*

3. METODOLOGI PENELITIAN

Desain penelitian yang digunakan adalah dengan menggunakan pendekatan kualitatif. Penelitian kualitatif adalah penelitian yang menekankan pada *quality* atau hal yang terpenting dari sifat suatu barang atau jasa (satori dan komariah, 2011).

3.1 Data Penelitian

Data-datapenelitian yang dikumpulkan untuk proses analisis dalam penelitian ini menggunakan data primer dari objek penelitian yaitu data *traffic* jaringan DNS server di Universitas Bina Darma Palembang dan data sekunder mengambil dari beberapa sumber buku, jurnal penelitian terdahulu.

Tabel 3.1 Data *Log* DNS

Data	Tanggal	File	Size
Dns_log	2Agustus 2018	Ns1	58,4 Mb
		Nslocal	64 Mb

Tabel 3.1 diatas adalah data sampel *log server* DNS yang diperoleh peneliti dari objek pebelitian yaitu Universitas Bindarma Palembang atas seizin pihak Universitas guna datadalam penelitian ini.

3.2 Teknik Analisis Data

Secara umum penelitian ini menggunakan metode penelitian *experiment*, yaitu melakukan pengujian menggunakan *tools* yang direkomendasikan dan tingkat akurasi pengujian menggunakan algoritma *Decesion Tree* dalam analisa *traffic* pada jaringan. Pengujian algoritma dilakukan dengan menggunakan data dari objek penelitian yaitu data *log* DNS dari server Universitas Bina Barma Palembang yang merupakan data *traffic* jaringan hasil monitor *traffic* dari DNS yang diolah dan diklasifikasikan menjadi beberapa jenis intrusi yang didefinisikan dalam label *class* seperti normal dan *anomaly* dan beberapa serangan yang lain, berikut tabel beberapa *attribute* dalam *traffic* jaringan yang sangat mempengaruhi deteksi *anomaly traffic* pada jaringan, Wijaya (2009).

Tabel 3.2 Beberapa Atribute Yang Mempengaruhi Traffic jaringan

1	Duration	Lama waktu koneksi (nilainya dari 0 sampai tak terhingga)
2	protocol_type	Tipe protokol, yaitu: ICMP / TCP / UDP
3	Service	Layanan jaringan yang digunakan, misalnya: HTTP, Telnet, FTP, IMAP, DNS, dsb
4	src_bytes	Jumlah byte data yang dikirimkan dari sumber ke tujuan (nilainya 0 s/d tak terhingga)
5	dst_bytes	Jumlah byte data yang dikirimkan dari tujuan ke sumber (nilainya 0 s/d tak terhingga)
6	Hot	Besarnya indikator "hot" (nilai rangananya 0 s/d 30)

7	logged_in	Keberhasilan user melakukan login (1 jika berhasil login, 0 jika tidak)
8	num_root	Jumlah akses root (nilainya berupa angka dari 0 s/d 6)
9	Count	Jumlah koneksi ke host yang sama di dua detik terakhir (nilainya 0 s/d 511)
10	dst_host_count	menghitung untuk host tujuan (nilainya 0 s/d 255)
11	dst_host_diff_srv_rate	tingkat layanan yang berbeda untuk host tujuan (nilainya 0, 0.01, 0.02, ... , 1)

Tabel 3.4 Proses Tahapan analisis

DNS Data collect and pre-handle	
DNS Nxdomain traffic Capture	Mengcapture atau mengumpulkan data <i>traffic</i> DNS yang akan di analisis
DNS Traffic Filter	Mengfiltering data <i>traffic</i> yang request ke server DNS
Domain Name Acces similarityanalysis	
Domain into groub	Pengelompokan dan membagi domain berdasarkan waktu, <i>source</i> ip dan <i>destination</i> ip yang merequest ke server (<i>Devide Domain Groub</i>)
Calculate Domain Domain Acces similarity for each groub	Melihat kesamaan <i>acces</i> domain dari groub yang telah dikelompokan dan menghitung berdasarkan kesamaannya
Suspicious DGA-Domain list name	Hasil yang didapat adalah <i>list name</i> domain berdasarkan <i>traffic</i> yang dianalisis apakah mengandung DGA atau <i>malicious traffic, anomaly traffic</i>
Result	
Hasil analisis	Yaitu hasil dari tahapan sebelumnya menggunakan tools Sawmil dan Even Log analyzer menghasilkan sebuah data atau temuan berdasarkan data penelitian.

Decision Tree	Hasil temuan berupa data dan diproses melalui tools rapid minner sebagai penegasan hasil analisis yang menyatakan data tersebut <i>anomaly</i> traffic normal dengan mengamati pola <i>tree</i> yang dihasilkan.
Pembahasan	Membahas hasil temuan yang telah didapat dan diuji menggunakan tools tersebut dan memahami apa yang dihasilkan dari data yang diteliti sehingga mendapatkan sebuah kesimpulan spekulasi terhadap hasil analisis.
Kesimpulan	Berisikan kesimpulan dari seluruh rangkaian penelitian , kesimpulan dari 3 <i>tools</i> yang digunakan , kesimpulan dari tujuan penelitian, kesimpulan pembahasan hasil penelitian.

Dengan data diatas sebagai data penelitian atau data analisis dengan parameter uji yaitu *traffic* normal *anomaly* dengan jumlah *sample* data *count* 3685 dan dianalisis menggunakan *tools* pada tahapan analisis Data dianalisis Menggunakan *Tools* dibawah ini ;

1. Even log Analyzer 11

Even log Analyzer adalah *tools network* analisis Pengelolaan log, audit, dan kepatuhan TI manajemen , yang akan digunakan untuk memproses data *log* DNS dari objek penelitian .

2. Sawmill

Sawmill tools menganalisis, memantau dan memperingatkan berbagai sistem. Digunakan untuk menormalisasikan data hasil *extract tools* sebelumnya dan melakukan pengelompokan data berdasarkan variabel penelitian.

3. Rappidminer

Rappidminer adalah aplikasi untuk memproses data perhitungan data dan melakukan pengujian data hasil analisis , yang digunakan dalam penelitian ini adalah *decidion tree* untuk pegujian keakurasian analisis *traffic* jaringan berdasarkan hasil data menggunakan dua *tools* sebelumnya.

3.3 Data Penelitian

Data-datapenelitian yang dikumpulkan untuk proses analisis dalam penelitian ini menggunakan data primer dari objek penelitian yaitu data *traffic* jaringan DNS server di Universitas Bina Darma Palembang dan data sekunder mengambil dari beberapa sumber buku, jurnal penelitian terdahulu.

4. HASIL ANALISIS DAN PEMBAHASAN

Hasil analisis yang didapat dari *tools* yang digunakan dalam dalam penelitian ini ada tiga *tools* yang diguakan yaitu ;

4.1 Sawmil

Hasil analisis menggunakan *tools Sawmil* adalah peneliti mendapatkan pengelompokan jumlah *traffic* atau lonjakan *request* tertinggi dari sekian banyak data sehingga mempermudah dalam pencarian dipetakan dalam tabel waktu akses sehingga dapat mempermudah dalam pencarian karena sudah mendapatkan informasi waktu *traffic* paling tinggi berdasarkan priode waktu hari,jam,menit dan detik dari data penelitian ditampilkan dalam bentuk grafik.

	↑ Hour of day	Messages
1	3:00 AM - 4:00 AM	1
2	6:00 AM - 7:00 AM	49
3	7:00 AM - 8:00 AM	939
4	8:00 AM - 9:00 AM	4,666
5	9:00 AM - 10:00 AM	8,367
6	10:00 AM - 11:00 AM	5,122
7	11:00 AM - noon	5,982
8	noon - 1:00 PM	4,708
9	1:00 PM - 2:00 PM	4,460
10	2:00 PM - 3:00 PM	1,128
11	3:00 PM - 4:00 PM	813
12	4:00 PM - 5:00 PM	250
	Total	36,485

Gambar 4.1 Tabel Waktu *traffic*

Aktifitas *request* paling tinggi pada jam 8:00 sampai 9:00 Am dengan jumlah *bits* 8367 (*count*) ,dengan ip *address* yang paling banyak *request* yaitu [84.194.217.172](#) dengan 325 *count* dengan *percentase* 69,43% dan paling rendah pada jam 3:00 sampai 4:00 Am dengan *count* 1 , dari data waktu *traffic* diatas selanjutnya akan dilihat ip *address* dari masing-masing *request* dan akan dilakukan pengelompokan berdasarkan beberapa kriteria tertentu, untuk melihat hasil *output report* menggunakan *tools sawmill* secara keseluruhan dapat dilihat pada Lampiran penelitian pengelompokan pada data *summary*.

4.2 Even log analyzer11

Pada gambar berikut adalah tampilan tools Even log analyzer11 inport data penelitian dan ditelusuri data log DNS server

4.2.1 NxDomain Traffic Capture

Nxdoamin sampel hasil *filteringtraffic* berdasarkan data *traffic* dengan *protocol* UDP dan jumlah *count* , hasil *filtering* terdapat beberapa NxDomain yang mencurigakan (*anomaly traffic*) berdasarkan waktu dan jumlah *request (count)* dan ip dari domain tersebut dilihat dari sampel domain dan dimasukan pada *tools Even Log Analyzer11* agar mempermudah pencarian domain tersebut dan domain-domain terdihasilkan hasil akan dikemlopokan menjadi beberapa sampel dan tabel *Nxdomain* untuk lebih jelas dapat dilihat pada Gambar 4.2 dibawah ini ;



Gambar 4.2 Pencarian Nxdomain

Dengan memasukan nama Domain yang terdapat pada waktu *traffic* tertinggi dan memasitkan bahwa *traffic* tersebut memiliki jumlah *request* atau *count* yang terbanyak berikut penelusuran domain-domain yang memiliki jumlah *traffic* yang tetinggi sehingga nantinya akan dipilih domain-domain yang tertinggi untuk ditelusuri ip *address* nya penelusuran domain dapat dilihat pada Gambar 4.3 dibawah ini ;



Gambar 4.3 Nxdomain www.gsjhehtqvin.com

Tabel 4.2 Hasil Capture NxDomain

Nama Nxdomain	Tanggal	Protocol	Size
www.gsjehtqvin.com	Aug - 2 - 2018	UDP	512 octect
www.ptblqwiz.com	Aug 2018	-2- UDP	512 octect
www.crvvrxfgsvohiy.com	Aug 2018	-2- UDP	512 octect
lcxeyzb.biz	Aug 2018	-2- UDP	512 octect
spo-msedge.net	Aug 2018	-2- UDP	512 octect
garenanoww.com	Aug 2018	-2- UDP	512 octect
mwlzwwr.biz	Aug 2018	-2- UDP	512 octect
ywvyfim.biz	Aug 2018	-2- UDP	512 octect
jjwelph.org	Aug 2018	-2- UDP	512 octect

4.2.2 Domain into Groub

Setelah melalui beberapa tahapan, tahap selanjutnya pengelompokan domain hasil *filtering* sebelumnya menjadi sebuah *groub* berdasarkan Nama Domain ,Protocol yang digunakan Tanggal, Waktu ,Panjang paket , Total perwaktu , Jumlah total keseluruhan *traffic*, peneliti mengambil 3 Domain yang tertinggi dan dicurigai *anomaly*, untuk lebih dapat dilihat pada Tabel dibawah ini ;

Tabel 4.3 Tabel Domain Groub

Nxdomain	Tanggal	Protocol	Waktu	Count	Total (count)	Size /s
www.gsjehtqvin.com	Aug 2 2018	UDP	09:00 13:01	2180 640	2820	512 octect
www.ptblqwiz.com	Aug 2 2018	UDP	11:01 13:01	341 520	861	512 octect
www.crvvrxfgsvohiy.com	Aug 2 2018	UDP	09:00 11:01	1702 430	2042	512 octect

4.2.3 Domain Name Access Similarity Analysis

Setelah pengelompokan domain menjadi sebuah group pada tahanan ini menganalisis kesamaan berdasarkan domain, ip address, jumlah paket, waktu paket yang telah didapat Tabel 5.5, pada tahapan ini peneliti membagi menjadi beberapa tahapan sebagai berikut;

1. Persamaan berdasarkan waktu akses

Analisis kesamaan berdasarkan tahapan sebelumnya terdapat 3 domain yaitu www.gsjhehtqvin.com, www.ptblqwjz.com, www.crvvrxfgsvohiy.com dilakukan menggunakan *toolsEvenlog analyzer11* dengan hasil akan dijelaskan pada Tabel 5.4 berikut;

Tabel 4.4 Persamaan Waktu Akses

<u>Nama Domain</u>	<u>Tanggal</u>	<u>Waktu</u>	<u>Count</u>	<u>Menit</u>
www.gsjhehtqvin.com	Aug - 2 -2018	13:01	640	13:01
www.ptblqwjz.com	Aug - 2 -2018	13:01	520	
www.gsjhehtqvin.com	Aug - 2 -2018	09:00	2180	09:14
www.crvvrxfgsvohiy.com	Aug - 2 -2018	09:00	1702	09:26
				09:35
				09:36
www.ptblqwjz.com	Aug - 2 -2018	11:01	341	11:01
www.crvvrxfgsvohiy.com	Aug - 2 -2018	11:01	430	

Tabel 4.5 Persamaan Waktu Akses Permenit

<u>Nama Domain</u>	<u>Jam</u>	<u>Count</u>	<u>Menit</u>	<u>Count</u>	<u>Jam</u>	<u>Nama Domain</u>
www.gsjhehtqvin.com	09:00	336	09:14	560	09:00	www.crvvrxfgsvohiy.com
		413	09:26	450		
		153	09:35	494		
		327	09:36	336		

Dari 2 Tabel diatas dapat dilihat iterasi dari kesamaan waktu akses dari domain-domain yang dicurigai, setiap domain memiliki kesamaan contoh www.gsjhehtqvin.com dan www.ptblqwjz.com memiliki kesamaan pada jam 13:01 dengan count yang berbeda dan menit yang sama yaitu 13:01, artinya semua paket diakses hampir dengan waktu yang bersamaan dengan jumlah yang berbeda-beda dalam beberapa menit perulangan, paket yang membedakan adalah jumlah count dan jedah waktu beberapa saat, dengan kata lain domain tersebut dalam waktu akses memiliki iterasi yang sama.



Gambar 4.4 Persamaan domain ip

Pada Gambar 4.4 dan 5.14 domain www.gsjhehtqvin.com yang sebelumnya pada gambar 4.2 memiliki ip 208.80.127.2 pada gambar 4.4 juga memiliki ip 208.94.148.2 menggambarkan bahwa ada persamaan domain dan ip dengan akses secara bersamaan aktifitas tersebut dapat dikategorikan sebagai *traffic anomaly* dengan karakteristik *Malware* yang memungkinkan domain memiliki banyak ip dan sebaliknya bahkan sebuah ip yang tidak *valid* atau tidak memiliki domain, maka tahapan selanjutnya yaitu menentukan domain dan ip yang terindikasi *Malware* atau *malicious* pada tabel *suspected domain*, untuk melihat kesamaan waktu akses, domain dan ip address secara detail pada graafik dapat dilihat pada *output report* menggunakan *tools even log analyzer* pada bagian lampiran, disana dapat dilihat kesamaan waktu traffic secara detail dalam bentuk grafik dari priode hari, jam, menit hingga akses perdetik.

4.2.4 Suspicious Domain List Name

Dari semua tahapan hasil analisis *domain alert malicious* ip dan *traffic anomaly*, dari data yang diteliti terdapat beberapa domain yang dapat penulis simpulkan telah terinfeksi *Malware* dan jumlah *traffic* yang *anomaly* pada tanggal 2 Agustus 2018 berdasarkan data yang diteliti, berikut 10 *List Domain Name* dengan jumlah *count* terbanyak yang dicurigai dan dinyatakan *anomaly* berdasarkan karakteristik *traffic* nya.

Tabel 4.6 Domain Anomaly

Domain Alert	Count	Anomaly
www.gsjhehtqvin.com	2820	Anomaly
www.crvvrxfgsvohiy.com	2042	Anomaly
www.eaaqsama.com	2056	Anomaly
www.ptblqwjz.com	861	Anomaly
lcxeyzb.biz	651	Anomaly
spo-msedge.net	441	Anomaly
garenanoww.com	576	Anomaly
mwlzwwr.biz	332	Anomaly
ywwyfim.biz	213	Anomaly
jjwelph.org	196	Anomaly

Tabel 4.6 adalah 10 domain terbanyak dari 36485 jumlah *traffic* DNS pada data log tantanggal 2-agustus-2018 sebagai data sampel dalam penelitian, 10 domain diatas adalah domain-domain yang dicurigai terindikasi *Malware* berdasarkan hasil analisis menggunakan *tools sawmil dan even log analyzer11* dan melalui beberapa tahapan analisis dapat disimpulkan bahwa domain-domain diatas *malicious* atau mengandung *Malware*, bahwa domain-domain tersebut memiliki jumlah *count* tertinggi dan memiliki banyak *ip address* dan sebaliknya satu *ip* memiliki beberapa domain seperti pada Gambar 4.6 sebagai contoh *ip* 208.80.127.2, dimiliki oleh tiga domain yaitu sebagai contoh domain www.gsjhehtqvin.com, www.ptblqwjz.com, www.crvvrxfgsvohiy.com dimiliki oleh *ip* 208.80.127.2 dalam waktu yang bersamaan, untuk lebih lebih jelas dibawah ini Tabel 5.7 menjelaskan kesamaan *ip* dan domain yang memiliki banyak *ip* dalam satu waktu yang bersamaan.

Tabel 4.7 List Domain IP Suspected

Domain	IP	Conut	Information	Status
www.gsjhehtqvin.com	208.80.125.2	282	Refused	Suspected
	208.80.124.2	282	Refused	Suspected
	208.94.148.2	282	Refused	Suspected
	208.80.126.2	282	Refused	Suspected
	208.80.127.2	282	Refused	Suspected

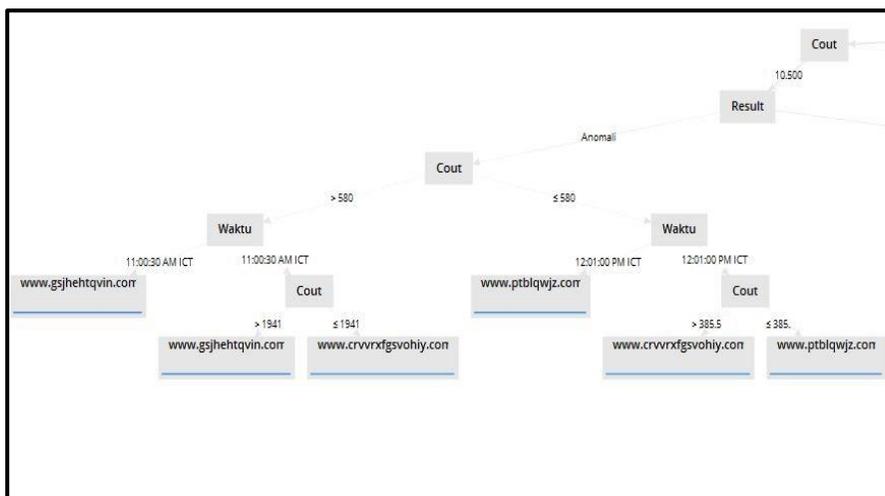
<i>www.crvvrxcfjgsvohiy.com</i>	208.80.127.2	204	Refused	Suspected
	208.80.126.2	204	Refused	Suspected
	208.94.148.2	204	Refused	Suspected
	208.80.125.2	204	Refused	Suspected
	208.80.124.2	204	Refused	Suspected
<i>www.ptblqwjz.com</i>	208.80.125.2	86	Refused	Suspected
	208.80.124.2	86	Refused	Suspected
	208.94.148.2	86	Refused	Suspected
	208.80.126.2	86	Refused	Suspected
	208.80.127.2	86	Refused	Suspected
<i>www.eaaqsama.com</i>	208.80.124.2	206	Refused	Suspected
	208.80.125.2	206	Refused	Suspected
	208.94.148.2	206	Refused	Suspected
	208.80.126.2	206	Refused	Suspected
	208.80.127.2	206	Refused	Suspected

Berdasarkan Tabel 4.7 diatas adalah domain-domain yang paling tinggi jumlah *count* nya dan hasil analisis menyatakan domain diatas mengandung *Malware* atau *anomaly* dan bahkan *malicious traffic* , satu domain memiliki lima ip *address* yang sama dan bahkan kelima domain tersebut memiliki ip yang sama dan dapat di simpulkan bahwa domain dan ip tersebut berasal dari sumber yang sama, berdasarkan informasi dari data *log* yang dianalisis domain dan ip tersebut berstatus *unexpected* yang artinya tak terduga atau *request* yang secara tiba-tiba dan *server* menyatakan *error* dan menolak *request* tersebut (*refused*) dan dari hasil tersebut peneliti juga mengelompokan ip *address* dengan *traffic* yang *anomaly* yang *suspected* dapat dilihat pada tabel 5.8 dibawah ini ;

Tabel 4.8 List Suspected IP Addres.

IP	Count	Information	Status
84.194.217.172	1536	Refused	Suspected
216.239.32.10	678	Refused	Suspected
216.239.34.10	677	Refused	Suspected
216.239.36.10	677	Refused	Suspected
216.239.38.10	678	Refused	Suspected

yang digunakan dalam perhitungan ini sesuai dengan metadata *log* DNS dan digunakan sebagai indikator maka *attribute* yang digunakan adalah *id*, *Nama domain*, *ip address*, *count*, *waktu*, *Tanggal* dan hasil perhitungan dengan *Result label* Normal dan *Anomaly* dijelaskan pada Gambar 5.20 *Output* perhitungan *decision tree* dibawah ini ;



Gambar 4.6 Hasil Perhitungan *Decision Tree*

Berdasarkan hasil perhitungan *Decision tree* menggunakan *tools rapidminer* dinyatakan count 10500 *anomaly* dibedakan dari waktu yaitu pukul 11:00 dan 12:00 dan dibagi menjadi 2 jalur *count* > dari 580 dan < 580 dinyatakan *Anomaly*, dengan turunan attribute “Waktu” terlihat ada 3 domain sama seperti sebelumnya bahwa 3 domain tersebut berdasarkan perhitungan *tools* ini juga dinyatakan *anomaly* yaitu domain *www.gsjhehtqvin.com*, *www.ptblqwjz.com*, *www.crvrxfsgsvohiy.com* sama dengan hasil analisis menggunakan *tools sawmill* dan *even log analyzer* bahwa 3 domain atau *traffic* domain tersebut *Anomaly* dan bisa jadi mengandung *Malware*, metode *Decision tree* ini hanya untuk penegasan dalam penentuan berdasarkan hasil dari dua *tools* sebelumnya yang menyatakan bahwa domain-domain tersebut terindikasi oleh *Malware* DGA, bahwa kategori yang bisa digunakan sebagai ciri untuk mendeteksi adanya “*malicious domain*” yang diakses lewat DNS. Kelompok fitur lainnya yang juga bisa dipakai untuk mendeteksi aktifitas yang jahat (*malicious activity*) termasuk fitur waktu (rasio akses domain, umur dari domain, pola pengulangan akses, dan kemiripan per hari), hasil jawaban DNS (jumlah negara unik, jumlah alamat IP unik, dan jumlah IP dalam satu domain), nilai TTL (nilai rata-rata TTL, standard deviasi TTL, dan jumlah nilai TTL yang unik), dan fitur nama *domain* (jumlah angka yang ada pada nama domain).

5. PEMBAHASAN

Dari hasil temuan tahapan analisis menggunakan 3 *tools* diatas didapat beberapa domain yang *suspected* sebagai *Malware* , berdasarkan perilaku dan karakteristik domain tersebut yang berpindah pindah dan menggunakan domain ip yang sama dalam waktu yang bersamaan dipastikan domain tersebut terindikasi algoritma DNS *genereting* atau *Malware* DGA (*Domain generating algorithm*) , DGA adalah algoritma digunakan untuk menggenerate nama domain yang bisa digunakan sebagai tempat komunikasi dengan server C&C. Teknik ini digunakan penyerang untuk menghindari *server C&C (Command & Control)* nya di *takedown*.

1. karakteristik DGA (*Domain Generation Algoritym*)
2. Menggunakan nama domain 2LD/ 3LD secara acak.
3. Peningkatan *request* yang banyak dari IP yang sama
4. Memiliki *Domain* dengan banyak ip dan sebaliknya
5. Ada yang membuat domain dengan kata-kata yg tidak bisa dibaca, ada juga yang mengandung kata-kata yang bisa dibaca bahwa kategori yang bisa digunakan sebagai ciri untuk mendeteksi adanya

Domain *www.gsjhehtqvin.com* mencoba masuk ke sistem server DNS Bina Darma dengan menggenere domain lainnya hingga kalau ditelusuri domain dan ip yang digunakan berasal dari sumber yang sama , namun sistem keamanan Bina Barma dapat mengidentifikasi serangan atau *request* yang gagal tersebut sehingga sistem keamanan UBD menolak *request* (*refused*) dapat dilihat pada Gambar dibawah ini



Gambar 4.7 *Unexpected Rcode Refused*

Gambar 5.1 diatas adalah DNS *query time status error unexpected Rcode Refused* artinya *request* yang secara tiba-tiba dan tak terduga ditolak oleh sistem keamanan *server* DNS Bina Barma , berarti sistem keamanan sudah mengantisipasi akan serangan tersebut dengan sistem keamanan yang tinggi, server DNS UBD memiliki sistem keamanan yang tinggi sehingga algoritma yang digunakan telah diketahui perilakunya dan tak dapat masuk kesistem dan tertolak, *www.gsjhehtqvin.com* Dengan ip 208.80.124.2 Masi tetap mencoba masuk tanpa menggunakan ip *address* seperti gambar dibawah ini



Gambar 4.8 Error Network Unreachable

Server juga menyatakan sebuah *error Network Unreachble* yang artinya domain tak tejangkau sulit dibaca oleh sever , kemudian domain tersebut berpindah mendapatkan atau menggunakan ip yang baru berbeda dari ip sebelumnya dalam waktu yang singkat seperti Gambar 5.2 dibawah ini;



Gambar 4.9 IP 208.80.125.2

Setelah menggunakan ip 208.80.24.2 tertolak dan masuk dengan tanpa ip address juga ditolak oleh sistem keamanan server selanjutnya domain tersebut menggunakan ip baru yaitu 208.80.25.2, sesuai dengan prilaku algoritma DGA akan tetapi sistem keamanan server Bina Barma lebih telah mengetahui perilaku yang *anomaly* tersebut dan algoritma GDA tersebut tak dapat masuk terdeteksi terlebih dahulu oleh server , berikut algoritma yang biasanya digunakan oleh penyerang untuk menyusup melalui domain,

```

1  from datetime import date
2  from hashlib import sha256
3
4  def dyre_dga(num, date_str=None):
5      if None == date_str:
6          date_str = '{0.year}-{0.month}-{0.day}'.format(date.today())
7
8          tlds = ['.cc', '.ws', '.to', '.in', '.hk', '.cn', '.tk', '.so']
9          hash = sha256('{0}{1}'.format(date_str, num)).hexdigest()[3:36]
10         replace_char = chr(0xFF & ((num % 26) + 97))
11
12         return '{0}{1}{2}:443'.format(replace_char, hash, tlds[num % len(tlds)])
13
14     todays_domains = [dyre_dga(i) for i in xrange(333)]

```

Gambar 4.10 Sumber Cisco

Gambar 5.4 diatas adalah contoh algoritma yang digunakan oleh domain yang yang teridikasi menggunakan algoritma DGA, *role* DGA menggunakan beberapa blok penyusun yaitu;

1. Benih, elemen dasar

2. Unsur yang berubah seiring waktu
3. Domain Tingkat Atas (TLD)

Benih dapat berupa frasa atau angka apapun dapat diubah oleh aktor penyerang ancaman secara praktis atau dengan cepat beralih ke versi tubuh baru atau domain ip yang berbeda, Benih dan elemen berbasis waktu digabungkan dalam algoritma tersebut untuk membuat nama domain dalam "tubuh" baru (domain) ini akan digabungkan dengan salah satu TLD yang tersedia. Yang perlu diperhatikan bahwa elemen berbasis waktu tidak harus selamanya tanggal dan waktu. bisa juga berupa hal lain yang berbeda dengan waktu,

5. KESIMPULAN DAN SARAN

5.1 KESIMPULAN

Hasil dari penelitian ini penulis dapat membuat kesimpulan bahwa DNS dapat menjadi salah satu indikator untuk menganalisis sebuah *traffic* pada jaringan dengan menganalisis aktifitas di jaringan tersebut, ada 4 set kategori yang bisa digunakan sebagai ciri untuk mendeteksi adanya "*malicious domain*" yang diakses lewat DNS, ditemukan *traffic anomaly* dan bahkan dapat dikatakan terindikasi mengandung *Malware DGA (Domain Generation Algorithm)* dimana algoritma tersebut dapat berubah atau berpindah domain dan ip secara cepat dalam waktu yang singkat seperti karakteristiknya dari *Malware* tersebut, dan itu membuktikan bahwa tanpa disadari DNS masi sangat rentan menjadi target serangan, DNS *log* sangat berguna membantu melihat *traffic anomaly* dengan beberapa tahapan analisis yang dilakukan, hasil penelitian ini juga mencega atau mempelajari *traffic* yang ada pada server DNS khusus nya diobjek penelitian.

Kelompok fitur lainnya yang juga bisa dipakai untuk mendeteksi aktifitas yang jahat (*malicious activity*) termasuk fitur waktu (rasio akses domain, umur dari domain, pola pengulangan akses, dan kemiripan per hari), hasil jawaban DNS (jumlah negara unik, jumlah alamat IP unik, dan jumlah IP dalam satu domain), nilai TTL (nilai rata-rata TTL, standard deviasi TTL, dan jumlah nilai TTL yang unik), dan fitur nama domain (jumlah angka yang ada pada nama domain). Tahanan-tahapan yang dilakukan dalam mendeteksi serangan *Malware*, DDoS DGA menghasilkan prosentase rata-rata pengenalan terhadap tiga kondisi jaringan (*Normal, slow Alert*) dengan Prosentase data *log Ns1* yang dihasilkan berada diatas rata-rata mengandung *anomaly traffic* dan *malicious Malware*

Hal ini menunjukkan bahwa pendekatan baru dalam mendeteksi serangan *Malware* dengan memanfaatkan analisis statistik terhadap *log* aktivitas jaringan dengan metode pengelompokan sampai domain *Alert* berfungsi sebagai fungsi deteksi yang berupaya mampu mengenali serangan *Malware*, DGA dan bahkan DDoS yang sangat merugikan pada sistem jaringan.

5.2 SARAN

Untuk menghasilkan tingkat pengenalan yang lebih baik lagi, maka untuk penelitian selanjutnya sebaiknya menambahkan beberapa parameter yang dapat dioptimasi yaitu tidak hanya DNS sebagai indikator penentuan dapat juga menambahkan indikator-indikator lainnya seperti seluruh *traffic* sebuah jaringan secara menyeluruh atau semua *service* pada jaringan tidak hanya DNS untuk mendapatkan informasi yang lengkap pada *traffic* jaringan dan juga dapat menggunakan metode-metode seperti *neurl network* dan memperbanyak jumlah data pelatihan, optimasi jumlah *neuron* dan *hidden layer* pada *neural network*, konfigurasi pelatihan *neural network* (*momentum*, *learning rate*, *epoch*, dan *goal mean square error*), penyesuaian fungsi pelatihan, dan fungsi aktivasi *layer neuralnetwork*. Diharapkan dengan adanya pendekatan baru dalam mengenali serangan di jaringan komputer bisa menjadi sebuah komplemen terhadap sebuah sistem keamanan jaringan.

UCAPAN TERIMAKASIH

Ucapan terimakasih Saya sampaikan untuk Dedy Syamsuar,Ph.D dan Yesi Novaria Kunang ,S.T.,M.Kom yang telah membimbing dan mendukung penuh penelitian ini.

DAFTAR PUSTAKA

- H.Choi, H. Lee, and H. K. (2009). *Botgad: detecting botnets by capturing group activities in network traffic,* in *Proceedings of the Fourth International ICST Conference on Communication System software and middlewaRE*.
- Jhohanes. (2018). The DGA of Pykspa. Retrieved from www.Jhohanes.com
- Kalista, P. (2016). *Konsep dan Teori Trafik*.
- Karima, A. (2012). *Deteksi anomali untuk identifikasi botnet kraken dan conficker menggunakan pendekatan rule based*. 2012(Semantik), 274–281.
- Sons, J. W. &. (2012). *CompTTA Network Study Guide 2nd Edition*. Indianapolis. 2.
- Wijaya, E. S., Syukur, A., Wahono, R. S., Thesis, J., Magister, P., & Informatika, T. (n.d.). *DETEKSI ANOMALI TRAFIK JARINGAN DENGAN MENGGUNAKAN METODE DECISION TREE*. 1–14.