ISSN: 2477-5126
e-ISSN: 2548-9354

Klasifikasi Paket Jaringan Berbasis Analisis Statistik dan *Neural Network*

Harsono[1], Muhammad Chambali[2], Arif Wirawan Muhammad[3]
[1]Jurusan Teknik Komputer, Politeknik Harapan Bersama, Tegal
[2] Jurusan Teknik Elektro, Politeknik Negeri Semarang, Semarang
[3] Jurusan Teknik Informatika, Politeknik Harapan Bersama, Tegal
[1,3]Jl. Mataram 9, Pesurungan Lor, Margadana, Tegal, 50272, Indonesia
[2]Jl. Prof. H. Soedarto, S.H., Tembalang, Semarang, 50275, Indonesia
email: [1]harsonopoltekhb@gmail.com, [2]mc.chambali.poltek@gmail, [3]arif.wirawan@poltektegal.ac.id

***Abstract* – Distributed Denial-of-Service (DDoS) is one of network attack technique which increased every year, especially in both of intensity and volume. DDoS attacks are still one of the world's major Internet threats and become a major problem of cyber-world security. Research in this paper aims to establish a new approach on network packets classification, which can be a basis for framework development on Distributed Denial-of-Service (DDoS) attack detection systems. The proposed approach to solving the problem on network packet classification is by combining statistical data quantification methods with neural network methods. Based on the test, it is found that the average percentage of neural network classification accuracy against network data packet is 92.99%.**

***Abstrak* – *Distributed Denial-of-Service* (DDoS) adalah sebuah serangan jaringan yang sangat sering terjadi dalam lingkup jaringan Internet, dimana intensitas dan volume DDoS terus meningkat setiap tahunnya. Serangan DDoS masih termasuk salah satu ancaman utama dunia Internet dan menjadi sumber utama pada masalah keamanan *cyber-world*. Penelitian dalam paper ini bertujuan untuk membentuk suatu pendekatan baru dalam kaitannya dengan klasifikasi paket jaringan, sehingga bisa menjadi sebuah *framework* pada pengembangan sistem deteksi serangan *Distributed Denial-of-Service* (DDoS). Pendekatan yang diusulkan untuk menyelesaikan permasalahan klasifikasi paket jaringan yaitu dengan mengkombinasikan metode kuantifikasi data secara statistik terhadap aliran paket data jaringan yang digabungkan dengan *neural network* sebagai basis *classifier*. Berdasarkan pengujian didapatkan bahwa rerata persentase akurasi klasifikasi *neural network* terhadap paket data jaringan Internet sebesar 92,99%.**

***Kata Kunci* – Klasifikasi, Statistik, *Neural Network*, *Framework*, Paket Data.**

I. PENDAHULUAN

Distributed Denial-of-Service (DDoS) adalah sebuah serangan jaringan yang sangat sering terjadi dalam lingkup jaringan Internet, dimana intensitas dan volume DDoS terus meningkat setiap tahunnya [1]. Pada saat ini, serangan DDoS masih termasuk salah satu ancaman utama dunia Internet dan menjadi sumber utama pada masalah keamanan *cyber-world* [2]. DDoS merupakan suatu ancaman permanen bagi *user*, organisasi, maupun infrastruktur yang ada dalam jaringan Internet. Disisi lain, serangan DDoS sangat berkontribusi atas timbulnya resiko baik dari segi kerahasiaan, integritas, serta ketersediaan *resource* yang disediakan dan dimiliki oleh sebuah organisasi pada jaringan Internet [3].

Klasifikasi paket jaringan yang melewati *router* pada suatu organisasi yang tekoneksi dengan jaringan Internet merupakan sebuah proses yang fundamental dan mutlak untuk dilaksanakan dalam rangka meminimalisir adanya resiko serangan DDoS [4]. Pada umumnya, proses deteksi dini terhadap serangan DDoS dilaksanakan oleh *Intrusion Detection System* (IDS) yang terpasang pada sistem router organisasi, namun teknik deteksi yang dilaksanakan oleh IDS dapat dikatakan jauh dari sempurna apabila dibandingkan dengan semakin bervariasinya teknik dan metode serangan *cyber* yang semakin modern [5]. Secara teknis, *Intrusion Detection System* bekerja dengan cara memonitor dan memberikan *flag* terhadap aktivitas mencurigakan yang terjadi dalam jaringan dan langsung di-*report* sebagai *alert*, sehingga menimbulkan dampak terhadap tingginya rata-rata deteksi yang bersifat *false-positive* serta ukuran volume *alert* yang terus membesar, karena *traffic* data yang ada dalam jaringan merupakan suatu hal yang bersifat *non-stationary* [6].

Intrusion Detection System (IDS) memiliki dua buah kelemahan dalam mengenali serangan DDoS. Kelemahan pertama dari *Intrusion Detection System* (IDS), dalam mengenali pola serangan DDoS disebabkan karena adanya defisit dalam protokol TCP/IP [7] sehingga berdampak pada mudahnya untuk memulai serangan DDoS dari sisi penyerang, baik dengan menggunakan *tools* yang ada pada sistem operasi misalnya *tools Ping*, maupun dengan menggunakan *advanced-tools* semisal LOIC. Sementara dari sisi lain, dengan penggunaan protokol TCP/IP, serangan DDoS menjadi terlalu lambat untuk disadari oleh korban karena protokol TCP/IP merupakan sebuah protokol yang umum digunakan untuk mengontrol komunikasi antar *device* dalam jaringan Internet. Kelemahan kedua dari *Intrusion Detection System* (IDS) terletak pada semakin berkembangnya teknik dan metode dalam melancarkan serangan DDoS yang dilaksanakan oleh *hacker*, dimana hal itu sangat sulit untuk diimbangi oleh sistem IDS [8].

Serangan DDoS dengan memfungsikan metode *SYN-Flood* merupakan salah satu contoh semakin berkembangnya teknik serangan DDoS. Dalam aktivitas jaringan Internet, aliran paket data yang memanfaatkan protokol *SYN* merupakan sebuah paket jaringan yang bersifat legal, karena

*) penulis korespondensi (Harsono)
Email: harsonopoltekhb@gmail.com

protokol *SYN* mutlak diperlukan dalam proses otentikasi komunikasi antar perangkat dalam jaringan Internet. Oleh karenanya, ketika protokol *SYN* dimanfaatkan untuk melancarkan serangan DDoS dengan cara *flooding* target, maka *Intrusion Detection System* (IDS) cukup sulit untuk mendeteksinya sebagai artefak abnormal dan berakibat pada tingginya *false-rate alert* yang dibangkitkan oleh *Intrusion Detection System* (IDS).

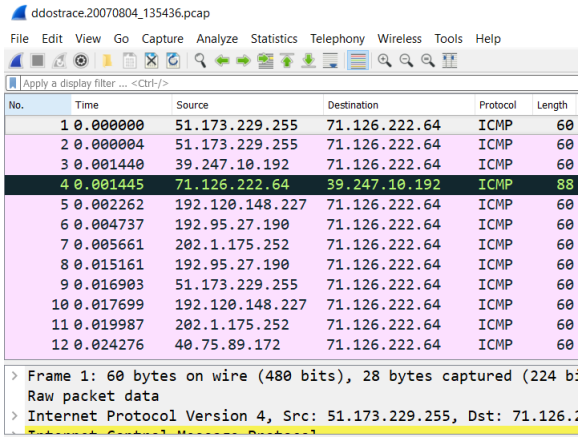
Selain dua kelemahan yang telah dipaparkan, secara umum pada *Intrusion Detection System* (IDS) berbasis *signature*, seiring dengan meningkatnya volume alert yang bersifat *false-positive* maka tentu akan berdampak kepada lambatnya pengamanan jaringan ketika benar-benar terjadi serangan DDoS yang berefek kepada rendahnya efisiensi mitigasi jaringan [9]. Berdasarkan masalah yang telah diutarakan sebelumnya, maka penelitian yang dituangkan dalam paper ini bertujuan untuk membentuk suatu pendekatan baru dalam kaitannya dengan klasifikasi paket jaringan, sehingga bisa menjadi sebuah *framework* pada pengembangan sistem deteksi serangan *Distributed Denial-of-Service* (DDoS). Pendekatan baru yang diusulkan, memanfaatkan metode analisis statistik sebagai fungsi kuantifikasi aktivitas yang digabungkan dengan metode *neural network* sebagai basis klasifikasi. Penelitian yang dilaksanakan mengambil data uji dan data pelatihan dari dataset paket jaringan DDoS *Attack* 2007 yang dirilis oleh *Center for Applied Internet Data Analysis* (CAIDA) [10] sebanyak 1000 data.

II. METODE PENELITIAN

Metode yang digunakan dalam penelitian ini terbagi menjadi beberapa langkah yaitu :

A. *Dataset*

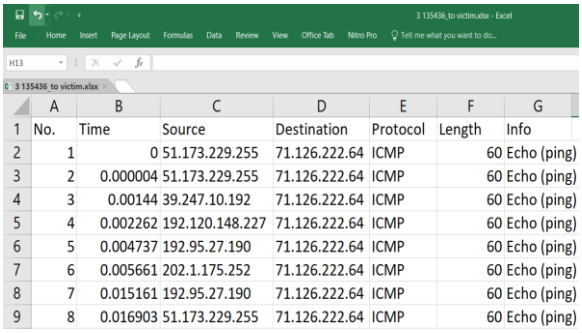
Mengambil dataset paket jaringan DDoS *Attack* 2007 yang dirilis oleh *Center for Applied Internet Data Analysis* (CAIDA) dalam format .pcap, seperti yang disajikan pada Gbr 1.



Gbr. 1 Pengambilan log.

B. *PraProcessing Data*

1) Konversi File: Melakukan konversi format file .pcap menjadi .csv sehingga data paket jaringan dapat diolah dengan perangkat lunak *spreadsheet* seperti yang terjadi pada Gbr 2.



	A	B	C	D	E	F	G
1	No.	Time	Source	Destination	Protocol	Length	Info
2	1	0	51.173.229.255	71.126.222.64	ICMP	60	Echo (ping)
3	2	0.000004	51.173.229.255	71.126.222.64	ICMP	60	Echo (ping)
4	3	0.00144	39.247.10.192	71.126.222.64	ICMP	60	Echo (ping)
5	4	0.002262	192.120.148.227	71.126.222.64	ICMP	60	Echo (ping)
6	5	0.004737	192.95.27.190	71.126.222.64	ICMP	60	Echo (ping)
7	6	0.005661	202.1.175.252	71.126.222.64	ICMP	60	Echo (ping)
8	7	0.015161	192.95.27.190	71.126.222.64	ICMP	60	Echo (ping)
9	8	0.016903	51.173.229.255	71.126.222.64	ICMP	60	Echo (ping)

Gbr. 2 Konversi format .pcap .csv.

2) Kuantifikasi Data: Melakukan kuantifikasi terhadap data paket jaringan yang telah dikonversi menjadi format .csv dengan memanfaatkan perhitungan statistik. Proses kuantifikasi dilaksanakan secara *fixed moving average window* [11] dengan jeda 5 detik. Proses kuantifikasi bertujuan untuk mencirikan karakteristik aktivitas jaringan dalam satu rentang waktu serta memudahkan proses pelatihan dan pengujian klasifikasi data dengan *neural network.* Perhitungan statistik yang digunakan adalah :

- Nilai rerata *(average)* panjang paket jaringan dalam satu *frame* waktu yang telah ditentukan.
- Nilai jumlah keseluruhan paket jaringan dalam satu *frame* waktu yang telah ditentukan.
- Nilai *varians* dari variabel jeda waktu kedatangan paket jaringan yang bersumber dari IP tertentu dalam satu *frame* waktu yang telah ditentukan. Nilai *varians* dihasilkan dari persamaan 1.

$$\sqrt{\frac{\sum(tn-\bar{t})^2}{n}} \quad (1)$$

Dengan tn=waktu kedatangan paket; $\bar{t}$=rerata waktu kedatangan paket; n=jumlah paket.

- Nilai *varians* dari variabel panjang paket jaringan yang bersumber dari IP tertentu dalam satu *frame* waktu yang telah ditentukan. Nilai *varians* dihasilkan dari persamaan 2.

$$\sqrt{\frac{\sum(pn-\bar{p})^2}{n}} \quad (2)$$

Dengan pn=panjang paket; $\bar{p}$=rerata panjang paket; n=jumlah paket.

- Nilai kecepatan paket dalam satu frame waktu yang telah ditentukan, yang dihitung dengan persamaan 3.

$$np * \frac{1}{T.akhir - T.awal} \quad (3)$$

Dengan np=jumlah paket; T.akhir=waktu akhir paket; T.awal=waktu awal paket.

- Nilai jumlah keseluruhan bit data dalam satu frame waktu yang telah ditentukan.

C. Pemodelan Sistem

1) Pemodelan: Melaksanakan pembentukan stuktur *neural network* dengan satu *hidden layer*, dengan jumlah *neuron* sebanyak 2n+1 dimana n adalah jumlah *neuron input* [12].

2) Pelatihan Data: Melaksanakan pelatihan *neural network* dengan 60% data dari hasil kuantifikasi, serta pengujian *neural network* dengan 40% data dari hasil kuantifikasi.

3) Analisis: Melaksanakan analisis kinerja klasifikasi paket jaringan dari metode yang diterapkan.

III. HASIL DAN PEMBAHASAN

Berdasarkan hasil kuantifikasi statistik terhadap data paket jaringan dihasilkan enam jenis karakteristik aktivitas jaringan, dimana keenam karakteristik tersebut digunakan sebagai input dari *neural network* yang memiliki struktur 6-13-2 dengan satu *input layer* yang terdiri dari enam *neuron*, satu *hidden layer* yang terdiri dari tiga belas *neuron*, serta satu *output layer* yang terdiri dari dua *neuron*. *Neural network* memiliki tiga kondisi *output* yang diterjemahkan dalam angka biner. Ketiga kondisi tersebut yang pertama adalah kondisi normal yang diterjemahkan dalam pasangan bilangan biner (0-0). Kondisi kedua adalah kondisi dimana jaringan diserang dengan slow DDoS yang diterjemahkan dalam pasangan bilangan (1-0). Sedangkan kondisi ketiga adalah kondisi dimana jaringan diserang dengan DDoS bervolume besar yang diterjemahkan dalam pasangan bilangan biner (1-1).

Pada penelitian ini digunakan satu jenis struktur *neural network* dengan konfigurasi *neuron hidden layer* 2n+1, dimana n adalah jumlah *input*, dengan alasan bahwa penggunaan satu *hidden layer* telah mencukupi untuk memecahkan permasalahan klasifikasi [13]. Secara ringkas, konfigurasi *layer* dan fungsi aktivasi *neural network* disajikan pada Tabel I.

TABEL I
KONFIGURASI *NEURAL NETWORK*

No	Layer	Jumlah Neuron	Fungsi Aktivasi
1.	Input	6	-
2.	Hidden	13	Logsig
3.	Output	2	Logsig

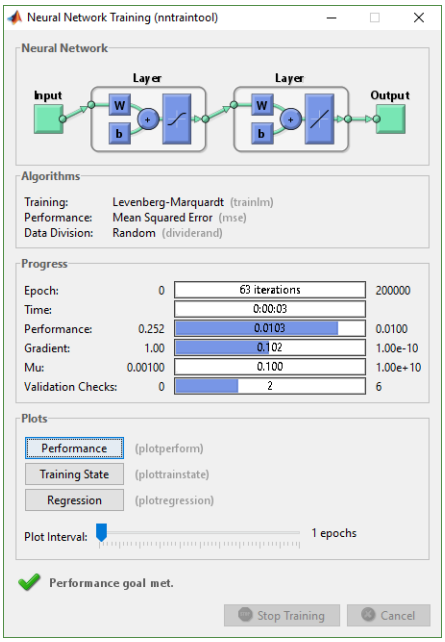
Neural network yang telah disajikan pada Tabel 1, dilatih dengan metode *backpropagation* dengan fungsi pelatihan *levenberg-marquardt*. Fungsi pelatihan *levenberg-marquardt* merupakan jenis fungsi pelatihan yang memiliki *error-rate* terkecil jika dibandingkan dengan fungsi pelatihan lain misalnya fungsi pelatihan *resillent-propagation*, maupun fungsi pelatihan *scaled-conjugate* [13]. Nilai konfigurasi yang digunakan terhadap *neural network* dalam penelitian ini yaitu:

- *Epoch* = 200000.
- *Learning rate* = 0,5.
- *Momentum* = 0,95.
- *Goal Mean Squared-Error* (MSE) = 0,01.

Pelatihan *neural network* yang telah dilaksanakan menghasilkan nilai *Mean Squared-Error* sebesar 0,0103 yang dicapai pada *Epoch* 63, dengan nilai gradien sebesar 0,0102. Nilai *Mean Squared-Error* yang dihasilkan dari proses pelatihan *neural network* telah sesuai dengan konfigurasi dari *Mean Squared-Error* (MSE) yang ditentukan sebelumnya. Kecilnya nilai MSE tersebut menandakan bahwa *neural network* telah mampu digunakan untuk menggeneralisasi suatu *input* baru.

Sedangkan dari sisi performa, *neural network* yang telah dilatih menghasilkan nilai *regresi R-test* sebesar 0,98957 yang berarti bahwa bobot-bobot koneksi antar *neuron* pada setiap layer *neural network* telah mampu memberikan hasil yang optimal dalam mengenali pola data *input.* Hasil pelatihan *neural network* disajikan pada Gbr 3. Sedangkan hasil plot performa *neural network* disajikan pada Gbr 4.

Langkah selanjutnya setelah melaksanakan pelatihan terhadap *neural network*, adalah menguji *neural network* dengan data uji untuk mendapatkan persentase rata-rata pengenalan terhadap tiga kondisi yang telah ditentukan sebelumnya.



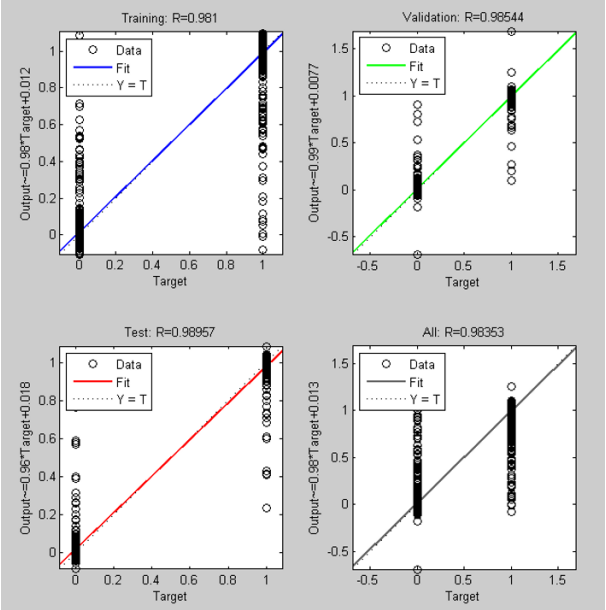
Gbr. 3 Hasil pelatihan *neural network.*

TABEL II
HASIL UJI PENGENALAN NEURAL NETWORK

No	Kondisi	Data Uji	Data Error	Prosentase Pengenalan
1.	Normal	133	9	93,23%
2.	Slow DDoS	133	12	90,97%
3.	DDoS	134	7	94,77%
Total Data		400	28	92,99%

Berdasarkan pengujian didapatkan bahwa rerata persentase akurasi klasifikasi *neural network* didapatkan sebesar 92,99% seperti yang tersaji pada Tabel II.

ISSN: 2477-5126
e-ISSN: 2548-9354



Gbr. 4 Hasil performansi pelatihan *neural network.*

Rerata persentase hasil pengujian klasifikasi *neural network* terhadap data *input* baru berada diatas 90%. Tingginya persentase tersebut mengindikasikan bahwa metode kuantifikasi data secara statistik terhadap aliran paket data jaringan yang digabungkan dengan *neural network* mampu digunakan untuk mengklasifikasi aktivitas paket data dalam jaringan Internet dan dapat dijadikan sebagai landasan ataupun *framework* dalam mengambangkan sistem deteksi serangan *Distributed Denial-of-Service* (DDoS).

IV. KESIMPULAN

Pendekatan yang diusulkan untuk menyelesaikan permasalahan klasifikasi paket jaringan dengan metode kuantifikasi data secara statistik terhadap aliran paket data jaringan yang digabungkan dengan *neural network* sebagai basis *classifier* menghasilkan nilai rerata persentase akurasi klasifikasi sebesar 92,99%, sehingga dapat disimpulkan bahwa pendekatan yang diusulkan dalam penelitian ini dapat dijadikan sebagai landasan ataupun *framework* dalam mengambangkan sistem deteksi serangan *Distributed Denial-of-Service* (DDoS). Untuk menghasilkan nilai rerata akurasi klasifikasi yang lebih baik, maka terdapat beberapa parameter *neural network* yang dapat dioptimalkan, yaitu: (1) Menambah jumlah data pelatihan; (2) Memvariasikan jumlah *neuron* serta *hidden layer neural network*; (3) Mengoptimasi konfigurasi *neural network* dari sisi *momentum; learning-rate; epoch;* serta *Mean Squared-Error*); (4) Mengoptimasi fungsi pelatihan, serta fungsi aktivasi *layer neural network*.

Diharapkan dengan terbentuknya paradigma baru dalam klasifikasi paket data jaringan Internet, dapat menjadi landasan serta komplemen terhadap *Intrusion Detection System* (IDS) dalam meminimalkan resiko terhadap serangan *Distributed Denial-of-Service* (DDoS) pada sebuah organisasi yang terhubung dengan dunia Internet.

UCAPAN TERIMA KASIH

Ucapan terima kasih penulis kepada pihak yang membantu ataupun memberikan dukungan terkait dengan penelitian yang dilakukan.

DAFTAR PUSTAKA

[1] N. Hoque, D. K. Bhattacharyya, and J. K. Kalita., 2015. Botnet in DDoS Attacks : Trends and Challenges. vol. 17, no. 4, pp. 2242–2270.
[2] A. Networks., 2016. World Wide Infrastrucure Security Report 2015,.
[3] S. Khattak, N. R. Ramay, K. R. Khan, A. A. Syed, and S. A. Khayam, 2014. A Taxonomy of Botnet Behavior . vol. 16, no. 2, pp. 898–924.
[4] I. Riadi, A. W. Muhammad, and Sunardi., 2016. Integrasi Metode Normalized Relative Network Entropy dan Neural Network Backpropagation (BP) untuk Deteksi dan Peramalan Serangan DDoS. *Natl. 4th APPPTM Indones. Conf.*, vol. 4.
[5] R. Khattak and Z. Anwar., 2016. D3TAC : Utilizing Distributed Computing for DDoS Attack Traffic Analysis on the Cloud.
[6] C. Fachkha and M. Debbabi., 2015. *.Darknet as a Source of Cyber Intelligence : Survey , Taxonomy and Characterization*, no. c.
[7] R. F. Fouladi, C. E. Kayatas, and E. Anarim., 2016. Frequency Based DDoS Attack Detection Approach Using Naive Bayes Classification. *39th Int. Conf. Telecommun. Signal Process.*, pp. 104–107.
[8] I. Riadi, A. W. Muhammad, and Sunardi., 2017. Neural Network-Based DDoS Detection Regarding Hidden Layer Variation. *J. Theor. Appl. Inf. Technol.*, vol. 95, pp. 1–9.
[9] E. Balkanli, J. Alves, and A. N. Zincir-Heywood., 2014. Supervised Learning to Detect DDoS Attacks. *IEEE Int. Conf. Adv. Comput. Commun. Informatics*.
[10] W. Bhaya and M. E. Manaa., 2014. A Proactive DDoS Attack Detection Approach Using Data Mining Cluster Analysis. vol. 5, no. 4, pp. 36–47.
[11] N. Furutani, T. Ban, J. Nakazato, J. Shimamura, J. Kitazono, and S. Ozawa., 2014. Detection of DDoS Backscatter Based on Traffic Features of Darknet TCP Packets. pp. 0–4.
[12] T. Zhao, D. C. T. Lo, and K. Qian., 2015. A Neural Network Based DDoS Detection System Using Hadoop and HBase. *Proc. - 2015 IEEE 17th Int. Conf. High Perform. Comput. Commun. 2015 IEEE 7th Int. Symp. Cybersp. Saf. Secur. 2015 IEEE 12th Int. Conf. Embed. Softw. Syst. H*, pp. 1326–1331.
[13] I. Riadi, A. W. Muhammad, and Sunardi., 2017. Network Packet Classification Using Neural Network Based on Training Function and Hidden Layer Neuron Number Variation. *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 6, pp. 1–4.