

Visualisasi Serangan *Denial Of Service* Dengan *Clustering* Menggunakan *K-Means Algorithm*

Napsiah

Computer Engineering Department
Faculty of Computer Science
Sriwijaya University, Inderalaya 30662
South Sumatera Indonesia
09121001065@students.ilkom.unsri.ac.id

Deris Stiawan

Computer Engineering Department
Faculty of Computer Science
Sriwijaya University, Inderalaya 30662
South Sumatera Indonesia
deris@unsri.ac.id

Ahmad Heryanto

Computer Engineering Department
Faculty of Computer Science
Sriwijaya University, Inderalaya 30662
South Sumatera Indonesia
hery@unsri.ac.id

Abstrak—Visualisasi menjadi salah satu solusi dalam menampilkan serangan di *network*. Dengan memvisualisasikan serangan, akan lebih mudah dalam mengenali dan menyimpulkan pola dari gambar visual yang kompleks. Target serangan *DoS* bisa ditunjukkan ke berbagai bagian jaringan, bisa ke *routing*, *web*, *electronic mail* atau *server DNS (Domain NameSystem)*. Tujuan dari serangan *DoS* membuat *server shutdown*, *reboot*, *crash* atau *not responding*. Pola serangan *DoS* pada dataset *ISCX* membentuk sebuah pola dimana banyaknya *IHost* yang hanya meng-*exploit* ke satu *server*. *Snort* mendeteksi adanya serangan *DoS* pada dataset *ISCX testbed 14 juni* sebanyak 42 *alert HttpDoS attack*. Persentase akurasi dari program *clustering* menggunakan algoritma *k-means* sebesar 97,83%, untuk *detection rate* nya sebesar 98,63%, dan *false alarm* dari program sebesar 0,02%. Sedangkan, nilai persentase akurasi dari *clustering* menggunakan algoritma *k-means* dengan tool *WEKA* sebesar 99,69%, *detection rate* 99,01% dan *false alarm* sebesar 3,70%.

Keywords— Visualisasi; *DoS*; *Clustering*; *K-means Algorithm*

I. PENDAHULUAN

Teknik serangan yang sering dilakukan oleh attacker dalam melumpuhkan sistem terbagi dalam beberapa macam, salah satu teknik serangan yang umum digunakan oleh attacker adalah *Denial of Service (DoS)*. Serangan *DoS* menghasilkan kerusakan yang sifatnya persisten, artinya kondisi *DoS* akan tetap terjadi walaupun attacker sudah berhenti menyerang, dan server akan kembali normal setelah di *re-start/reboot* [1]. Serangan *DoS* mudah untuk diimplementasi, akan tetapi sulit untuk mencegah dan melacaknyanya.

Pada penelitian [2], [3], [4], [5] membahas permasalahan terhadap visualisasi serangan secara otomatis menggunakan *parallel coordinate attack visualization (PCAV)*. Penelitian ini, mendeteksi serangan internet dalam skala besar yang tidak diketahui seperti internet worms, *DoS attack* dan aktifitas *network scanning*. *PCAV* menampilkan *traffic* jaringan pada bidang koordinat paralel menggunakan informasi seperti *source IP address*, *destination IP address*, *source port* dan *packet length*. Penelitian [2], [5] menjelaskan, bahwa setiap jenis serangan secara signifikan membentuk pola yang unik. Pada penelitian [6], melakukan penelitian terhadap serangan *DoS* dengan cara mengklasifikasikan serangan

menggunakan jaringan syaraf tiruan *LVQ (Learning Vector Quantization)*, dengan menghasilkan tingkat keberhasilan yang tinggi untuk kondisi normal, *PING flood* dan *UDP flood*. Dengan demikian, penelitian akan difokuskan pada salah satu pendekatan yaitu *attack visualization* terhadap serangan *Denial of Service (DoS)* dengan *clustering* menggunakan *K-Means Algorithm*.

Dataset yang digunakan pada penelitian ini yaitu dataset (*Information Security Center of eXcellence*) *ISCX* yang dikembangkan oleh fakultas ilmu computer, universitas *newBrunswick* dari tahun 2009 sampai tahun 2011. Kumpulan data simulasi yang terdapat pada dataset *ISCX* adalah infiltrasi jaringan dari dalam, *HTTPDoS*, *DDoS* menggunakan *IRCbotnet*, dan *brute-force SSH*.

II. METODOLOGI PENELITIAN

A. Pengenalan Pola Paket Data

Langkah pertama dalam pengenalan pola paket data adalah menganalisa antara skenario serangan dari dataset dengan hasil *traceroute*. Selanjutnya, akan dibandingkan dengan program perhitungan paket data yang dominan disetiap baris paketnya. Program perhitungan paket data dilakukan secara *off-line*, menggunakan dataset dalam format *csv* sebagai data inputan. Acuan perhitungan pada program berdasarkan daftar *features* yang terdapat pada dataset *ISCX* (tabel 1), dimana *features* ini diperoleh dari *extraction raw* paket data kedalam bentuk *csv*.

TABLE I. DAFTAR FEATURES DATASET *ISCX*

No	Nama Features	No	Nama Features
1	<i>appName</i>	11	<i>Source TCP flags</i>
2	<i>Total source bytes</i>	12	<i>Destination TCP flags</i>
3	<i>Total destination bytes</i>	13	<i>Source</i>
4	<i>total destination packet</i>	14	<i>Protocol name</i>
5	<i>Total source packet</i>	15	<i>Source port</i>
6	<i>Source payload as base 64</i>	16	<i>Destination</i>
7	<i>Source payload as base UTF</i>	17	<i>Destination port</i>
8	<i>Destination payload as base 64</i>	18	<i>Start date time</i>
9	<i>Destination payload as base UTF</i>	19	<i>Stop date time</i>
10	<i>Direction</i>		

Prosiding
ANNUAL RESEARCH SEMINAR 2016
 6 Desember 2016, Vol 2 No. 1

ISBN : 979-587-626-0 | UNSRI

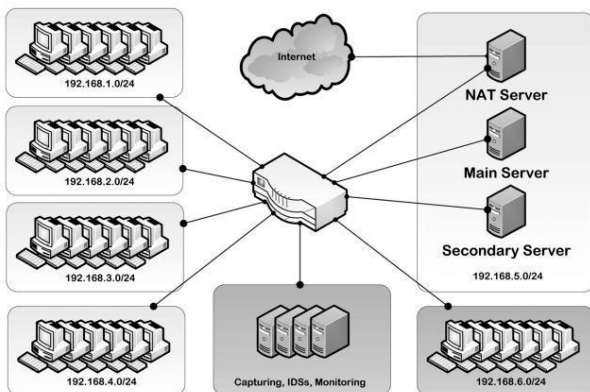
http://ars.ilkom.unsri.ac.id

B. Skenario Dataset ISCX

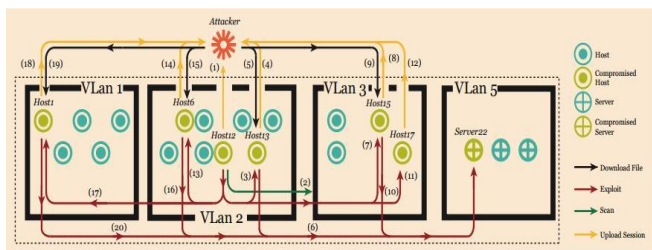
Dataset ISCX memiliki arsitektur jaringan testbed (gambar 1) yang terdiri dari 21 *windows workstation* yang saling berhubungan, dua mesin berbasis *linux Ubuntu* dan satu mesin diinstal *windows server 2003*. Tahapan dalam menghasilkan dataset ISCX terdiri dari empat elemen :

6. *Probe*, yang bertujuan untuk mengumpulkan informasi
7. *Identifikasi vulnerability*
8. *Mempertahankan akses dengan menciptakan backdoors*
9. *Kemampuan untuk secara efektif menutupi jalur penyerangan.*

Dataset ISCX yang digunakan dalam penelitian ini adalah hasil testbed 14 juni yang merupakan serangan *HTTPDoS*. Skenario serangan *HTTPDoS* (gambar 2) dirancang tanpa membanjiri jaringan, sehingga *bandwidth* yang dibutuhkan rendah. ISCX memanfaatkan *slowloris* sebagai alat utama dalam skenario seranga *HTTPDoS*.



Gambar 1. Arsitektur Jaringan Testbed ISCX [7]



Gambar 2. Ilustrasi Skenario HTTPDoS Dataset ISCX [7], [8]

C. Perhitungan Paket Data Dominan

Pada tahapan pengenalan pola dari perhitungan paket data dominan menggunakan sebuah program yang akan melakukan *string matching* terhadap setiap paket data. Dari program inidapat dilihat bagaimana pola paket normal dan pola paket serangan *DoS*. Berikut *pseudocode* programnya .:

Pseudocode : Perhitungan Probability dari serangan Denial of Service (DoS) pada Dataset ISCX

```

Input : dl
Output : Banyaknya kemungkinan
        HttpDoS attack

Step 1: Connect to
Database
for dataset connect with
Microsoft Access
P ← Microsoft.Jet.OLEDB.4.0
    pc = p + ds + ep + E
End for
Od conn = new Od(pc)

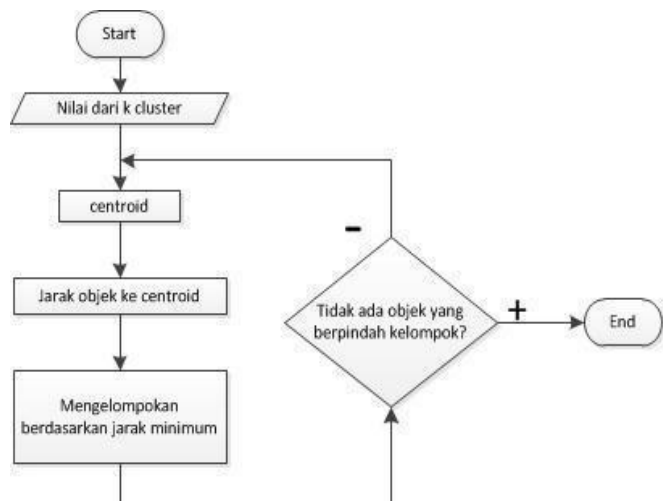
Step 2: proses data
if features= A
read dl
query ← SELECT*
count(StartDateTime)as
Jumlah from dl
where appName=HTTPWeb
groupBy features
orderBy count DESC
End if

Step 3: fill data gridview
ODA myDA = new ODA(query,conn) DT
dt ← new DT
myDA.fill(dt) dG.DS = dt
    
```

D. Clustering K-Means Algorithm

Langkah awal dari *k-means clustering* adalah menentukan nilai *k* sebagai jumlah *cluster* yang ingin dibentuk, kemudian mengasumsikan *k centroid* (titik pusat *cluster*) awal secara *random*. Langkah selanjutnya dari algoritma *k-means* yang harus dilakukan sampai objek benar-benar stabil (tidak ada objek yang berpindah kelompok) yaitu :

- 4 Menentukan koordinat *centroid*
- 5 Menentukan jarak dari masing-masing objek ke *centroid*
- 6 Mengelompokan objek berdasarkan jarak minimum (sampai menemukan *centroid* terdekat).



Gambar 3. Proses Clustering K-means

Dari pengolahan dataset ISCX 14 juni menggunakan program pengenalan pola paket data dominan, maka dari data tersebut diambil secara acak *centroid* awal untuk data normal dan data serangan sebagai pola dalam meng-cluster data pada algoritma K-means. Selanjutnya dilakukan perhitungan jarak antara paket satu dan paket lainnya menggunakan rumus *Euclidean distance* [9], [10]. Berikut pseudocode dari program clustering menggunakan algoritma K-means :

Pseudocode : Program clustering menggunakan algoritma k-means pada Dataset ISCX 14 juni

```

Input : F = {m1, m2, ... mn} #dataset ISCX
K = banyak cluster
pola={p1, p2, ..., pn}
Output : dos = cluster attack
         normal = cluster normal
for p1 ∈ pola do
    p1 ← mi ∈ F
end
def kmeans do
    a (m1) ← sqrt (mn - p1)2 n ∈ {1...n}
    return a
end
def hitung centroid baru do
    h (m1) ← i/len (cluster)
    return h
end
centroidbaru ← none
cluster ← []
i77 ← none
while pola !=centroidbaru
    i77+1
    i++
for i in range (len(F))
    
```

```

for h in range (len(cluster_attack))
for k in range (len(cluster_normal))
jarak ← kmeans (m1, p1)
if jarak = min(jarak) then
    cluster ← m1
end
polaupdate ← newcentroid(clusterattack)
matrikattack ← polaupdate
polaupdate ← newcentroid(clusternorm)
matriknormal ← polaupdate
end
    
```

Langkah berikutnya setelah melakukan *clustering* paket data adalah memvisualisasikan paket data tersebut kedalam dua *cluster* yaitu paket data *attack* dan paket data normal. Berikut pseudocode program visualisasi nya :

Pseudocode : Program visualisasi di Dataset ISCX 14 juni

```

Input:F
    cluster
Output : plot
for cluster ∈ F do
    cluster ← F
    cluster=attack and cluster=normal
Return plot
Plot.show
end
    
```

III. HASIL DAN PEMBAHASAN

Visualisasi serangan DoS menggunakan clustering dengan algoritma k-mean diterapkan untuk mempermudah dalam mengenali pola serangan DoS. Sehingga, dengan cara memvisualisasikan akan lebih mudah dalam mengatasi serangan yang ada.

3.1. Pengujian Traceroute Dataset ISCX

Berikut gambar 4 akan menampilkan hasil dari *trace routedataset ISCX testbed* 14 juni. Hasil dari *traceroute* padagambar 4 berikut menunjukkan, bahwa skenario dataset ISCX :

1. Dalam melakukan skenario serangan *HTTPDoS, attacker* memanfaatkan koneksi dari *host*
 - a. 192.168.2.112, 192.168.2.113, 192.168.3.115,
 - b. 192.168.3.117, 192.18.1.101 dan 192.18.2.106 yang bertindak sebagai penyerang untuk melakukan *exploitke IP server* 192.168.5.122 melalui *backdoor*.
2. Kemudian *attacker* melakukan koneksi dengan cara *upload session ke host*.

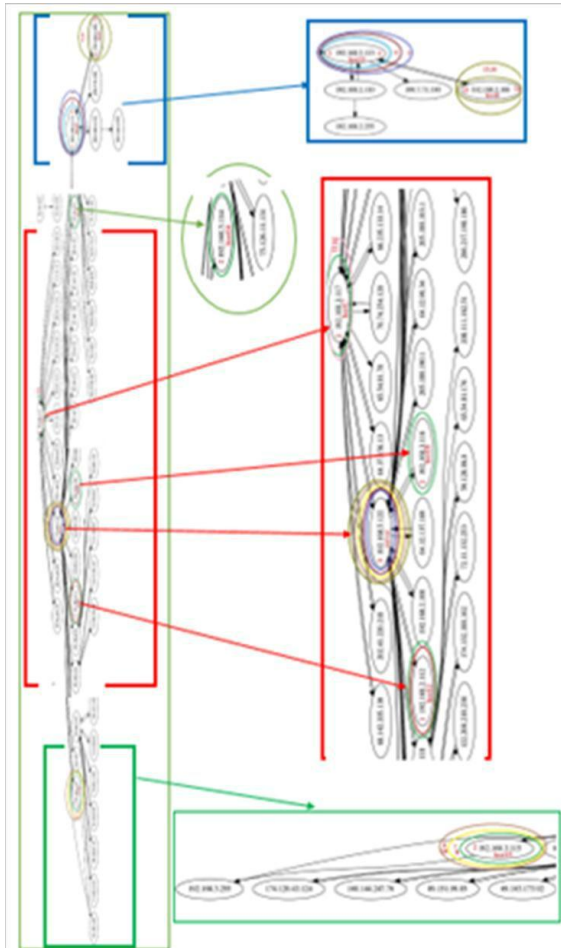
Prosiding ANNUAL RESEARCH SEMINAR 2016

6 Desember 2016, Vol 2 No. 1

ISBN : 979-587-626-0 | UNSRI

http://ars.ilkom.unsri.ac.id

3. Selanjutnya host 192.168.2.112 melakukan scan terhadap host 192.168.2.113, 192.168.3.115 dan 192.168.3.117. Scan dilakukan untuk mengetahui vulnerability dari sistem agar dapat di exploit.
4. Setelah mengetahui vulnerability, maka attacker akan me-remote sistem yang kemudian dengan leluasa attacker dapat meng-exploit serangan ke sistem.
5. Dan setelah attacker melakukan serangan terhadap server, maka host yang telah digunakan oleh attacker akan men-download file mailicious dari remote server yang telah dikendalikan oleh attacker.



Gambar 4. Hasil Traceroute Dataset ISCX Testbed 14

B. Pengujian Program Perhitungan Paket Data Dominan

Dilihat dari hasil pengujian program perhitungan paket data yang dominan pada gambar 5 berikut, Pola baris paket data pada dataset ISCX mengacu pada karakteristik serangan

Denial of Service [6], [7] :

- [1] *Attacker* memanfaatkan koneksi dari *IP source* sistem melalui backdoor untuk meng-exploit server.
- [2] *Port numbers* yang digunakan secara acak dari jumlah paket palsu.
- [3] Ukuran *window* dan *packet length* yang tetap selama serangan dilakukan.
- [4] *Flags* dalam *protocol TCP* dimanipulasi hanyamelakukan *SYN* dan *ACK*.
- [5] *HTTP requests* dibanjiri melalui port 80.

Gambar 5. Hasil Perhitungan Paket Data Dominan Pengujian Menggunakan Snort

Sebagai dasar untuk memvalidasi apakah benar hasil dari pengenalan pola dari program diatas adalah serangan DoS, maka akan dibandingkan dengan hasil dari engine IDS yang dapat membedakan antara paket data normal dan paket data serangan DoS. Berikut hasil pengujian :

TABLE II. JUMLAH ALERT TERDETEKSI PADA PENGIJIAN ISCX DATASET (14 JUNI 2010)

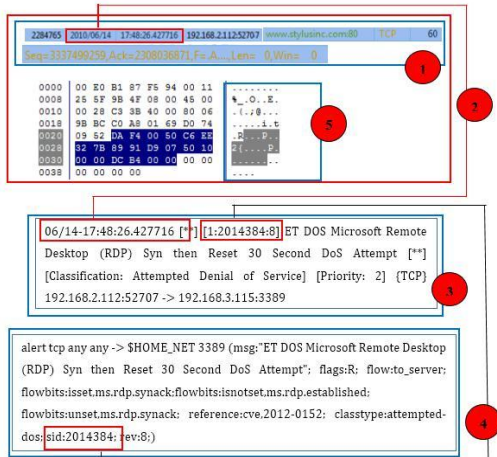
No.	Klasifikasi alert Terdeteksi	SID	Priority	Total
1	Misc activity	1:2925:3	3	18264
2	Generic protocol command decode	1:1748:8	3	7310
3	Attempted adm privilege gain	1:2546:14	1	3582
4	Attempted information leak	1:1201:13	2	770
5	Access to potentially vulnerable web	1:1721:18	2	284
6	Attempted DoS	1:2014384:8	2	42

Prosiding
ANNUAL RESEARCH SEMINAR 2016
 6 Desember 2016, Vol 2 No. 1

ISBN : 979-587-626-0 | UNSRI

http://ars.ilkom.unsri.ac.id

Berikut ekstraksi dan korelasi data hasil pengujian untuk dataset ISCX :



Gambar 6. Ekstraksi dan Korelasi Data Hasil Pengujian Snort Dataset ISCX

- b) Menentukan nilai dari k cluster, k cluster dalam
- c) program clustering menggunakan k -mean ini terdiri dari dua k cluster yaitu cluster attack dan cluster normal.
- d) Menentukan centroid awal, centroid awal pada program ini ditentukan secara acak dari hasil program sebelumnya yaitu program perhitungan paket data dominan.

Gambar 9. Pola Paket Dataset ISCX 14

F. Pengujian Dataset menggunakan Algoritma K-Means Clustering

Pada tahapan pengujian menggunakan algoritma k-means clustering, dilakukan dua macam percobaan. Percobaan pertama dilakukan clustering dataset menggunakan program dengan bahasa pemrograman "python". Sedangkan, pada tahapan pengujian kedua dataset dilakukan percobaan menggunakan tool yang sudah sering digunakan oleh para penelitian [11], [12] yaitu "WEKA".

- e) Menghitung jarak paket data ke centroid menggunakan rumus *Euclidean distance* [9], [10].

$$d(x, y) = |x - y| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

Dari perhitungan menggunakan rumus diatas maka didapatkan hasil jarak setiap paket ke centroid :

d) 1. Pengujian Dataset ISCX dengan Program Clustering menggunakan Algoritma K-means

a) Normalisasi data

Gambar 7. Paket Dataset ISCX 14 Juni.csv before- After Normalisasi

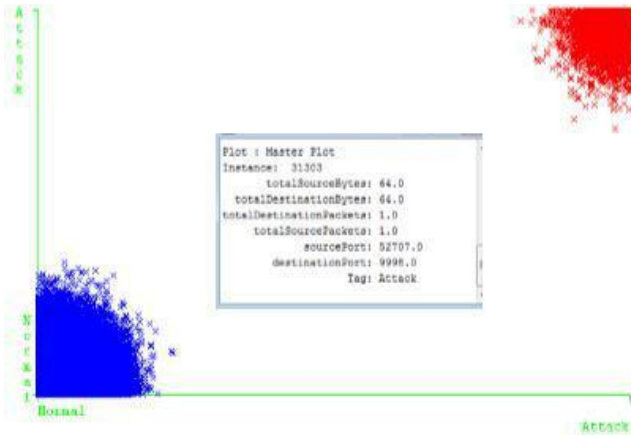
Gambar 10. Hasil Perhitungan Jarak Antar Paket di Dataset ISCX 14 Juni

Prosiding
ANNUAL RESEARCH SEMINAR 2016

6 Desember 2016, Vol 2 No. 1

ISBN : 979-587-626-0 | UNSRI

http://ars.ilkom.unsri.ac.id



Gambar 15. Visualisasi Serangan DoS dengan Clustering menggunakan algoritma K-means Dataset ISCX dengan WEKA

Hasil akhir *clustering* pada gambar diatas menggambarkan setiap paket pada dataset, dimana bentuk visual normal dan *attack* ditampilkan berdasarkan sumbu x dan sumbu y yang digunakan. Berikut hasil perhitungan *confusion matrix* dari *clustering* menggunakan algoritma *k-means* yang terdeteksi untuk setiap *cluster*.

TABLE V. PERHITUNGAN *CONFUSION MATRIX* TOOL WEKA

Jenis Paket	Jumlah Paket	Accuracy %	Detection %	False Alarm %
TN	53907	99,69	99,01	3,70
FP	20			
FN	151			
TP	2010			

IV. KESIMPULAN DAN SARAN

2.1. Kesimpulan

Serangan *Denial of Service* di dataset *ISCX* memiliki pola sebagai berikut : banyaknya *IP host* yang hanya meng-*exploit* ke satu *IP server*, *HTTP request* dibanjiri melalui port 80, ukuran *window* dan *packet length* yang tetap selama serangan, flags dalam protokol TCP dimanipulasi hanya melakukan *SYN* dan *ACK*, *port numbers* secara acak dari jumlah paket palsu. Sedangkan, paket normal membentuk

pola yaitu : banyaknya satu *source* ke banyak *destination*, banyak *source* ke satu *destination* dan satu *source* ke satu *destination*. Persentasi akurasi dari program *clustering* menggunakan algoritma *k-means* sebesar 97,83%, untuk *detection rate* nya sebesar 98,63%, dan hasil perhitungan *confusion matrix* untuk *false alarm* dari program sebesar 0,02%. Nilai persentase akurasi dari *clustering* menggunakan algoritma *k-means* dengan *tool WEKA* yang dihitung menggunakan *confusion matrix* menghasilkan nilai *accuracy* 99,69%, *detection rate* 99,01% dan *false alarm* sebesar 3,70%.

B. Saran

Pada penelitian lanjutan, ada baiknya mencoba untuk melakukan visualisasi menggunakan dataset dengan topologi sendiri dan secara *real-time*.

REFERENSI

1. J. J. Siregar, "Web Denial Of Service Attack," no. 9, pp. 1199–1205.
2. Y.-J. Yang and Y.-H. Liu, "A DoS Attack Situation Visualization Method Based on Parallel Coordinates," 2012 IEEE 12th Int. Conf. Comput. Inf. Technol., pp. 340–344, 2012.
3. H. Choi, H. Lee, and H. Kim, "Fast detection and visualization of network attacks on parallel coordinates," *Comput. Secur.*, vol. 28, no. 5, pp. 276–288, 2009.
4. H. Kim, I. Kang, and S. Bahk, "Real-time visualization of network attacks on high-speed links," *IEEE Netw.*, vol. 18, no. 5, pp. 30–39, 2004.
5. H. Kim, I. Lee, J. Cho, and J. Moon, "Visualization of Network Components for Attack Analysis," 2009.
6. P. A. R. Kumar and S. Selvakumar, "Distributed denial of service attack detection using an ensemble of neural classifier," *Comput. Commun.*, vol. 34, no. 11, pp. 1328–1341, 2011.
7. M. Kale, "DDoS Attack Detection Based on an Ensemble of Neural Classifier," vol. 14, no. 7, pp. 122–129, 2014.
8. R. F. Malik and V. Puspita, "Classification Denial Of Service (DoS) Attack Using Artificial Neural Network Learning Vector," no. August, pp. 20–21, 2014.
9. B. Santoso, "Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis." Yogyakarta: Graha Ilmu, 2007.
10. K. Sholeh, B. D. Setiawan, and I. Cholissodin, "Implementasi Metode K-Means Clustering untuk Pembangunan Aturan Fuzzy pada Klasifikasi Ketahanan Hidup Penderita Kanker Payudara," pp. 1–9.
11. N. Sharma, A. Bajpai, and R. Litoriya, "Comparison the various clustering algorithms of weka tools," vol. 2, no. 5, pp. 73–80, 2012.
12. S. Jain, "K-means Clustering Using WEKA Interface," 2010.