BAB I
PENDAHULUAN

1.1. Latar Belakang Masalah

Motivasi dalam penelitian ini berawal dari laporan (ID-SIRTII/CC, 2014) terkait serangan yang sering terjadi pada jaringan internet khususnya Indonesia, serangan tersebut diantaranya adalah: EXPLOIT, DoS, SQL, WEB-MISC, BOTNET-CNC, DDoS, BAD-TRAFFIC, ORACLE, WEB-CLIENT, POLICY. Fokus pada penelitian ini adalah jenis serangan *DoS* yang memiliki cara kerja dengan cara mengirimkan permintaan massal pada *server* sehingga tidak mampu melayani permintaan dengan baik (Kak, 2015). Serangan *DoS* terjadi ketika penyerang membanjiri jaringan dengan data, sering kali pengguna tidak menyadari bahwa sistemnya menjadi target, (Kurniawan, 2012) hal ini disebabkan karena kesulitan dalam membedakan adanya serangan dengan tidak ada serangan. Menurut (Tan, Jamdagni, He, Nanda, & Liu, 2014) mengenali serangan *DoS* dapat dilakukan dengan memantau lalu lintas jaringan, yang bertujuan untuk mengenali ada tidaknya serangan namun ternyata masih ditemui beberapa kendala.

Berdasarkan penjelasan diatas maka penelitian ini bertujuan untuk mengenali bahaya atas serangan pada jaringan dengan mengelompokkan *log* menjadi 3 (tiga) tingkat bahaya, yaitu: rendah, sedang & tinggi yang dievaluasi berdasarkan frekuensi data atribut *totlength* & *totTCPlength* berdasarkan jam, dengan melakukan analisis tingkat kerapatan (densitas) data berdasarkan nilai *Index Davies-Bouldin* (*DBI*). Nilai *index DBI* diharapkan dapat menunjukan tingkat kerapatan data pada *cluster* serta keterpisahan *cluster*, analisis yang dilakukan pada tingkat kerapatan dan keterpisahan berdasakan nilai DBI yang didapatkan diharapkan dapat menunjukan adanya serangan *DoS*.

Penelitian dilakukan di Usaha Kecil dan Menengah (UKM) penyedia akses internet Mandala Citra Media di Surakarta selama 2 (dua) minggu, perusahaan yang baru berusia 1 (satu) tahun dalam menjalankan bisnisnya membutuhkan pengembangan dalam pengelolaan jaringan agar kualitas layanan terjaga. Proses penelitian dimulai dari masukan berupa basisdata *log* yang didapatkan selama proses pemantauan, kemudian basisdata tersebut dikelompokan dengan teknik *clustering* menggunakan algoritma *density k-means* yang bertujuan dalam mengelompokkan tingkat bahaya rendah, sedang dan tinggi. *Clustering* dengan *density k-means* dimanfaatkan untuk membantu menangani kompleksitas dalam mengolah basisdata besar.

Clustering berada pada urutan ke 4 (empat) sebagai teknik yang digunakan dalam mengelompokkan data (S. H. Liao, Chu, & Hsiao, 2012) untuk memudahkan dalam mengenali ada tidaknya serangan *DoS*. *Clustering* akan melakukan pemisahan data ke dalam sejumlah kelompok menurut karakteristik tertentu dalam pekerjaan pengelompokkan berdasarkan perhitungan jarak data yang lebih dekat dibandingkan dengan data yang lain (Prasetyo, 2012, 2014), penentuan *clustering* dalam penelitian ini dikarenakan karakteristik basisdata *log* memiliki ukuran dan *record* besar, sehingga diperlukan pengelompokkan informasi atas basis data *log* untuk mengetahui ada tidaknya serangan *DoS*, *clustering* dapat dilakukan dengan algoritma *density k-means*. Algoritma *Density k-means* pada intinya adalah aktifitas pengelompokkan catatan-catatan yang memiliki kemiripan atribut akan dikelompokkan kedalam salah satu dari sekian kelompok. Adapun catatan-catatan yang kurang memiliki kesamaan atribut akan ditempatkan pada kelompok yang berbeda (Susanto & Suryadi, 2010), dalam statistika dan pembelajaran mesin, algoritma *k-means* merupakan metode analisis kelompok yang mengarah pada pemartisian *N* obyek pengamatan ke dalam *K* kelompok (*cluster*) dimana setiap obyek pengamatan dimiliki oleh sebuah kelompok dengan *mean* (rata-rata) terdekat (Prasetyo, 2012). *Clustering* merupakan

pekerjaan yang memisahkan data (vektor) ke dalam sejumlah kelompok menurut karakteristik masing-masingnya. Tidak diperlukan label kelas, untuk setiap data yang diproses dalam *clustering* karena nantinya label baru bisa diberikan ketika *cluster* sudah terbentuk, karena tidak adanya label kelas untuk setiap data, maka *clustering* sering disebut juga pembelajaran tidak terbimbing (*unsupervised learning*) (Prasetyo, 2014). Data yang telah tergabung dalam *cluster* yang terbentuk kemudian ditentukan seberapa dekat (*density*) data dengan *centroid*-nya.

Ada tidaknya serangan dalam jaringan internet dapat diketahui dengan memantau aktifitas jaringan, aktifitas tersebut disebut sebagai *log* transaksi (Hariyanto, 2012), *log* transaksi menghasilkan format dan struktur yang beragam sehingga *log* transaksi dapat dibuat tergantung informasi yang diperlukan, pemanfaatan informasi ini digunakan untuk mendukung administrator mengenali apa yang terjadi (Larsen & Jensen, 2003). *Log* berasal dari lalu lintas jaringan berfungsi untuk mengenali ada tidaknya serangan *DoS*. *Log* disimpan dalam format asli berbentuk teks kemudian disimpan dalam basis data. *Log* memiliki ukuran besar, oleh karena itu perlu dilakukan beberapa tindakan untuk mempermudah proses penyimpanan dan pencarian informasi dalam basis data tersebut. Peningkatan hasil penyimpanan dan pencarian basis data diperlukan untuk mempercepat mengenali ada tidaknya serangan *DoS*. *Log* dalam penelitian ini dihasilkan oleh `tcpdump`. Hasil luaran `tcpdump` disimpan ke dalam berkas teks dan basis data, berkas teks akan digunakan untuk melakukan proses verifikasi dalam tahapan akhir penelitian ini mengenai banyak paket data berdasarkan jam terjadinya suatu insiden.

1.2. Perumusan Masalah

Berdasarkan pada latar belakang yang telah dijelaskan sebelumnya, maka penelitian ini merumuskan permasalahan sebagai berikut:

a. Apakah algoritma *density k-means* dapat digunakan dalam mengelompokkan basisdata *log* berdasarkan frekuensi *totlength* & *tottcplength*) yang dikelompokan berdasarkan jam yang bertujuan mengelompokkan tingkat bahaya serangan *DoS* berdasarkan hasil analisis densitas data.
b. Apakah algoritma *density k-means* dengan validasi *cluster* Davies Bouldin Index dapat menunjukan nilai densitas data & separasi *cluster* dari *log* basisdata serangan *Denial of Service*.

1.3. Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

a. Mengelompokkan basisdata *log* serangan *denial of service*.
b. Melakukan analisis kerapatan data pada *cluster* & keterpisahan antar *cluster log* basisdata serangan *Denial of Service*.

1.4. Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

a. Administrator

   Memudahkan dalam mengelompokkan *log* terkait serangan *denial of service*.
b. Akademisi

   Sebagai referensi dalam pengembangan masalah-masalah deteksi abnormal.

1.5. Metodologi Penelitian

Dalam melakukan penelitian ini, langkah-langkah yang digunakan yaitu Perancangan Skenario & Penerapan Skenario, Metodologi yang digunakan penelitian ditunjukan pada **gambar 1.1.**

Gambar 1.1. Skema Analisis Serangan

- Perancangan Skenario

  Bertujuan untuk menyusun hal-hal yang dibutuhkan dalam penelitian yang bertujuan agar jalannya penelitian terarah untuk mencapai tujuannya.
- Penerapan Skenario

  Bertujuan untuk melaksanakan hal-hal yang telah direncanakan agar tujuan penelitian tercapai.

1.6. Batasan Peneltian

Agar tujuan dalam penelitian ini tercapai maka batasan yang ditentukan dalam penelitian ini adalah:

1. Penelitian dilakukan dilakukan pada jaringan LAN UKM Mandala Citra Media di Surakarta.
2. Serangan yang dianalisis merupakan serangan *Denial of Service* `TCP SYN Flooding` yang menyerang *file server* Mandala Citra Media Surakarta.
3. Analisis dilakukan secara *off-line* terhadap *log*.
4. Serangan yang melalui protokol FTP menggunakan *port* 21.
5. Berkas *log* digunakan sebagai petunjuk adanya serangan *denial of service*.
6. Tidak membahas kinerja algoritma *clustering*.
7. Tidak menentukan asal lokasi serangan.

1.7. Sistematika Penulisan

Bagian ini menjelaskan secara garis besar pokok bahasan pada tiap bab yang dikerjakan dalam penelitian ini, yaitu:

BAB I PENDAHULUAN

Menjelaskan latar belakang permasalahan, perumusan permasalahan, tujuan dari penelitian, manfaat atas penelitian, metodologi dalam penelitian, batasan bagi penelitian dan sistematika penulisan.

BAB II LANDASAN TEORI

Menjelaskan penelitian terkait *clustering* dan teori yang digunakan dalam mendukung penelitian.

BAB III METODOLOGI PENELITIAN

Menjelaskan langkah-langkah dalam melakukan penelitian, langkah-langkah dalam penelitian dibagi menjadi 2 (dua) bagian yaitu: **Perancangan Skenario** dan **Penerapan Skenario**.

BAB IV HASIL DAN PEMBAHASAN

Menjelaskan mengenai hasil dari penelitian yang telah dilakukan dalam menganalisis basisdata *log*.

BAB V PENUTUP

Menjelaskan kesimpulan dari hasil penelitian yang telah dilakukan yang sesuai dengan tujuan penelitian.