BAB I
PENDAHULUAN

1.1 LATAR BELAKANG

Dalam suatu trafik jaringan, biasanya terdapat keanehan/kejangaalan yang biasa disebut dengan Anomali Trafik. Anomali Trafik ini sering diartikan dengan penyimpangan atau keanehan yang terdapat pada suatu trafik jaringan. Anomali Trafik pada trafik jaringan meliputi *Denial of Service* (DoS), *Distributed Denial of Service* (DDoS) dan *flash crowds*. Serangan *Denial of Service* (DoS) adalah suatu jenis serangan terhadap sebuah komputer atau *server* didalam suatu jaringan internet dengan cara menghabiskan sumber (*Resources*) computer tersebut sehingga komputer tersebut tidak dapat menjalankan fungsinya dengan baik dan benar sehingga secara tidak langsung mencegah komputer lain mengakses komputer tersebut [1] [2]. Seiring perkembangan zaman, serangan *Denial Of Service* (DoS) sudah berkembang mennjadi serangan yang terdistribusi yang disebut *Distributed Denial of Service* (DDoS). *Distributed Denial of Service* (DDoS) merupakan salah satu serangan yang sangat ditakuti di dunia internet saat ini. Teknik serangan Distributed *Denial of Service* (DDoS) lebih canggih dibandingkan teknik yang digunakan oleh *Denial of Service* (DoS), yakni meningkatkan serangan beberapa kali dengan menggunakkan banyak komputer sekaligus, sehingga dapat mengakibatkan *server* yang diserang atau jaringan yang diserang menjadi tidak berfungsi sama sekali [2]. Diketahui dari sumber (RISK 2013) serangan DDoS berlangsung 8% dari total serangan di Internet. Memang sangat kecil presentasenya, yaitu hanya 8%, walaupun dengan presentase yang kecil, hal ini tidak bisa dianggap remeh, karena serangan ini dapat sangat merugikan target serangan. Target serangan ini berupa *server* dan *server* lainnya dalam suatu jaringan [1].

Sedangkan untuk *flash crowd*, *flash crowd* bukanlah merupakan suatu serangan seperti *Denial of Service* (DoS) ataupun *Distributed Denial of Service* (DDoS) melainkan berupa peningkatan trafik yang sangat tinggi secara signifikan dalam suatu jaringan sehingga tidak dapat diakses dalam waktu rentan tertentu [1]. Kejadian *Flash Crowds* ini dapat terjadi kapan saja, contohnya yaitu pada suatu kejadian seperti bencana alam, peluncuran produk. Breaking news, dll. Maka pada saat itulah terjadi peningkatan akses yang sangat tinggi ke suatu *server* [1]. *Flash Crowd* ini juga dapat dikatakan sebagai situasi dimana saat banyak *user* mengakses sebuah *website* dalam waktu yang sama.

Dalam mendeteksi adanya anomaly trafik, informasi 5-*tuple* dari *IP* menjadi karakteristik acuan dalam menganalisis. Dimana diantaranya, *protocol type, source IP address, destination IP address, source port, dan destination port* [1] [2]. Pada penelitian sebelumnya [4], dalam mendeteksi anomaly trafik menggunakan metode *Data Mining* dengan algoritma *K-means*, dikatakan berhasil dalam membedakan trafik normal dengan trafik anomaly. Namun, algoritma *K-Means* masih sangat sensitif terhadap adanya *outlier* dan *cluster* yang dihasilkan cenderung berbentuk oval [5]. Dari kelemahan algortima *K-Means* tersebut, maka digunakan algoritma ISODATA dalam mengatasi *outlier*. Proses pembelahan yang terdapat pada algoritma ISODATA dapat mengatasi hasil *cluster* yang cenderung berbentuk oval yang dihasilkan oleh algoritma *K-Means* [5]. Algoritma ISODATA ini merupakan perkembangan dari algoritma *K-Means* dimana terdapat proses – proses baru seperti penggabungan *cluster* dan pembelahan *cluster* serta kepadatan suatu *cluster* dapat dikontrol dengan algoritma [6] [7] [8] [9].

Dilihat dari masalah yang terdapat pada suatu trafik jaringan tersebut, dirasa penting membangun sistem yang dapat mendeteksi dan membedakan antara serangan DDoS dan anomali *flash crowds*. Maka dibuatlah sebuah sistem deteksi yang dapat membedakan antara serangan DDoS dengan *Flash Crowd* menggunakan algoritma ISODATA dan penambahan metode berbasis pengukuran jarak Manhattan Distance [4] [6] [7] [8] [9] [11]. Pada dataset yang digunakan sistem juga akan dimodifikasi dengan menggunakan metode *windowing* [15]. Dimana, pada penerapannya sistem dapat bekerja dengan baik dalam mendeteksi trafik normal dan trafik anomali.

1.2 RUMUSAN MASALAH

Dalam penelitian anomali trafik ini terdapat metode-metode yang mendukung dalam pendeteksian tersebut. Dimana salah satunya yaitu metode *Clustering*. *Clustering* merupakan salah satu teknik pengelompokan data berdasarkan kesamaan karakteristik data [3]. Kelebihan dari metode *clustering* ini mampu mengatur jumlah *cluster* fleksibel. Terdapat Algoritma untuk mengelompokkan suatu data dalam metode ini, salah satunya yaitu Algoritma ISODATA. ISODATA ini merupakan perkembangan dari Algoritma K-means. Algoritma ISODATA memiliki proses - proses seperti pengabungan *cluster*, pembagian *cluster* dan penghapusan *cluster* [6] [7] [11]. Dengan metode ini kepadatan suatu *cluster* dapat dikontrol dengan

algoritma. Dalam Algoritma ini juga terdapat parameter – parameter yang mendukung untuk melakukan prosesnya.

Dalam proses *clustering* dengan algoritma ISODATA, ada yang dinamakan *outlier* dalam dataset yang dipakai, *outlier* yaitu suatu objek yang berada jauh dari kumpulan objek [11]. Proses *clustering* akan membentuk *cluster* baru untuk objek *outlier*. Hal ini akan membuat *cluster* menjadi banyak dan variansi antar *cluster* semakin mengecil. Pada penelitian sebelumnya [5], algoritma *K-Means* yang digunakan pada penelitian tersebut dikatakan bahwa algoritma *K-Means* masih sangat sensitif terhadap adanya *outlier* dan cluster yang dihasilkan cenderung berbentuk oval.

Pada penelitian ini, sistem yang dibangun menggunakan algoritma ISODATA dan metode pengukuran jarak *Manhattan Distance* ke dalam sistem deteksi anomali trafik dan mengatasi objek *outliers* [11] [12]. Serta untuk mengetahui pengaruh dari metode pengukuran jarak *Manhattan Distance* pada algoritma ISODATA dari tingkat keakuratan algoritma berdasarkan parameter *Detection Rate* (DR), *Accuracy* (ACC), *False Positive Rate* (FPR) dan waktu proses pada algoritma ini dengan menggunakan *Manhattan Distance.*

1.3 TUJUAN

Tujuan dari penelitian ini yaitu merancang sistem deteksi anomali trafik menggunakan algoritma ISODATA *clustering*, dan menggunakkan metode pengukuran jarak *Manhattan Distance* pada Algoritma ISODATA kemudian dianalisis hasil perfomansi dari algoritma ISODATA berdasarkan parameter *Detection Rate* (DR), *Accuracy* (ACC) dan *False Positive Rate* (FPR). Kualitas sebuah *cluster* juga diukur dengan menggunakan metode *Dunn Index*. Selain itu, waktu proses dari algoritma ISODATA menggunakan metode pengukuran jarak *Manhattan Distance* juga akan diukur. Sistem juga akan dimodifikasi dengan menggunakan metode *windowing*.

1.4 BATASAN MASALAH

Dalam penelitian ini kita menggunakan metode *clustering* dengan Algoritma ISODATA dalam perancangan sistem deteksi anomali trafik dengan hanya sebatas mendeteksi dan tidak membahas mengenai pencegahan terhadap serangan yang ada pada suatu jaringan. Serta menggunakan metode berbasis jarak *Manhattan distance* pada Algoritma ISODATA dalam

proses pembelahan, penggabungan dan penghapusan *cluster* yang terdapat pada Algoritma ISODATA. Dalam mendeteksi suatu dataset serangan, sistem dimodifikasi dengan menggunakan metode *windowing* dimana sistem dapat membagi dataset agar memperkecil suatu kesalahan deteksi. Dalam penelitian ini juga digunakan dataset KDD Cup 1999 dan dataset Darpa 1998 real*time* yang sudah terekam (ter-capture) *berupa network log connection* untuk trafik normal dan serangan DDoS, untuk trafik yang mengalami *flash crowd* menggunakan dataset dari World Cup 1998. Dan juga dianalisis berdasarkan parameter *detection rate* (DR), *false positive rate* (FPR), dan *accuracy* (ACC).

1.5 METODELOGI PENYELESAIAN MASALAH

a. Studi Pustaka

Melakukan pencarian materi – materi dan refrensi melali buku, jurnal, internet, ataupun media lain yang berkaitan dengan permasalahan yang dibahas serta mendapatkan pemahaman mengenai topik yang dibahas.

b. Pengumpulan Data

Pengumpulan data trafik yang berhubungan dengan topik deteksi anomali. Data trafik dapat dikategorikan menjadi dua kategori yaitu data normal/*training* dan data observasi/pengujian. Data – data tersebut dapat berupa data real trafik yaitu data internet sebagai lingkungan implementasi sistem deteksi, data *real traffic* itu sendiri bisa berupa data *realtime* dan *non-realtime* [1]. Dalam penelitian ini kami menggunkan dataset pengujian diantaranya dataset Darpa 1998, KDDcup 1999, dan World Cup 1998.

c. Pre-Researh

Preprocessing adalah suatu proses untuk normalisasi [8] sebuah data trafik agar mudah/cocok untuk digunakan pada proses pendeteksian. Tujuan dalam proses *preprocessing* dataset ini yaitu untuk sebuah visualisasi dari distribusi fitur trafik dalam suatu trafik data pada data set yang seperti informasi *5-tuple*, besarnya paket, dan banyakny paket yang dikirim dll [1]. Prosesnya yaitu memisahkan *flow* pada data trafik yang telah ditentukan parameternya. Parameter pemisahannya yaitu dari IP tujuan yang diterima, jumlah paket dengan sumber IP yang sama, dan waktu sampling. Pada tugas akhir ini telah di lakukan *pre-*

research. Dari proses ini didapatkan ekstraksi fitur dan model trafik sesuai dengan anomali trafik (normal, *Flash Crowd*, DDoS).

d. Perancangan Sistem

Merancang sistem deteksi anomali trafik menggunakan algoritma *clustering* ISODATA.

e. Pengujian dan Analisis

Pengujian sistem deteksi anomli trafik menggunakan algoritma *clustering* ISODATA menggunakan dataset yang sudah dikumpulkan diantaranya dataset KDD Cup 1999, dataset Darpa 1998, dan dataset World Cup 1998, serta menganalisis hasil dasi sistem deteksi anomali dan analisis pengaruh dari penggunaan *Manhattan Distance* pada algoritma *clustering* ISODATA.

f. Penyusunan Laporan Tugas Akhir

Tahap ini, melakukan penyusunan laporan tugas akhir serta mengumpulan hasil dan dokumentasi yang diperlukan, formal laporan mengikuti kaidah penulisan yang benar dan sesui ketentuan – ketentuan yang ditetapkan oleh institusi.

1.6 SISTEMATIKA PENULISAN

BAB I PENDAHULUAN

Bab ini membahas latar belakang, masalah, tujuan, batasan masalah, dan metodelogi penyelesaian masalah.

BAB II KAJIAN PUSTAKA

Bab ini membahas prinsip kerja algoritma *clustering* ISODATA, prinsip kerja manhattan distance dan dunn index, serta istilah – istilah yang terkait dengan judul.

BAB III PERANCANGAN SISTEM

Bab ini menjelaskan proses perancangan sistem deteksi.

BAB IV PENGUJIAN DAN ANALISIS

Bab ini membahas pengujian sistem dengan dataset yang sudah dikumpulkan, serta membahas analisis keluaran dari sistem yang dirancang.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan akhir mengenai hasil perancangan dan analisis yang diperoleh serta saran dan harapan untuk pengembangan lebih lanjut.