PERANCANGAN DETEKSI ANOMALI *TRAFFIC* UNTUK INVESTIGASI LOG MENGGUNAKAN METODE K-MEANS *CLUSTERS*

Fadhilah Dhinur Aini[1*], Imam Riadi[1] dan Rusydi Umar[3]
[1]Program Studi Teknik Informatika, Universitas Ahmad Dahlan
Jl. Prof. Dr. Soepomo, Umbulharjo, Janturan Yogyakarta 55164.
[3]Program Studi Sistem Informasi, Universitas Ahmad Dahlan
Jl. Prof. Dr. Soepomo, Umbulharjo, Janturan Yogyakarta 55164.
[*]Email : fadhilahdhinuraini@gmail.com

Abstrak

Informasi teknologi saat ini, sangat penting bagi suatu institusi dapat diakses oleh para pengguna dari mana dan kapan saja. Kemunculan dari berbagai akses informasi dapat menimbulkan masalah baru ialah terjadi pemeliharaan validitas data tersebut. Deteksi merupakan sistem untuk mendeteksi aktifitas yan bersifat mengganggu akses informasi. Identifikasi yang dilakukan jika terjadi gangguan kepada administrator sehingga dapat melakukan suatu tindakan terhadap pengganggu tersebut. Sistem yang digunakan dalam penelitian ini yaitu Deteksi Anomali Traffic untuk Investigasi Log menggunakan metode K-means Clusters. Traffic Anomaly memberikan suatu pengguna jaringan internet membuat secara resmi tidak dapat melakukan akses. Karena hal tersebut dapat mengalami peningkatan jumlah akses jaringan bagi pengguna dan sewaktu-waktu dapat serangan dari akses jaringan yang tidak dikenali oleh pengguna sebelumnya. Sebab itu dibutuhkan adanya suatu sistem untuk mendeteksi dan mengenali setiap anomaly trafik tersebut. Dalam penelitian ini akan menggunakan sistem pendeteksian dengan metode K-means Cluster. K-means Cluster dapat membantu mendeteksi data log, kemudian mengelompokan kedalam cluster, sehingga akan membentuk sebuah pola yang akan digunakan untuk proses pendeteksian.

Kata kunci : *Anomaly Traffic, Forensik, K-means Cluster, Log File*

1. PENDAHULUAN

Sistem keamanan data saat ini masih kurang pengamanan, oleh karena itu dibutuhkan pengamanan dengan melakukan pendekatan dengan beberapa metode serta tools yang akan di gunakan. *Anomaly Traffic* salah satu pendeteksian yang dilakukan dengan melihat hasil *Detection rate* yang tertinggi dan terkecil dari pengujian data. *Anomaly Traffic* pada jaringan internet biasanya terjadi membuat pengguna dapat melakukan akses sebagaimana mestinya. Dalam *Anomaly Traffic* dapat mendeteksi peningkatan jumlah akses pengguna dan sewaktu – waktu akan terjadi sebuah serangan dari pihak lain terhadap jaringan tersebut. Anomaly mempunyai dua metode untuk melakukan deteksi yaitu *Signature* dan *Anomaly Traffic Based*. Pada signature menggunakan database pengenalan suatu anomali sehingga jika terdapat anomaly baru tidak akan terdeteksi. Sedangkan, *Anomaly* tidak menggunakan database tetapi menggunakan pembelajaran pola yang terjadi, sehingga jika terdapar *Anomaly* baru akan bisa terdeteksi (Zulfadhilah dkk, 2016).

Algoritma *Clustering* adalah *K-means*. *K-means* merupakan algoritma untuk menentukan sistem deteksi dan mengelompokkan data berdasarkan statik data dengan data lainnya untuk membentuk cluster, dengan statik yang tertinggi pada sebuah cluster dan statik rendah pada cluster yang berbeda (Purwanto dkk, 2014).

Log record atau file log merupakan sumber informasi yang penting bagi sistem forensik. log file mengandung banyak informasi tentang aktivitas pengguna atau administrator yang di rekam melalui sistem jaringan. file log akan menunjukkan aktivitas yang tela di lakukan oleh pengguna seperti siapa, jam, perangkat yang digunakan untuk melakukan serangan data.

Penelitian ini menggunakan metode *K-means Cluster* agar dapat memberikan informasi dari hasil pendeteksian atau menginvestigasi log dengan menentukan sebuah pola yang mudah untuk dipahami. Sehingga clustering merupakan salah satu metode dalam data mining yang digunakan untuk menyelesaikan masalah pendeteksian. Metode ini dapat menginvestigasi atau menganalisa data log, kemudian disatukan berdasarkan cluster-cluster. Maka akan terbentuk sebuah patern yang digunakan untuk menentukan mana yang sifatnya intrusive dan mana tidak. Tujuan penelitian ini, memberikan keamanan bagi administrator untuk memudahkan invenstigasi log dengan

mengunakan *Anomaly Traffic* melalui metode *K-means Clustering*. Hasil dari investigasi log melalui metode *K-means Clustering* dengan *Anomaly Traffic* berupa pola traffic data serangan yang di lakukan oleh penyerang menggunakan akses jaringan dari perangkat lain.

2. METODOLOGI

Dalam penelitian ini, melakukan simulasi pendeteksian serangan atau anomaly yang terjadi dalam sebuah jaringan, dikenal dua istilah pedeketan yaitu *Instrusion Detections System* (IDS) dan *Instrusion Prevention System* (IPS) (Purwanto, 2014). Pedekatan keduanya merupakan suatu sistem yang bekerja mengawasi keadaan jaringan dan memberikan pemberitahuan kepada sistem administrator apabila terdeteksi serangan. *Anomaly Traffic* dapat investigsi log dengan menggunakan metode *K-means Clustering* yang akan menghasilkan berupa data yang akan di bentuk pola dengan pendeteksian anomaly seperti statistik.

Pada penelitian simulasi ini menggunakan metode *K-means Clustering*. *K-means* merupakan penentuan jumlah cluster yang digunakan, jadi untuk proses penentuan cluster dilakukan oleh sistem data dan Threshold. Proses dari alur alogritma modified k-means menggunakan metode initilasasi cluster (Riadi, 2016). Metode algoritma K-means menentukan sebuah kelompok seperti dibawah ini:

1. Menentukan sebuah kelompok k-means dalam bentuk pola serta menentukan jumlah cluster k-means dengan dilakukan beberapa pertimbangan sebagai konseptual untuk menentukan berapa banyak *Cluster* tersebut.
2. Menghasilkan titik pusat cluster dengan memulai secara acak. Penentuan setroid awal dilakukan pada objek yang di gunakan sebagai *K-means Cluster*, lalu menghitung cluster dengan menggunakan rumus sebagai berikut:

$$v = \frac{\sum_{i=1}^{n} x_i}{n} \; ; \text{i=1,2,..,n} \qquad (1)$$

v : cluster centroid
n : hitungan objek yang dibentuk oleh cluster
xi: initialisasi objek

3. Mencari suatu objek pada sentroid cluster dengan menggunakan perhitungan kalkulasi.
4. Mendapatkan lokasi objek pada sentroid cluster dengan dua objek interaksi menggunakan metode K-means.
5. Interaksi antar dua objek yang dilakukan oleh sentroid baru menghasilkan data sentroid positif.
6. Dengan melakukan 3 kali percobaan sentroid sehingga menghasilkan data yang sama.

2.1. LITERATUR

Penelitian yang terkait yang pernah dilakukan oleh:

Penelitian ini dilakukan (Zulfadhilah, 2016). " *Cyber Profiling Using Log Analysis and K-Means Clustering*. Penelitian ini algoritma K-Means digunakan sebagai algoritma untuk proses pembuatan profil maya. Algoritma K-Means yang digunakan sesuai dengan harapan dari penelitian ini, karena memiliki proses algoritmik sederhana dengan tingkat akurasi yang baik. Tetapi algoritma *K-means* memiliki kelemahan, yaitu proses pembuatan nilai awal pusat *random* awal. Ini dapat menyebabkan perbedaan dalam hasil klaster. Hasil analisis *Log* dataset menggunakan algoritma *K-means* untuk proses *cyber* profiling menunjukkan bahwa algoritma tersebut memiliki aktivitas kelompok berdasarkan data pengguna internet mengunjungi situs *web*. Pengelompokan ini dibagi menjadi tiga, yaitu kunjungan rendah, sedang, dan tinggi.

Penelitian ini dilakukan (Wayan, 2015) "*Design and Analysis of Anomaly Detection Based Clustering Modified K-Means Algorithm with Timestamp Initalizations Sliding Window*". Penelitian ini dilakukan untuk menghasilkan modified k-means menggunkan Timestamp Initialization dapat digunakan sebagai algoritma data *traffic* menggunakan 9 feature dengan similaritas tinggi untuk sebuah *cluster*.

ISBN 978-602-99334-9-9

Penelitian ini dilakukan (Fahana J, 2017). "Pemanfaatan *Telegram* sebagai Notifikasi Serangan untuk Keperluan Forensik Jaringan". Penelitian ini menghasilkan perancangan sistem dan simulasi yang telah dibuat menunjukkan bahwa *IDS* bekerja dengan baik. *IDS* dapat mendeteksi serangan dengan memanfaatkan *Snort*.

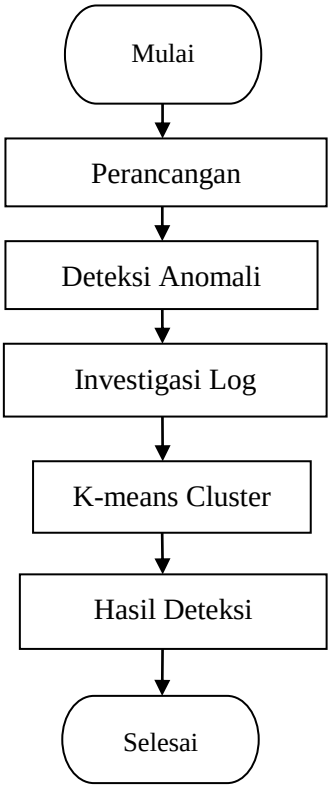
Penelitian ini dilakukan (Harjono, 2013). "*Honeyd untuk Mendeteksi Serangan Jaringan di Universitas Muhammadiyah Purwokerto*". Penelitian ini bertujuan untuk mendeteksi adanya serangan di dalam jaringan Universitas Muhammadiyah Purwokerto menggunakan *Honeyd*. Hasilnya bahwa serangan yang terdeteksi oleh honeyd berasal dari sejumlah host yang mempunyai alamat *IP private* dari jaringan internal UMP.

Penelitian ini dilakukan (Fadlil A, 2017). " *DDos Attacks Classification using Numeric Attribute-based Gaussian Naïve Bayes*. Penelitian ini menghasilkan klasifikasi *Gaussian Naive Bayes* adalah salah satu metode yang dapat digunakan untuk memproses atribut numerik pada layanan jaringan komputer. Numerik atribut seperti IP masuk dan paket panjang adalah utama fitur untuk mengetahui akses yang terjadi di jaringan komputer.

3. HASIL DAN PEMBAHASAN

Pada penelitian ini cara kerja sistem deteksi *Anomaly Traffic* untuk membantu investigasi log menggunakan metode *K-means* adalah menginvestigasi *Log* dengan *Anomaly Traffic* menggunakan metode *K-means Clustering* dapat mendeteksi terjadinya serangan dari pengguna yang melalui jaringan data forensik. data atau log ini akan menampilkan hasil berupa bentuk pola grafik dari *Anomaly Traffic* dan pola statistik *cluster*. *Anomaly Traffic* mendeteksi jalur sistem grafik yang ada pada sebuah jaringan forensik yang dapat mendeteksi serangan menggunakan perangkat komputer.

Adapun diagram alur tahap penelitian ini yang dilakukan dalam pemanfaatan anomaly traffic dan metode K-means cluster sebagai berikut Gambar 1:



Gambar 1. Diagram Alur Penelitian

Tahapan dalam diagram alur penelitian sebagai berikut :

1. Perancangan
   Perancangan merupakan suatu sistem yang bertujuan untuk menganalisa, menilai, serta memperbaiki dan menyusun suatu sistem. Perancangan termasuk dalam suatu metode teknik yang dilakukan dalam mengaplikasikan kerangka kerja.
2. Deteksi Anomaly Traffic
   Anomaly detections merupakan suatu monitoring untuk memantau pergerakan yang terjadi pada sistem jaringan. Jika terjadi penyerangan terhadap sistem maka anomaly traffic akan mendeteksi jumlah peningkatan pada jaringan tersebut.
3. Investigasi Log
   Investigasi log merupakan sistem data yang memiliki banyak informasi tentang aktivity administrator yang direkam melalui sistem jaringan.
4. K-means Cluster
   k-means cluster dapat memberikan hasil berupa pola statistik yang akan memudahkan investigasi log dalam menyelesaikan masalah pendeteksian.
5. Hasil Deteksi
   Hasil deteksi ini berupa gabungan antara *Anomaly Traffic* yang memantau investigasi yang terjadi pada log. Sehingga hasil tersebut dapat di gunakan dalam sebuah metode *K-means* dengan mencari proses statik tertinggi dan terendah sebagai bukti investigasi tersebut.

3.1. SIMULASI SISTEM

Sebuah rancangan simulasi untuk mendapatkan hasil serangan dengan menggunakan *Anomaly Traffic* melalui metode *K-means*. Pada penelitian ini dibuat dalam sebuah gambaran antara attacker dan korban menggunakan satu jaringan dalam Gambar 2:

Gambar 2. Rancangan Sistem

Pada gambar 2 ini menjelaskan tentang cara persiapan untuk melakukan pendeteksian yang akan dilakukan oleh user menggunakan sebuah perangkat keras berupa laptop yang terhubung melalui jaringan internet kemudian terhubung dengan *tool wireshark*, *packer tracer* untuk mendeteksi serangan.

4. KESIMPULAN

Deteksi *Anomaly Traffic* untuk investigasi *log* menggunakan metode *K-means Clustering* melakukan deteksi sistem yang dapat digunakan melalui sebuah *Traffic Anomaly* yang memberikan pola *Traffic* dari serangan dan berapa banyak pengguna yang melakukan serangan pada *log* dengan hasil investigasi menggunakan metode *Clustering* yang juga dapat menentukan pola data statistik tertinggi dan rendahnya suatu serangan.

DAFTAR PUSTAKA

Fadlil, A., Riadi, I., and Aji, S., (2017), DDos Attacks Classification using Numeric Attribute-based Gaussian Naive Bayes, *International Journal of Advanced Computer Science and Applications (IJACSA)*, Vol. 8, No. 8, pp.42-50.

Fahana, J., Rusydi, U., dan Faizin, R., (2017), Pemanfaatan Telegram Sebagai Notifikasi Serangan untuk Keperluan Forensik Jaringan., *Journal of Information Systems* hal 6-14.

Harjono., dan Wicaksono, A.P., (2013)., Honeyd untuk Mendeteksi Serangan Jaringan., Universitas Muhammadiyah Purwokerto.

Ardymulya, I., Riadi, I., (2016), *Danial Of Service Log Analysis Using Density K-Means Method, Journal Of Theoretical & Applied Information Technology*, Vol 83 Issue 2, ISSN: 1992-8645

ISBN 978-602-99334-9-9

M.W. Arif., Riadi, I., Sunardi., (2017), Deteksi Serangan Ddos Menggunakan Neural Network dengan Fungsi Fixed Moving Average Window, JISKA.

Purwanto, Y., Kuspriyanto,. Hendrawan,. Dan Raharjo, B., (2014), Traffik Anomaly Detection in Ddos Flooding Attack., *International Conference on Telecommunication System Service and Application*, vol 8, pp. 313-318.

Wayan., Purwanto, Y., Suratman, F. Y., Perancangan dan Analisis Deteksi Anomali Berbasis Clustering Menggunakan Algoritma Modified *K-Means* dengan *Timestamp* Initialization Pada Sliding Window., Telkom University, Bandung.

Zulfadhilah, M., Prayudi, Y., and Riadi, I., (2016), Cyber Profiling Using Log Analysis and K-Means Clustering A Case Study Higher Education In Indonesia, *Internasional Journal of Advanced Research in Computer Science and Applications*, 7(7), pp. 430-435.