Implementasi Splunk dalam Membangun *Security Information and Event Management* Berdasarkan *Log Firewall (studi kasus: Jaringan UII)*

Wahluf Abidian
Program Studi Informatika
Universitas Islam Indonesia
Yogyakarta, Indonesia
17523141@students.uii.ac.id

Mukhammad Andri Setiawan
Program Studi Informatika
Universitas Islam Indonesia
Yogyakarta, Indonesia
andri@uii.ac.id

***Abstrak*— Badan Sistem Informasi Universitas Islam Indonesia saat ini sedang dalam tahap menerapkan teknologi *enterprise*. Pada bidang keamanan, UII telah menerapkan *Next-Generation Firewall* (NGFW) dengan data yang dihasilkan memiliki ukuran yang masuk dalam kategori *big data.* Data yang besar tentunya menyimpan informasi yang besar pula. Sebagai bentuk upaya untuk meningkatkan pengamanan dan informasi pada jaringan UII. Penulis menerapkan teknologi SIEM yang digunakan untuk melakukan penggalian informasi data *Log* yang dihasilkan NGFW. *Tools* yang digunakan merupakan *next-generation* dari SIEM yaitu Splunk. *Log Firewall* dengan *traffic type* diolah menjadi visualisasi *cluster* dan *non-cluster* yang dikumpulkan pada sebuah *dashboard*. Diterapkan pula *alert system* yang diintegrasikan dengan *bot* Telegram. Tahapan dari penelitian meliputi studi literatur, perancangan, pengumpulan data, implementasi, pengujian, dan analisis. Penulis mendapatkan temuan bahwa visualisasi yang dibuat dapat memberikan kemudahaan pembacaan informasi maupun kemudahan mendapat informasi bagi administrator. Selain itu, potensi data yang didapat dari *Log Firewall* dapat menjadi acuan bagi pihak-pihak yang memiliki kepentingan dalam menentukan kebijakan.**

Kata Kunci—log; enterprise; firewall; SIEM; NGFW; cluster; dashboard; Alert;

1 PENDAHULUAN

Universitas Islam Indonesia adalah salah satu perguruan tinggi yang berkomitmen meningkatkan teknologi informasi dengan mendirikan Badan Sistem Informasi atau disingkat BSI. Secara umum BSI bertugas mengawal perencanaan, pengembangan, operasi, serta layanan sistem dan teknologi informasi untuk kurang lebih 30.000 pengguna [1]. Agar dapat memberikan layanan secara profesional, hendaknya BSI berfokus untuk menjamin *confidentiality* (kerahasiaan), *integrity* (konsistensi, akurasi, dan validitas data), *availability* (ketersediaan) atau biasa disebut CIA. Ketiga komponen tersebut merupakan dasar dari keamanan informasi [2].

Dalam menjaga kerahasiaan data, BSI telah mengimplementasikan *next-generation firewall* (NGFW) dari perusahaan *Palo Alto Network*. Teknologi *next-generation firewall* (NGFW) merupakan penerus dari *firewall* tradisional yang memiliki banyak kekurangan. Fitur pada NGFW mencakup segala fitur yang terdapat pada *firewall* tradisional dan terdapat fitur tambahan seperti *Intrusion Prevention System* (IPS), *Deep Packet Inspection* (DPI), *Application Control*, *Directory Integration*, dan *Encrypted Traffic Inspection* [3]. Dengan jumlah pengguna yang kurang lebih mencapai 30.000 pengguna, tentu saja catatan log pada *firewall* memiliki jumlah data berjumlah besar.

Untuk membantu administrator BSI UII dalam membaca *Log Firewall* dengan jumlah yang besar, solusi yang dapat ditawarkan adalah membuat *Log Monitoring System*. *Security Information Event and Management* atau disingkat SIEM merupakan sistem informasi terpusat dan digunakan untuk mengumpulkan *log* yang nantinya memberikan hasil berupa visualisasi *Log Monitoring* guna mempermudah pembacaan informasi *Log* [4]. Namun, akan timbul masalah dalam visualisasi *Log Monitoring* apabila jumlah data log berukuran besar. Informasi yang ditampilkan akan memiliki persebaran acak sehingga cenderung sulit untuk dibaca maupun dipahami.

Solusi untuk menangani jumlah data yang besar atau saat ini biasa disebut *big data* salah satunya adalah dengan pendekatan *data mining*. Data yang besar nantinya akan dikelompokkan atau dikluster dengan hitungan secara matematika dan statistik sebelum data ditampilkan. Oleh karena itu, visualisasi *Log Monitoring* akan dapat lebih mudah dipahami [5]. Dalam menerapkan *data mining* salah satu teknik yang digunakan adalah menggunakan *machine learning*. Dalam melakukan pengelompokan atau *cluster* dibutuhkan algoritma sebagai dasar dari penerapan teknik *machine learning*. Pemilihan algoritma salah satunya harus didasarkan dengan *tools* yang digunakan dalam mengimplementasikan teknologi *Security Information and Event Management* (SIEM) [6]. Pada penelitian kali ini alat yang akan digunakan sebagai solusi yang penulis tawarkan adalah menggunakan aplikasi Splunk yang merupakan *next-generation* SIEM *product*. Splunk mendukung penerapan *data mining* dengan teknik *machine learning* yang dibutuhkan untuk mengelola *Log Firewall* dalam melakukan klasterisasi agar visualisasi *Log Firewall* mudah untuk dipahami. Splunk juga memiliki fitur penyaringan, pencarian, masukan, modifikasi, pelaporan, dan penghapusan data.

Pada Splunk terdapat *add-on* berupa aplikasi, salah satunya adalah *Splunk Machine Learning Toolkit* yang memiliki fitur untuk melakukan kegiatan *machine learning*, yaitu *clustering*. Pada bagian *clustering* yang terdapat pada *add-on* Splunk, terdapat dua variasi untuk melakukan eksperimen *clustering* yaitu *smart clustering* dan *cluster number events*. Pada variasi *clustering* tersebut, Splunk memberikan 4 pilihan algortima *machine learning* yaitu DBSCAN, Birch, K-means, dan SpectralClustering [7]. Setelah melakukan perbandingan dari keempat algoritma tersebut, pada penelitian kali ini algoritma yang akan digunakan adalah K-means. K-means merupakan algoritma *machine learning* dengan type *Unsupervised Learning*. Algoritma K-means dipilih karena dalam hal melakukan komputasi, lebih cepat dibanding dengan algoritma klustering lainnya, hasil algoritma ini juga sederhana untuk menjelaskan dan memahami, dan algoritma ini biasanya menghasilkan *cluster* yang lebih *tighter* dibanding *hierarchical clustering* [7].

Splunk memiliki fitur *alert real-time* yang dapat digunakan untuk membantu memantau *event* atau serangan yang terjadi. Menyiapkan *alert* berdasarkan *critical control* keamanan dapat memberi informasi keamanan ketika penyerang mencoba melewati kontrol atau ketika perangkat yang berpotensi tidak aman atau tidak sah memasuki jaringan [8]. Dalam menerapkan *alert*, sebaiknya pemberitahuan *alert* dapat diketahui tidak hanya saat sedang memantau Splunk saja. Sebab administrator tidak mungkin selalu berada di tempat yang sama. Oleh karena itu dibutuhkan integrasi *alert* dengan sebuah aplikasi yang dapat memberikan kemudahan penerimaan informasi saat *alert* muncul. Telegram merupakan sebuah aplikasi yang dapat diakses pada *smartphone* ataupun perangkat komputer. Telegram memiliki fitur *bot* yang dapat diintegrasikan dengan Splunk untuk membantu kemudahan menerima informasi saat *alert* muncul. Penggunaan *bot* Telegram juga telah diadopsi oleh *team* BSI UII untuk melakukan pemantauan suhu *server*. Oleh karena itu, penerapan *bot* Telegram sudah sangat familiar di lingkungan BSI UII.

2 LANDASAN TEORI

2.1 Landasan Teori

Pemanfaatan IT salah satunya sebagai tempat mengumpulkan sumber informasi serta berbagai data penting yang merupakan aset dari organisasi. Informasi ataupun aset yang terkumpul merupakan sebuah nilai (*Value*) yang berharga dari suatu organisasi yang harus dilindungi. Semakin besar sebuah organisasi tentunya akan semakin besar dalam memanfaatkan infrastruktur IT, hal ini akan membuat sistem IT semakin kompleks dan terdistribusi sehingga akan membuat departemen IT mengalami kesulitan dalam melakukan pengelolaan serta kegiatan *monitoring*. Penerapan teknologi mutakhir tentunya diperlukan untuk mempermudah departemen IT dalam melakukan pekerjaannya.

2.1.1 Keamanan Informasi

Keamanan Informasi merupakan bidang serta aktivitas profesional multidisiplin yang berkaitan dengan pengembangan dan juga implementasi mengenai mekanisme keamanan dari berbagai aspek secara teknis maupun organisasional. Tindakan minimal yang harus diterapkan oleh organisasi untuk mengamankan informasi saat ini biasa dikenal dengan Kerahasiaan, Integritas, dan Ketersediaan atau biasa disebut dengan CIA. Kerahasiaan bertugas memastikan privasi dari data dengan membatasi akses melalui enkripsi yang terotentikasi. Integritas bertugas menjamin bahwa informasi tepat dan akurat serta terpercaya. Ketersediaan bertugas memastikan informasi selalu dapat diakses oleh pihak atau pengakses yang berwenang [9].

2.1.1.1 Confidentiality (Kerahasiaan)

Kerahasiaan memiliki istilah lain yaitu privasi. Sebuah perusahaan hendaknya memiliki kebijakan untuk membatasi akses ke informasi atau data yang hanya dapat diakses oleh pihak atau staf yang berwenang. Salah satu solusi dari hal tersebut adalah membagi data menurut tingkat keamanan atau sensitivitas dari informasi tersebut. Terdapat metode untuk memastikan kerahasiaan yang mencakup enkripsi data, ID nama pengguna dan kata sandi, otentikasi dua faktor, dan meminimalisir dalam menyebarkan informasi yang sensitif.

2.1.1.2 Integrity (Integritas)

Integritas merupakan keakuratan, konsistensi, dan keandalan dari data selama masa pakainya. Data yang dikirimkan tidak boleh berubah dan juga tidak dapat diubah oleh seseorang atau entitas yang tidak berwenang ataupun tidak sah. Perizinan pada file serta kontrol untuk akses pengguna dapat mencegah dari pengakses yang tidak sah. Selain itu, kontrol versi juga dapat digunakan untuk mencegah perubahan dari ketidaksengajaan pengguna yang berwenang. Cadangan data diperlukan untuk membantu apabila data ingin dikembalikan dari risiko kerusakan seluruh data, dan *hashing checksum* dapat digunakan untuk memverifikasi integritas data selama data dalam pengiriman.

2.1.1.3 Availability (Ketersediaan)

Memelihara peralatan, memperbaiki perangkat keras, selalu melakukan *update* sistem operasi dan perangkat lunak, serta mencadangkan data untuk memastikan ketersediaan jaringan dan data bagi pengguna yang berwenang. Rencana yang telah disusun hendaknya dijalankan degan cepat misalnya dalam hal memulihkan dari bencana alami atau akibat perbuatan manusia. Peralatan (*tools*) atau perangkat lunak (*software*) keamanan, seperti *Firewall* akan berguna untuk melindungi dari gangguan akibat serangan seperti DoS (penolakan layanan). DoS terjadi apabila penyerang mencoba membuat sumber daya tidak berdaya dalam memberikan respon yang mengakibatkan layanan tidak tersedia bagi pengguna.

2.1.2 IT Security Risk Management

Risk Management atau manajemen risiko merupakan proses untuk melakukan identifikasi celah keamanan (*vulnerabilities*) serta ancaman yang mengancam sumber informasi yang biasanya digunakan oleh organisasi dalam upaya mencapai tujuan bisnisnya dan untuk menentukan respon/tindakan

pencegahan apa saja yang diperlukan apabila risiko tersebut terjadi atau bahkan dapat mengurangi risiko dengan catatan tingkatnya dapat diterima didasarkan pada nilai dari sumber informasi organisasi tersebut [5].

IT risk management berupaya untuk mencoba melindugi *Confidentiality*, *Integrity*, dan *Availability* (CIA) dengan meminimalisir dampak yang nantinya mungkin akan muncul dan memberikan efek pada *Confidentiality* dari informasi, *Integrity* dari data yang terdapat pada sistem, dan *Availability* yang berasal dari infrastruktur sistem [10].

2.1.3 Security Information and Event Management (SIEM)

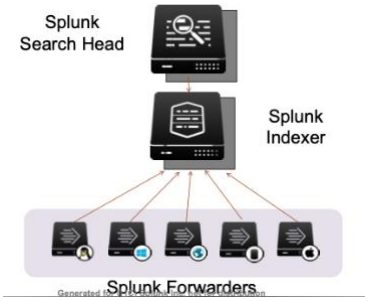
Security Information and Event Management atau biasa disebut dengan SIEM merupakan sebuah teknologi yang dapat mendeteksi berbagai ancaman dan insiden dari keamanan dengan mengumpulkan *Log real-time* dan melakukan analisis sejarah *Log* keamanan dari berbagai jenis tipe log dan berasal dari berbagai sumber data dari perangkat yang berbeda-beda.

SIEM mendukung analisis dan juga investigasi pada insiden melalui analisis dari data historis dan berbagai alat keamananan dari organisasi seperti Router, IDS/IPS, UTM, Firewall, Server, dan lain-lain. Inti dari kemampuan teknologi SIEM yaitu cakupan yang luas dalam mengumpulkan *Log* dan juga kemampuan dalam melakukan korelasi dan melakukan analisis kejadian dari berbagai sumber jenis log atau perangkat yang berbeda-beda [11].

2.1.4 Splunk

Splunk adalah *platform* perangkat lunak untuk melakukan pencarian, menganalisis, dan memvisualisasikan data yang dihasilkan mesin dan biasanya berasal dari *computers, network device, logs, sensor, databases*, dan lain-lain yang menyusun infrastruktur IT suatu organisasi.

Splunk bekerja melalui 3 perangkat penyusun dasarnya seperti yang ditunjukkan pada Gambar 1. Pertama adalah setiap perangkat yang meghasilkan data dipasang alat bernama Splunk *Forwarders* yang berfungsi untuk menghimpun data dari perangkat dan mengirimnya menuju Splunk *Indexer* yang merupakan tempat berkumpulnya data dari berbagai perangkat. Untuk kebutuhan tingkat atas yaitu berkaitan dengan klien dan visualisasi maka alat yang bekerja adalah Splunk *Search Head*.



Gambar 1. Cara Kerja Splunk

2.1.5 Next-Generation Firewall (NGFW)

Next-Generation Firewall atau biasa disebut NGFW lebih kuat dibandingkan dengan Traditional *Firewall*. NGFW memiliki kemampuan Traditional *Firewall* dan juga memiliki sejumlah fitur tambahan untuk menangani lebih banyak variasi kebutuhan organisasi dan memblokir lebih banyak potensi ancaman. *Firewall* ini disebut dengan "*Next-Generation*" untuk membedakan dari *Firewall* lama (Traditional *Firewall*) yang tidak memiliki kemampuan tambahan tersebut [12]. Segala macam *Firewall* mencatat segala macam jenis kegiatan ke dalam *Log Firewall*. Karena terdapat fitur tambahan pada NGFW, maka ukuran dari *Log Firewall* cenderung lebih besar dibanding Traditional *Firewall*.

2.1.6 Data Mining

Data mining merupakan proses dalam menggunakan teknik metematik, statistik, kecerdasan buatan, dan *machine learning* dalam melakukan ekstraksi dan melakukan identifikasi yang bermanfaat serta pengetahuan yang terkait dari berbagai *big data* atau maha data [11]. Tujuan dari dilakukannya *data mining* adalah untuk memahami lebih jauh terkait perilaku data yang sedang diamati (deskripsi) dan sebagai acuan untuk memperkirakan kondisi yang nantinya akan terjadi (prediksi).

2.2 Penelitian Terkait

[5] Dalam penelitiannya melakukan konsolidasi dan visualisasi *Log Server* yang berasal dari Badan Sistem Informasi, Universitas Islam Indonesia. Hasil dari penelitian ini adalah *Log* dapat dikonsolidasi dan terkumpul menjadi satu dengan bantuan Filebeat, *Log* dapat dipecah untuk disesuaikan dengan kebutuhan administrator menggunakan Logstash, dan *Log* yang telah dipecah berhasil divisualisasikan dengan memanfaatkan aplikasi Kibana. Namun peneliti memberikan saran untuk penelitian berikutnya agar menambahkan variasi dari *Log* yang dikelola dan pemanfaatan fitur *machine learning* yang tersedia pada pada aplikasi Kibana.

[13] Dalam penelitiannya dalam melakukan klasifikasi *alert* pada *Intrusion Detection System* menggunakan algoritma K-means. Penulis menggunakan *tools* IDS yaitu Snort dalam proses identifikasi serangan dan menggunakan *tools* Matlab untuk proses *clustering* dengan algoritma K-means. Hasilnya peneliti mendapatkan hasil nilai klasifikasi serangan menggunakan algoritma tersebut, namun penulis mencatat beberapa kelemahan dari algoritma K-means dalam penelitiannya yaitu kurangnya kejelasan mengenai bagaimana menentukan jumlah kluster (k) yang terbaik.

[14] Dalam penelitiannya melakukan integrasi untuk visualisasi *Log* dari *honeypot* dengan sistem manajemen *honeypot* paling mutakhir yaitu *Modern Honey Network* (MHN) dengan memanfaatkan *tools* Splunk. Peneliti mendapatkan hasil bahwa MHN mudah diterapkan karena telah memiliki API untuk membaca data penyerang. Hasil yang diperoleh pun mempermudah keterbacaan aktivitas yang telah tercatat pada *Log*.

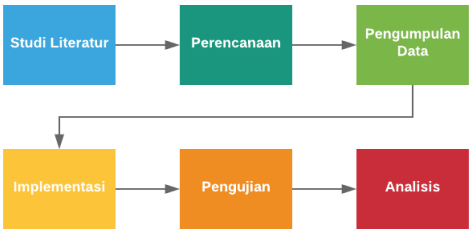
[15] Dalam kajian literaturnya yang membahas mengenai *Security Information and Event Management* (SIEM) dengan pendekatan Data Mining. Menunjukkan bahwa berbagai teknik

data mining dapat diterapkan pada SIEM untuk meningkatkan kemampuan dari sistem.

Dari penelitian [5], [13], [14], [15] penulis mengadopsi teknologi *Security Information and Event Management* (SIEM) yang digunakan untuk melakukan manajemen *Log*. Aplikasi Splunk penulis gunakan sebagai alat untuk menerapkan teknologi SIEM. Pendekatan Data Mining dengan teknik *machine learning* penulis adopsi untuk meningkatkan keterbacaan informasi dari *Log Firewall.* Perbedaan yang sangat menonjol dari penelitian terkait sebelumnya adalah pada bagian data yang diolah, tujuan pengolahan data, serta pendekatan yang penulis gunakan yang merupakan kombinasi dari penelitian terkait serta saran dari penelitian terkait.

3 METODOLOGI

Penelitian ini menggunakan metode penelitian seperti yang ditunjukkan pada Gambar 2:



Gambar 2. Metode Penelitian

3.1 Studi Literatur

Tahapan pertama dalam penelitian ini adalah studi literatur untuk pencarian masalah, pencarian referensi, dan landasan teori sebagai dasar untuk melakukan penelitian. Studi literatur dilakukan guna melakukan *review* terhadap makalah, jurnal ataupun hasil penelitian yang sejenis, mengeksplorasi sumber pustaka yang dikutip oleh tulisan yang dibaca dengan bahasan yang sejenis. Penulis mengadopsi beberapa teknologi yang digunakan dan mengambil beberapa saran penelitian dari studi literatur sebagai masukan dalam meneliti.

3.2 Perencanaan

Setelah melakukan studi literatur, penulis melakukan perencanaan secara mendalam untuk solusi ditawarkan. Perencanaan meliputi sumber atau literatur yang penulis adopsi, pemilihan teknologi yang akan digunakan, dan pemilihan *tools* yang akan digunakan.

3.3 Pengumpulan Data

Data yang akan digunakan merupakan data yang tercatat pada *Log Firewall* Badan Sistem Informasi, Universitas Islam Indonesia data dalam waktu satu hari secara penuh yang merupakan hasil pilihan dari hari-hari lainnya di Universitas Islam Indonesia. Data *Log Firewall* yang digunakan adalah type *Traffic* dan merupakan data sekunder.

3.4 Implementasi

Pada bagian implementasi akan terdiri dari beberapa metode sebagai berikut,

3.4.1 Instalasi Splunk Web

Dalam penelitian kali ini, penulis menggunakan versi *trial* dari Splunk dan merupakan Splunk Web, sebab penggunaan aplikasi Splunk untuk kebutuhan jangka panjang adalah berbayar. Adapun kebutuhan system untuk dapat menggunakan versi *Trial* dari Splunk Enterprise ditunjukkan pada Tabel 1.

Tabel 1. KEBUTUHAN SISTEM

Sistem Operasi	Windows / Linux / Mac Os
Processor	Minimal 2-core 64-bit CPU at 2GHz
RAM	Minimal 4 *gigabyte*
Web Browser	Versi paling *update* dari Chrome / Firefox / Safari
Port Splunk Web	8000

3.4.2 Instalasi Add-on Aplikasi Palo Alto Firewall

Data yang akan diteliti merupakan data yang berasal dari *Firewall Palo Alto Network*. Oleh karena itu, untuk menambah keterbacaan *fields* dari data *Log Firewall* (*traffic type*) sehingga fields yang terdeteksi akan lebih banyak. Maka dibutuhkan instalasi *Add-on* Aplikasi *Palo Alto Network*.

3.4.3 Instalasi Aplikasi Splunk Machine Learning Toolkit

Untuk proses klasterisasi dengan pendekatan *machine learning*, dibutuhkan aplikasi yang dapat melakukan proses komputasi dengan *machine learning*. pada Splunk telah terdapat aplikasi berupa *Add-on* untuk pengolahan tersebut yang bernama "*Splunk Machine Learning Toolkit".*

3.4.4 Instalasi Aplikasi Telegram Alert Action

Pada *alert system* yang akan diterapkan dan akan diintegrasikan dengan *bot* Telegram. Maka instalasi aplikasi "*Telegram Alert Action*" bertujuan untuk memudahkan dalam melakukan integrasi tersebut.

3.4.5 Upload Data

Splunk Web versi *trial* memberikan pembatasan untuk upload data yaitu sebesar 500 *megabyte* untuk satu waktu upload. Data *traffic Log Firewall* dalam satu hari memiliki ukuran 25 – 45 *gigabyte*. Maka salah satu solusi agar data dapat diolah adalah dengan membagi data menjadi bagian-bagian yang berukuran dibawah 500 *megabyte*. Penulis melakukan pemecahan data dengan memanfaatkan *tools* dari linux yaitu *tools* "*Split*". *Syntax* untuk memecah file ditunjukkan pada Gambar 3.

```
Split -b <ukuran file>m –additional-suffix=<ekstensi
file> <Nama File> <nama file setelah dipecah> --verbose
```

Gambar 3. *Syntax Split* File

3.4.6 *Search & Reporting*

Pada bagian ini, hal yang dilakukan adalah membuat "*Rule*" untuk menampilkan visualisasi dari data *Log Firewall* agar mudah dipahami atau dibaca. Untuk melakukan hal ini, yang harus dilakukan adalah masuk pada bagian "*Search & Reporting*" dari tampilan awal Splunk Web. Berikut merupakan *Rule* yang akan digunakan:

```
index=* sourcetype="pan:traffic"
| chart count over app:technology by action
```

Gambar 4. *Rule Detection Traffic by Action*

```
index=* sourcetype="pan:traffic"
| stats dc(user) as total_pengguna
```

Gambar 5. *Rule Detection All User by User*

```
index=* sourcetype="pan:traffic"
| stats count by client_location
| sort -count
```

Gambar 6. *Rule Detection Most Popular Country by client_location*

```
index=* sourcetype="pan:traffic"
| iplocation client_ip
| geostats count by client_location
```

Gambar 7. *Rule Detection User Location by client_location*

```
index=* sourcetype="pan:traffic"
| chart count over app:technology by action
```

Gambar 8. *Rule Splunk Web Login Attemps*

```
index=* sourcetype="pan:traffic" user=* user!=""
| stats count(eval(action="success")) as successes
count(eval(action="failure")) as failures by user,
app
| where successes>0 AND failures>100
```

Gambar 9. *Rule Detection Brute Force Attack*

3.4.7 *Klasterisasi*

Pada bagian ini, hal yang dilakukan adalah mencoba melakukan klasterisasi dari data "*Log Firewall*". Kegiatan ini dilakukan dengan menggunakan aplikasi *Splunk Machine Learning Toolkit* yang sebelumnya telah di-*install*. Algoritma untuk *preprocessing* ialah *Standard Scaler* dan Algoritma *Clustering* adalah *K-means*.

3.4.8 *Membuat Alert dan Integrasi ke bot Telegram*

Hasil yang diperoleh dan sesuai dengan kebutuhan administrator akan disimpan menjadi sebuah *alert.* Saat *alert* akan disimpan, terdapat *Rule* yang harus ditentukan sebagai pemicu kapan sebuah informasi akan dikirim menjadi sebuah *alert.* Pembuatan *bot* pada aplikasi Telegram memanfaatkan *FatherBot* dari Telegram. Kemudian konfigurasi untuk integrasi dengan Splunk adalah dengan memasukkan "Token ID" dan "Chat ID"

3.5 *Pengujian*

Kegiatan utama dari penelitian ini adalah menghasilkan tiga macam keluaran, ketiga keluaran tersebut yaitu pertama visualisasi *non-cluster*. Kedua visualisasi dengan *cluster*. Ketiga integrasi *alert* dengan *bot* Telegram. Pengujian dilakukan dengan menguji *Rule* pada data *Log Firewall* pada hari yang berbeda.

3.6 *Analisis*

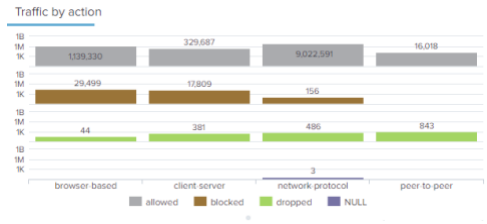
Tahapan analisis akan dilaksanakan setelah pengujian dilakukan. Analisis dilakukan untuk mengamati hasil dari visualisasi *non-cluster* dan visualisasi dengan *cluster*. Hasil dari analisis ini yang nantinya akan menjadi acuan bagi penulis untuk menuliskan kesimpulan.

4 PEMAPARAN HASIL

4.1 *Visualisasi Log Firewal non-Cluster*

Dari *Rule* yang telah diterapkan, maka hasil visualisasi adalah sebagai berikut:

Rule pada Gambar 4, dengan memilih jenis visualisasi "*Column Chart"*, memberikan hasil visualisasi seperti pada Gambar 10.



Gambar 10. *Visualization Traffic by Action*

Rule pada Gambar 5, dengan memilih jenis visualisasi "*Single Value"*, memberikan hasil visualisasi seperti pada Gambar 11.

Gambar 11. *Visualization All User by user*

Rule pada Gambar 6, dengan memilih jenis visualisasi "*Word cloud",* memberikan hasil visualisasi seperti pada Gambar 12.

Gambar 12. *Visualization Most Popular Country by client_location*

Rule pada Gambar 7, dengan memilih jenis visualisasi "*Cluster Map",* memberikan hasil visualisasi seperti pada Gambar 13.

Gambar 13. *Visualization User Location by client_location*

Rule pada Gambar 9, dengan memilih jenis visualisasi "*Statistics Table"* menghasilkan *No Results Found*

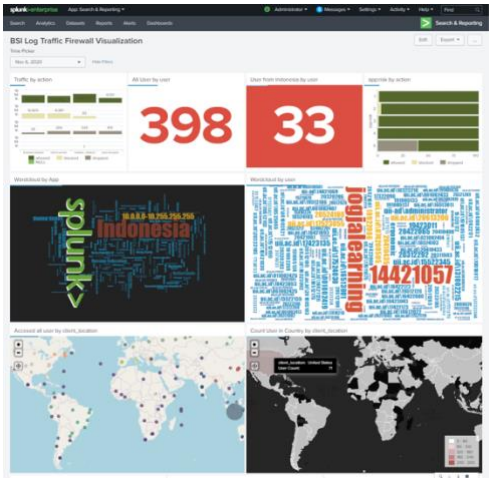
Dalam proses pengujian, Penulis menguji apakah Rule bersifat *universal* sehingga dapat untuk melakukan deteksi di setiap hari. Karena keterbatasan waktu dan ruang penyimpanan. Sebab data *Log Firewall* setiap harinya berukuran lebih dari 30 *gigabyte*. Penulis melakukan pengujian pada dua hari yang berbeda. Pemilihan hari ditentukan berdasarkan ukuran *Log Firewall* yang lebih besar dibanding hari-hari lainnya. Hari pertama adalah Rabu, 28 Oktober 2020. Kemudian hari kedua adalah Jum'at, 06 November 2020. Tabel 2 dan Tabel 3 menunjukkan perbandingan dari kedua hari tersebut:

Tabel 2. DASHBOARD RABU, 28 OKTOBER 2020

Waktu	Rabu, 28 Oktober 2020
Data Size	40.8 *gigabyte*
Total Events	104.484.528 events
Problems	Tidak ada

Tabel 3. DASHBOARD JUM'AT 06 NOVEMBER 2020

Waktu	Jum'at, 06 November 2020
Data Size	43.3 *gigabyte*
Total Events	111.866.547 events
Problems	Tidak ada



Dari hasil Tabel 2 dan Tabel 3, tidak terdapat masalah sehingga dapat disimpulkan bahwa *Rule* yang telah ditetapkan bersifat *universal* karena dapat diintegrasikan pada hari yang berbeda. Menurut penulis, hal ini terjadi karena proses yang dilakukan hanya sekedar melakukan pencarian data yang telah ada kemudian ditampilkan dalam bentuk visualisasi, tidak ada proses perhitungan matematika ataupun statistika.

4.2 Visualisasi Log Firewall dengan Cluster

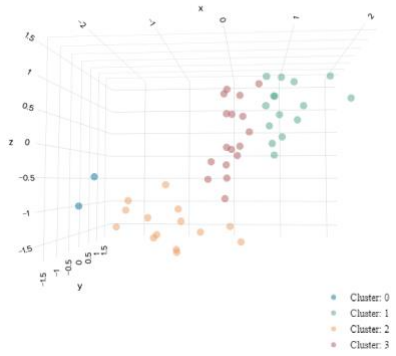
Penulis membuat klasterisasi untuk visualisasi dengan *cluster* dari *Log Firewall* berdasarkan *field* "*app:category*". *Field* tersebut dipilih dengan harapan penulis akan mendapatkan kecenderungan *users* dalam menggunakan aplikasi. Untuk menambah variasi informasi, dalam visualisasi akan dikombinasikan dengan *field "_time"*. Pada *field* "*app:category*" hanya dipilih lima kategori dari enam kategori aplikasi. Lima kategori tersebut dipilih untuk mempermudah dalam mengetahui kecenderungan *user* dalam menggunakan aplikasi. Dalam penerapan *clustering*, fitur yang digunakan yaitu *Smart Cluster*, *Rule* yang digunakan ditunjukkan pada Gambar 14:

```
index=* sourcetype="pan:traffic"
| timechart count by app:category
| fit StandardScaler "business-systems"
"collaboration" "general-internet" "media"
"networking" with_mean=true with_std=true
| fit KMeans "SS_business-systems"
"SS_collaboration" "SS_general-internet" "SS_media"
"SS_networking" k=4
| eval clusterId= "Cluster: " + cluster,
x='SS_general-internet', y='SS_collaboration',
z='SS_business-systems'
```

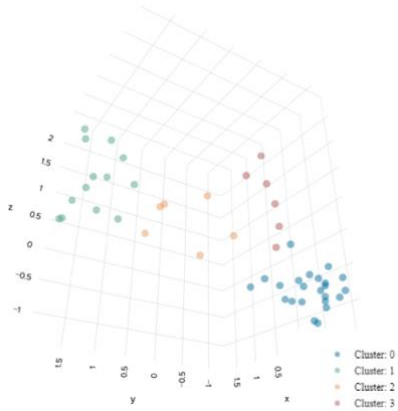
Gambar 14. *Rule* Splunk Clustering with *K-Means Algorithm*

Rule pada Gambar 14 menghasilkan visualisasi yang ditunjukkan oleh Tabel 4 dan Tabel 5.

Tabel 4. CLUSTER PADA RABU, 28 OKTOBER 2020

Waktu	Rabu, 28 Oktober 2020
Algorithm Preprocessing	*Standard Scaler with respect to Mean and Standart Deviation*
Algorithm Clustering	*K-Means with K = 4*
Problems	Tidak ada
	

Tabel 5. CLUSTER PADA JUM'AT, 06 NOVEMBER 2020

Waktu	Jum'at, 06 November 2020
Algorithm Preprocessing	*Standard Scaler with respect to Mean and Standart Deviation*
Algorithm Clustering	*K-Means with K = 4*
Problems	Tidak ada
	

Visualisasi pada Tabel 4 dan Tabel 5, memberikan gambaran persebaran *cluster*. Masing-masing warna mewakili jenis dari kluster sebagaimana ditunjukkan pada *legend* dari gambar tersebut. Untuk nilai dari masing-masing anggota *cluster* dapat dilihat ketika kursor diarahkan pada salah satu *cluster*. Namun visualisasi jenis tersebut tidak dapat menggambarkan bagaimana kecenderungan pengguna dalam melakukan akses berdasarkan waktu. Jenis visualisasi yang cocok digunakan untuk melihat detail waktu adalah "Downsampled Line Chart".

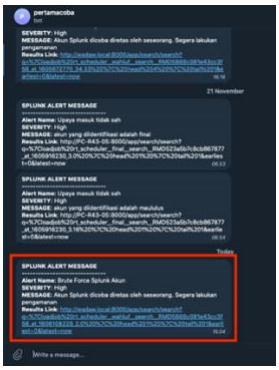
Dari hasil Tabel 4 dan Tabel 5 tidak terdapat masalah sehingga dapat disimpulkan bahwa *Rule* pada Gambar 14 bersifat *universal* karena dapat digunakan pada data dangan hari yang berbeda. Menurut penulis hal ini dapat terjadi karena variabel dari *field* dipilih dengan tepat. Variabel yang dipilih selalu muncul pada setiap data *Log Firewall* pada hari apa saja.

4.3 Integrasi Alert dengan bot Telegram

Penulis membuat tiga Jenis *alert.* Bagian *alert* yang pertama adalah mendeteksi upaya masuk yang tidak sah pada akun *Splunk Web.* Bagian *alert* yang kedua adalah memberitahu informasi jumlah pengguna apabila melebihi batas dari pengguna seharusnya dalam kurun waktu yang telah ditentukan. Bagian *alert* yang ketiga adalah mendeteksi serangan *Brute Force.* Gambar 8 menunjukkan *Rule* yang digunakan pada bagian pertama dan Gambar 15 merupakan konfigurasi *Rule* yang memicu kapan *alert* akan dikirimkan.

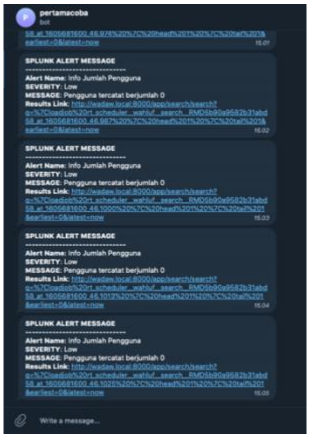
Gambar 15. *Rule* pemicu agar *Alert terkirim*

Dalam melakukan pengujian, penulis mencoba melakukan kesalahan selama empat kali dalam kurun waktu satu menit saat melakukan *log-in*. Gambar 16 menunjukkan hasil dari pengujian tersebut.



Gambar 16. Notifikasi pada *bot* Telegram

Pengujian *alert* pada bagian kedua, hasilnya tidak muncul. Hal ini terjadi karena data yang akan dibaca harus merupakan data yang sedang berlangsung (*present*) sedangkan data *Log Firewall* yang penulis gunakan merupakan data yang telah lalu (*past)*. Untuk mamastikan apakah integrasi dengan *bot* telah berhasil, maka penulis merubah pengaturan pada bagian *Trigger Alert When* menjadi "*Per-result*", Gambar 17 menunjukkan hasil setelah perubahan pada *Rule* pemicu.



Gambar 17. Notifikasi Jumlah Pengguna

alert memberikan notifikasi ke *bot* Telegram, namun user yang terdeteksi adalah nol. Hal tersebut terjadi karena deteksi dilakukan secara *Real-Time*, maka tidak ada data yang dibaca sehingga data *user* berjumlah nol. Untuk pengujian pada bagian *alert* yang ketiga adalah sama seperti kasus pada bagian kedua, mendapatkan hasil nol.

5 KESIMPULAN

Berdasarkan hasil yang telah dipaparkan, penulis berhasil melakukan penelitian sesuai dengan tujuan awal yaitu membuat *visualisasi non-cluster*, *visualisasi* dengan *cluster*, dan membuat *alert* yang terintegrasi dengan *bot* Telegram. Penulis mendapati bahwa hasil dari visualisasi memberikan kemudahan bagi administrator dalam mengerti dan mendapat apa saja yang terjadi pada *traffic* jaringan UII. Bahkan dapat melakukan identifikasi data *Log Firewall* yang telah berlalu untuk keperluan investigasi. Penulis dapat menarik kesimpulan bahwa penerapan teknologi *Security Information and Event Management* dengan *tools* Splunk pada level organisasi yang telah menerapkan sistem *enterprise* sangatlah dibutuhkan. Selain dapat memudahkan administrator dalam pembacaan dan kemudahan didalam mendapatkan informasi pada *traffic* jaringan UII yang terjadi secara *real-time*. Penerapan teknologi SIEM akan dapat memberi banyak masukan bagi pihak-pihak yang memiliki kepentingan dalam menentukan kebijakan karena tren penggunaan perangkat pintar semakin meningkat. Informasi-informasi seperti kebiasaan pengguna aplikasi atau informasi penting lainnya akan dapat diperoleh dari *Log Firewall* jaringan UII. Penulis memberikan saran agar penelitian selanjutnya lebih berfokus pada penggalian informasi memanfaatkan *machine learning.* Selain itu, pada bagian *bot* Telegram dapat dilakukan penambahan fitur untuk melakukan *request* informasi agar administrator tidak hanya menerima informasi apabila terdapat *alert* saja.

6 REFERENCES

[1] A. BSI, "Sekilas Tentang BSI," 2017. https://bsi.uii.ac.id/sekilas-bsi/ (accessed Sep. 10, 2020).

[2] H. Azaim and W. Nuryanto, "Mengenal Confidentiality, Integrity, dan Availability Pada Keamanan Informasi," *Netsec.Id*, 2017. https://netsec.id/confidentiality-integrity-availability-keamanan-informasi/ (accessed Nov. 10, 2020).

[3] Cloudflare, "What Is a Next-Generation Firewall (NGFW)? | NGFW vs. FWaaS," *CloudFlare*, 2020. https://www.cloudflare.com/learning/cloud/what-is-a-next-generation-firewall/ (accessed Nov. 10, 2020).

[4] S. Perciballi, "Design Correlation Rules to Get the Most Out of Your SIEM," *Paloalto Network*, 2015. https://blog.paloaltonetworks.com/2015/08/design-correlation-rules-to-get-the-most-out-of-your-siem/ (accessed Sep. 10, 2020).

[5] Y. B. Erwinsyah, "Konsolidasi dan Visualisasi Log Server BSI UII Menggunakan Elk Stack," 2019.

[6] J. Chancey and J. Penn, "Analyticsdriven Security Power Next Generation SIEMS," Accenture, Dublin, 2018.

[7] Splunk, "Splunk® Machine Learning Toolkit User Guide 5.2.0," *Splunk Inc*, 2020. https://docs.splunk.com/Documentation/MLApp/5.2.0/User/WelcometoMLTK (accessed Sep. 10, 2020).

[8] D. Brown, "Finding Bad with Splunk," *SANS.edu*, 2016.

[9] N. Cisco, "Pendahuluan Tentang Keamanan Cyber," 2020. http://static-course-assets.s3.amazonaws.com/CyberSec2.1/id/index.html#1.2.1.2 (accessed Nov. 12, 2020).

[10] A. Haikal, "Pentingnya IT Risk Management Dalam Mendukung Keberlangsungan Bisnis," 2017. https://netsec.id/pentingnya-it-risk-management/ (accessed Nov. 13, 2020).

[11] T. Efrain, J. E. Aronson, and T. P. Lian, *Decision Support Systems and Intelligent Systems Edisi Bahasa Indonesia Jilid 1*. Yogyakarta: Andi, 2005.

[12] Cloudflare, "What Is a Next-Generation Firewall (NGFW)? | NGFW vs. FWaaS," 2020. https://www.cloudflare.com/learning/cloud/what-is-a-next-generation-firewall/ (accessed Nov. 13, 2020).

[13] F. C. Wulur and I. Sembiring, "Klasifikasi Alert pada Intrusion Detection System Menggunakan Algoritma K-Means," no. December, pp. 2–4, 2015.

[14] Rakhmadhani, Syaifuddin, and Z. Sari, "Integrasi Visualisasi Modern Honey Network (MHN) dengan Splunk," *SENTRA 2019*, pp. 167–172, 2019.

[15] A. R. Zope, A. Vidhate, and N. Harale, "Data Mining Approach in Security Information and Event Management," *Int. J. Futur. Comput. Commun.*, vol. 2, no. 2, pp. 80–84, 2013, doi: 10.7763/ijfcc.2013.v2.126.