DETEKSI INTRUSI PADA JARINGAN KOMPUTER BERDASARKAN ANALISA PAYLOAD MENGGUNAKAN METODE OUTLIER

Ahmad Fajar al Kharis

***Abstract*—Sebagian besar sistem deteksi intrusi biasanya hanya mengulas tentang header dari sebuah packet, sementara bagian payloadnya tidak diperhatikan.Untuk mencegah serangan yang tidak diketahui dari internet maka diperlukan suatu IDS yang mampu menganalisa bagian header dan juga bagian payloadnya. Pada proyek akhir ini akan didesain dan diimplementasikan suatu sistem deteksi penyusupan jaringan yang memiliki kemampuan untuk mendeteksi adanya serangan dengan menganalisa payload dari suatu packet dengan menggunakan metode outlier. Yaitu dengan melakukan normalisasi pada fitur fitur yang dipilih, kemudian menghitung standard deviasi untuk menentukan batas bawah dan batas atas. Data yang berada diluar range batas bawah dan atas akan dianggap sebagai outlier. Selain itu juga digunakan keyword payload untuk menentukan apakah payload tersebut termasuk serangan atau bukan. Diharapkan metode outlier ini mampu mengenali serangan melalui analisa fitur pada packet dengan akurat. Dan menjadi salah satu metode yang digunakan pada software IDS diantara metode-metode lain yang sudah dibuat atau sedang dikembangkan.**

***Index Terms*—IDS, payload, outlier, tcpdump**

I. Introduction

Dalam era teknologi informasi saat ini, hampir seluruh informasi yang penting bagi sebuah institusi dapat diakses oleh para penggunanya. Keterbukaan akses tersebut memunculkan berbagai masalah baru, antara lain : pemeliharaan validitas dan integritas data/informasi tersebut, jaminan ketersediaan informasi bagi pengguna yang berhak, pencegahan akses informasi dari yang tidak berhak, dan pencegahan akses sitem dari yang tidak berhak.

Sistem pertahanan pada jaringan komputer terhadap aktivitas gangguan umumnya dilakukan oleh seorang admin, dengan bantuan software IDS (Intrusion Detection System). Keberadaan IDS sangat membantu kerja soarang admin jaringan, karena aliran data pada sebuah jaringan sangat banyak dan prosesnya berlangsung 24jam.

Saat ini berbagai jenis IDS telah dikembangkan baik bersifat open source maupun yang komersial, namun secara garis besar IDS terbagi dalam 2 kategori dalam mengenali pola serangan, yaitu yang berbasis signature dan yang berbasis anomaly. IDS yang berbasis signature menggunakan database yang berisi data dan ciri dari sebuah serangan. Metodenya adalah dengan mencocokan setiap data yang lewat dengan data yang ada pada database, sehingga jika ada kesamaan ciri dengan data yang ada pada database, maka data yang lewat tersebut dianggap sebagai sebuah serangan. Keuntungan dengan metode ini adalah kemampuan IDS yang mampu mengenali secara tepat data-data pada jaringan yang berpotensi menimbulkan kerusakan. Namun IDS ini akan sangat lemah apabila data serangan yang lewat belum dikenali, atau belum ada database-nya. Oleh karena itu IDS tipe ini membutuhkan update database secara kontinyu sesuai dengan munculnya tipe atau pola serangan yang baru.

Sedangkan IDS yang berbasis pada anomaly bersifat lebih fleksibel, karena dapat mengenali pola serangan baru tanpa harus meng-update database pola serangan. IDS yang berbasis pada anomaly memiliki sebuah kecerdasan buatan yang mampu mendeteksi dan mengenali sebuah serangan. Anomaly pada dasarnya adalah mencari data yang menyimpang dari sekumpulan data normal. IDS yang berbasis anomaly menggabungkan metode analisis dan statistik untuk mengenali penyimpangan tersebut. Kelemahan dari metode ini adalah kemungkinan salah identifikasi pada data yang diolah.

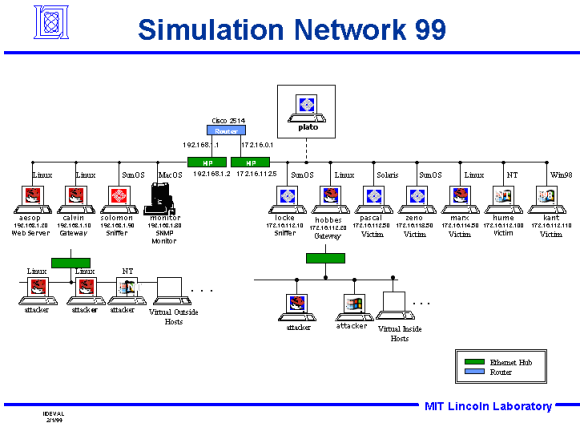
II. Methodology

Terdapat dua proses utama yang dilakukan dalam penelitian ini, proses yang pertama adalah proses capture packet pada jaringan dan menuliskan data yang telah di-capture tersebut kedalam sebuah file. Proses capture disini menggunakan software tcpdump yang mampu mengcapture dan menuliskan data yang akan dianalisa pada proses selanjutnya. Proses yang kedua adalah proses analisa dari data yang sudah disimpan pada file. Sebelum melakukan analisa, program akan melakukan ekstraksi fitur, karena data yang dismpan dalam file masih berbentuk binary dan tidak terbaca. Seperti pada penelitian Like Zhang, analysis of payload based application level network anomaly detection,Proses ekstraksi hanya dilakukan pada packet TCP dan terfokus pada bagian payload, maka kami hanya meng-ekstrak fitur berikut ini : Packet Length, Source Port, Payload Size, dan Payload. Setelah itu dilakukan proses normalisasi pada fitur yang telah diekstrak,sehingga menghasilkan resultan dari data tersebut. Nilai resultan ini akan digunakan untuk menghitung nilai threshold sekaligus untuk menentukan apakah data yang lewat termasuk serangan atau bukan. Khusus untuk fitur payload dilakukan proses pemberian keyword terlebih dahulu. Biasanya kata pertama dari sebuah payload selalu didahului dengan ”keyword parameter” seperti ”GET/index.html” atau “EHLO Jupiter.cherry.org”. Kata pertama inilah yang akan diambil dan diinisialisasi, misalnya kata “EHLO” diberi nilai 79, sedangkan kata ”GET” diberi nilai 32. Kata pertama tersebut sebenarnya adalah sebuah protokol dan merupakan protokol umum yang sudah diketahui. Pada umumnya sebuah payload yang mengandung serangan biasanya berisi kata yang aneh atau asing dan tidak umum seperti ”.tcshrc”. Kata yang tidak dikenali akan diberi nilai tersendiri. Pada bagian akhir program akan menampilkan output berupa

data yang mengandung anomaly dan juga menampilkan data-data yang menyertainya seperti tanggal dan waktu kejadian, alamat IP tujuan dan asal, port tujuan dan port asal, dan juga payloadnya itu sendiri. Tipe payload yang termasuk serangan bisa dicatat untuk update database serangan yang telah dikenali. Hal ini dapat membantu admin dalam mengambil keputusan selanjutnya dan juga membantu menjaga keamanan jaringan computer.

III. Results

Pada bab ini program akan diuji dengan menggunakan beberapa data set, data set yang digunakan adalah data set second week training DARPA'99. Data set DARPA'99 merupakan data packet yang dicapture pada jaringan selama 22 jam oleh MIT Lincolin Laboratory. Terdiri dari dua buah data, yaitu : inside.tcpdump dan outside.tcpdump. inside.tcpdump, merupakan data packet yang dikumpulkan dari jaringan lokal, sedangkan outside.tcpdump berisi data packet yang dikumpulkan dari luar jaringan local. Di dalam kedua data ini terdapat data serangan yang sudah teridentifikasi. Sehingga kemampuan program dalam menganalisa dan mengenali serangan dapat diuji. Berikut adalah skema dari simulasi yang dilakukan oleh MIT Lincolin Laboratory untuk mendapatkan data set DARPA'99 selama 4 hari.



Pada pengujian kali ini digunakan data set DARPA'99 yang dicapture dari jaringan pada hari senin, 8 maret 1999. Didalam data set ini terdapat 7 serangan yang telah teridentifikasi. Berikut adalah data serangan yang telah diketahui :

Date	Start time	Source IP	Name
03/08/1999	05:14:25	172.016.112.100	NTinfoscan
03/08/1999	06:00:23	172.016.114.050	pod
03/08/1999	09:31:43	172.016.112.100	back
03/08/1999	12:09:18	172.016.112.50	httptunnel
03/08/1999	15:57:15	172.016.112.50	land
03/08/1999	04:03:52	192.168.001.020	secret
03/08/1999	04:33:52	192.168.001.010	ps attack

Setelah data inside.tcpdump dan outside.tcpdump di masukan ke dalam program, maka hasilnya adalah sebagai berikut :

```
root@ubuntu:/home/alkhareelst# g++ TA.cpp -o TA
root@ubuntu:/home/alkhareelst# ./TA inside.tcpdump outside.tcpdump

File: inside.tcpdump
            Anomaly No.1
Waktu Kejadian          :       03/08/1999 05:14:25
Source IP               :       172.016.112.100
Destination IP          :       172.016.114.050
Source Port             :       000.000.000.020
Destination Port        :       000.000.008.145
Payload Size            :       12692
Payload keyword         :       ^@/*^M^  * Na

            Anomaly No.2
Waktu Kejadian          :       03/08/1999 06:00:23
Source IP               :       172.016.114.050
Destination IP          :       172.016.113.204
Source Port             :       000.000.009.020
Destination Port        :       000.000.000.079
Payload Size            :       5
Payload keyword         :       ^@Al^M^

            Anomaly No.3
Waktu Kejadian          :       03/08/1999 09:31:43
Source IP               :       172.016.112.100
Destination IP          :       172.016.114.148
Source Port             :       000.000.000.020
Destination Port        :       000.000.108.178
Payload Size            :       12692
Payload keyword         :       ^@/*^M^  * Na

End of file

212 TCP streams still open


File: outside.tcpdump
            Anomaly No.4
Waktu Kejadian          :       03/08/1999 04:33:52
Source IP               :       192.168.001.020
Destination IP          :       192.168.001.010
Source Port             :       000.000.005.145
Destination Port        :       000.000.000.053
Payload Size            :       49
Payload keyword         :       ^@^@.W^A^@^@^A^@

            Anomaly No.5
Waktu Kejadian          :       03/08/1999 04:33:52
Source IP               :       192.168.001.010
Destination IP          :       192.168.001.020
Source Port             :       000.000.000.053
Destination Port        :       000.000.005.145
Payload Size            :       497
Payload keyword         :       ^@^A^@^@^A^@

End of file

175 TCP streams still open

root@ubuntu:/home/alkhareelst#
```

Dari hasil uji coba diatas, program mampu mengenali 5 dari 7 serangan yang ada. Pengujian akan terus dilakukan dengan menggunakan data set yang berbeda, hal ini bertujuan untuk menambah keyword payload, sehingga akurasi program dalam mengenali payload yang berisi serangan semakin bertambah. Berikut adalah tabel hasil pengujian pertama.

IV. Conclusions

Setelah melakukan beberapa percobaan maka didapatkan kesimpulan sebagai berikut : Program mampu mengenali adanya serangan dengan baik. Program mampu mengenali serangan baik serangan dengan tipe lama, maupun serangan dengan tipe yang baru. Program mampu mengeksploitasi fitur yang ada pada header maupun payload dari suatu packet. Terdapat delay yang cukup lama saat program melakukan analisa, hal ini bergantung pada jumlah data yang dianalisa. Kecepatan dari analisa juga bergantung pada hardware yang digunakan (processor dan RAM)

V. REFERENCES

1) John E Dickerson, Julie A Dickerson. Fuzzy Network Profiling for Intrusion Detection. Electrical and Computer Engineering department Iowa State University Ames.
2) Matthew V. Mahoney, Philip K. Chan. PHAD: Packet Header Anomaly Detection for Identifying Hostile Network Traffic. Department of Computer SciencesFlorida Institute of Technology. Florida Institute of Technology Technical Report CS-2001-04
3) Eleazar Eskin. Anomaly Detection over Noisy Data using Learned Probability Distributions. Computer Science Department, Columbia University.
4) James R. Binkley , Suresh Singh. An Algorithm for Anomaly-based Botnet Detection. Computer Science Dept Portland State University Portland OR USA
5) Lilis fauizah.Pendeteksian serangan jaringan computer berbasis IDS snort dengan algoritma clustering K-means.Jurusan teknik informatika politeknik elektronika negri surabaya-ITS.
6) Penyiapan data(preprocessing).Modul ajar EEPIS-ITS.
7) Ali ridho barakbah. Modul ajar Cluster analysis.EEPIS-ITS
8) O'Reilly. Learn Network step by step. Network troubleshooting tool 2004. http://hell.org.ua/Docs/oreilly/tcpip2/tshoot/ch04_04.htm.
9) WinPCap manual ebook. http://www.winpcap.org/
10) Windump manual ebook. http://www.winpcap.org/windump/default.htm
11) http://en.wikipedia.org/wiki/Intrusion_detection_system
12) http://en.wikipedia.org/wiki/IPv4
13) http://en.wikipedia.org/wiki/Anomaly
14) 2000 DARPA intrusion detection evaluation data set. Windows NT attack data set.http://www.ll.mit.edu/mission/communications/ist/corpora/ideval/data/2000data.html
15) KDD-cup 1999. Knowledge Discovery and Data Mining 1999 .http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html
16) 1999 DARPA intrusion detection evaluation data set.http://www.ll.mit.edu/mission/communications/ist/corpora/ideval/data/1999/training/week2/index.html