

Case Study of an Anomalous Traffic Detection on the Aggregation Points of Enterprise Network

Yoohee Cho, Yihan Kim

KT Network R&D Laboratory

463-1 Jeonmin-dong, Yuseong-gu, Daejeon, South Korea

[yoohee,yihan}@kt.com](mailto:{yoohee,yihan}@kt.com), Fax: +82 42 870 8491, Tel: +82 42 870 8139

Abstract— Our network infrastructure is exposed to persistent threats of DDoS and many unknown attacks. These threats threaten the availability of ISP's network and services. This paper proposes network-based anomalous traffic detection method and presents an anomalous traffic detection system, its architecture and main function blocks. Every five minutes, traffic information and security events are gathered in a central server and new arrival traffic is compared with the already generated baseline traffic. This approach is exploring an anomalous traffic detection based on time series traffic modeling and analyzing traffic by time of day, day of week, and special days. To improve the accuracy of detection, we analyze flow information and security events. We developed an anomalous traffic detection system and deployed on the aggregation points of enterprise network.

Keywords— intrusion detection, anomaly, traffic monitoring

I. INTRODUCTION

Guaranteeing of seamless and secure traffic through the Internet is becoming a critical issue because the daily life of corporation and individual customers has been more heavily depending on the Internet in these days. This paper proposes network-based anomalous traffic detection method and presents an anomalous traffic detection system, its architecture and main function blocks

Anomalous traffic detection is one of the intrusion detection techniques. Traditionally, intrusion detection techniques can be categorised into signature detection and anomaly detection [1][2]. Signature detection uses patterns of well-known attacks to match and identify known intrusions. They search for malicious code in every network packets. If signature is well defined, it is easy to detect a malware, but if unknown or novel attacks appear, it takes much time to understand a signature and apply it. Now a day detection method tends to be further complicated since attackers are likely to evolve intelligent techniques of payload encoders and encrypted data channels. Anomaly detection relies on the concept of a baseline for ordinary network behavior. This method can determine anomalous network behavior by comparing predefined and acceptable

model; baseline traffic. Baseline traffic is automatically self-learned or specified by network administrator, or both. It is difficult to define rules and needed for detailed knowledge of normal network behavior. If the rules have been built once, engine can scale more flexible than signature-based detection. But anomaly detection has high false positive rate, it needs to deliberate a learning period to adjust baseline traffic parameters.

This approach is exploring an anomalous traffic detection based on time series traffic modeling and analyzing traffic by time of day, day of week, and special days. In order to complement the high false positive rate of anomaly detection, we also reference security events data from IDSs that are located on the aggregation points of enterprise network.

II. MOTIVATION AND CONTRIBUTION

ISPs have invested on security solutions and services for their customers but, investment budget is increasingly needed to protect the persistent threats such as worm, virus, DDoS (Distributed Denial of Service) and even unknown attacks. This paper is to offer an anomalous detection system that includes aspects of both an anomalous traffic monitoring and signature-based detection techniques. While we are collecting SNMP data, flow statistics, security events from the real network and observing the ingress/egress traffic of enterprise customer networks, it quite a difficult task to determine symptoms of anomalous traffic that allows sending alarms to operators. Therefore, we need to make a baseline traffic based on time of day, day of week and special days with a sufficient learning period. Every five minutes, traffic information and security events are gathered in a central server and new arrival traffic is compared with the already generated baseline traffic. To improve the accuracy of determination, we analyze flow information and security events to understand traffic characteristics at that moment. In this paper, we presented an Anomaly Traffic Detection System (ATDS) which using SNMP traffic volume as baseline traffic and referring to flow statistics and security events at a time. This ATDS has been installed in the enterprise network and indicated an anomalous traffic symptom to users and system operators as a service.

The rest of this paper is organized as follows. In section 3, we present the network environments and general traffic view of enterprise customer network. In section 4, the overall anomaly detection system architecture and the function blocks of it are given. We show a commercial system implemented using this approach in section 5 and finally, we conclude in section 6.

III. NETWORK ENVIRONMENTS

Figure 1 shows access network as an experimental network used throughout this paper, where enterprise customers' access network, not individuals'. Most of them have fixed C class IP blocks(/24). In Figure 1, we collect SNMP data from aggregation switches and security events from IDSs located transparently in line. There is a central server farm in the network and all data were gathered in it. Each access router is enabled to export the flow and is set the sampling rate at 1:500.

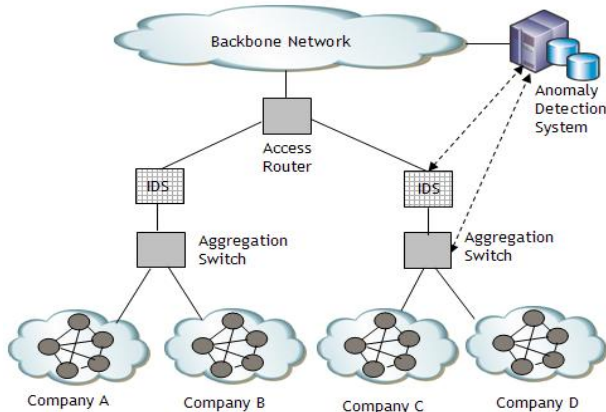


Figure 1. Network environment

Figure 2 and Figure 3 shows the aggregated BPS(Bits per Second) and FPS(Flows per Second) of the ingress and egress traffic on the aggregation points of enterprise network, where a unit X axis indicates five minutes duration and it begins on Sunday midnight(00:00) to Saturday 11:55 p.m.

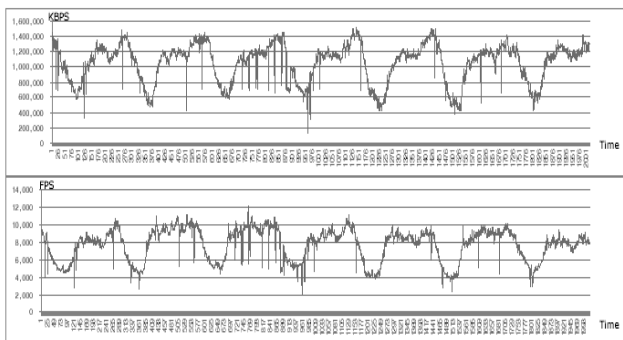


Figure 2. General traffic view of campus access network

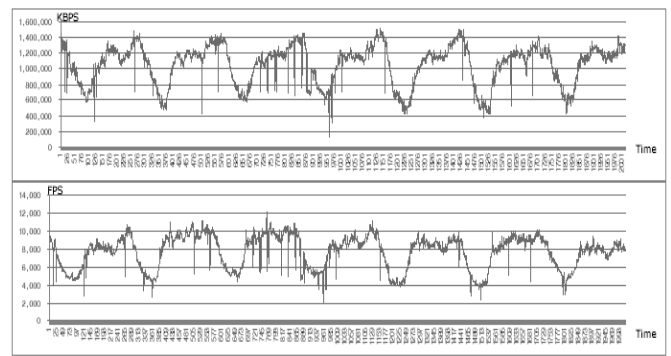


Figure 3. General traffic view of enterprise access network

From the figures, traffic data was collected during one week and was aggregated five campus access networks and four SMB (Small to Medium) enterprise networks. From Figure 2 and 3, we note that there are time-based and day-based repeated patterns of the volume and flows. The traffic pattern on different access routers might be obviously diverse according to the connecting stub networks. Therefore, we try to set up baseline traffic for each customer network based on a time series.

IV. ANOMALY DETECTION SYSTEM ARCHITECTURE

This section supposes an overall architecture of Anomalous Traffic Detection System (ATDS). It is necessary to define each process from collecting packets of the network to detecting anomalous traffic step by step. From Figure 4, Packet Collecting is gathering SNMP, flow and security events from the network and auditing data on memory of system to process real timely. In Aggregating Packets step, raw data arriving in real-time are aggregated every five minutes interval into one record.

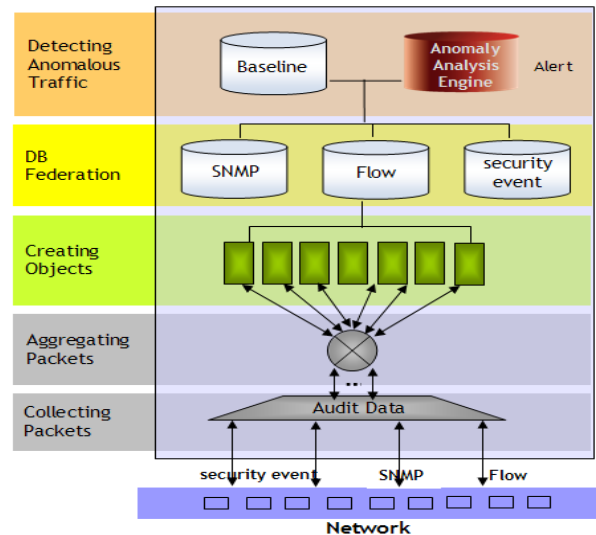


Figure 4. Overall architecture of ATDS

Creating Objects make meaningful data scheme based on predefined analysing templates. In this step, we can extract new important meaning from joining several objects. Whenever

objects are created, they are loaded onto an associated database table and there are three types of database; SNMP, flow and security events. Because we measure traffic data from the real network, data volume is enormous that we keep data separated in each database on set purpose. In final step, baseline traffic is made based on normal network usage of customer traffic during a specific learning time and Anomaly Analysis Engine does discriminate an anomalous traffic comparing baseline traffic, flow and security events also are referred to determine an anomaly symptom. If anomaly symptom is occurred, ADTS system dispatches an alert to administrator and users.

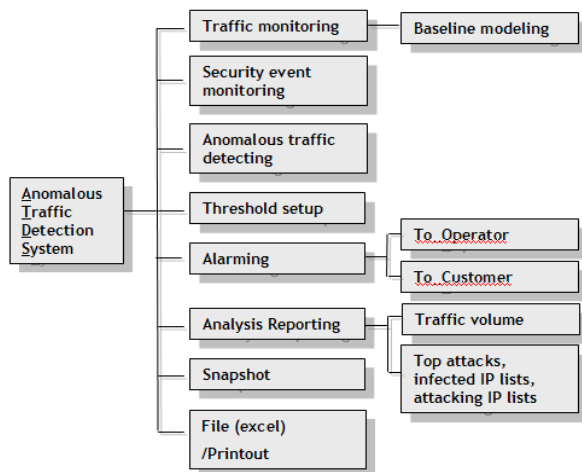


Figure 5. Function block of ADTS

Figure 5 shows the function blocks of ADTS. Traffic monitoring is observing customer's traffic volume (BPS) and flow every five minutes. As a result of traffic monitoring, we make baseline traffic for each customer.

Last week(w-4):	X _{01:00}	X _{01:05}	X _{01:10}	X _{01:15}	X _{01:20}	X _{01:25}	X _{01:30}	X _{01:35}	X _{01:40}	X _{01:45}	X _{01:50}	X _{01:55}
Last week(w-3):	X _{01:00}	X _{01:05}	X _{01:10}	X _{01:15}	X _{01:20}	X _{01:25}	X _{01:30}	X _{01:35}	X _{01:40}	X _{01:45}	X _{01:50}	X _{01:55}
Last week(w-2):	X _{01:00}	X _{01:05}	X _{01:10}	X _{01:15}	X _{01:20}	X _{01:25}	X _{01:30}	X _{01:35}	X _{01:40}	X _{01:45}	X _{01:50}	X _{01:55}
Last week(w-1):	X _{01:00}	X _{01:05}	X _{01:10}	X _{01:15}	X _{01:20}	X _{01:25}	X _{01:30}	X _{01:35}	X _{01:40}	X _{01:45}	X _{01:50}	X _{01:55}
This week (w)	X _{01:00}	X _{01:05}	X _{01:10}									

Figure 6. Four-weeks traffic data

Figure 6 shows four weeks traffic data to make baseline traffic. We note $X_{HH:MM}$ as the bits per seconds(bps) of some hour(HH) and some minute(MM) while minutes express five minute intervals. $T_{baseline} = \sum(X_{HH:MM})/N$ (N is the number of X), baseline traffic is calculating average of last four weeks BPS for each enterprise customer traffic. Whenever a week is passing by, moving last four weeks average is applied. From Figure 6, $X_{01:10}$ of this week is compared with average of last four weeks traffic data in the same column of the table. When we evaluate the baseline traffic, it is needed to remove outlier values to get an unbiased result. In facts, it is too difficult to define the range of outlier value.

Security event monitoring is collecting security events from IDS. IDS is carrying out signature based detection and it generates events like source IP, destination IP, source port number, destination port number, protocol and category of threat, degree of threat and threat name, etc. After gathering events and analysing them, we are generating outputs such as top threats, severity of each threat, suspect's IP lists and infected IP lists, etc.

In threshold setup process, users can input threshold of traffic volume that determining the point of time to receive alarm message using Short Message Service (SMS) from the system. Users set up the threshold value that how many times real-time traffic larger than baseline traffic. We also use the other condition that look in the traffic using flow such as what service and how many flows per second a user create. We examine thoroughly if security events have a critical threat or not at the same time and traffic volume and critical events together could provide more accurate detection result.

Alarming is sent by two kinds of methods like SMS and Email and they are sent for system operators and customers. Also, system gives web based reporting services including traffic volume, top attacks and infected IP lists.

Snapshot is very unique function of this system; it is taking a snapshot of web report and is storing image files in a directory. If some users want to see the past snapshot images, we can see the files using the web browser again.

Because of ATDS developed as a security service for enterprise customers, it consists of some additive function such as file output and printout. Some user wants to do data processing using Microsoft Excel so that excel formatted output is provided.

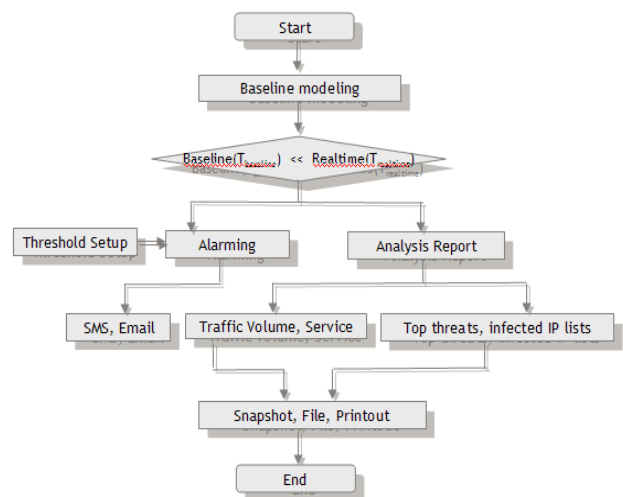


Figure 7. Anomalous traffic detecting procedure

Figure 7 shows an anomalous traffic detection procedure that it starts from baseline modelling, anomalous traffic detection to alarming and analysis reporting. At first, we need to make baseline traffic for each customer based on time of day, day of week and special days with a sufficient learning period about four weeks. If new arrival traffic (real-time traffic) exceeds its baseline traffic, ATDS checks the threshold value and sends

alarming by SMS and Email. ADTS also gives a detailed analysis report such as traffic volume, service type, top threats and infected IP lists in form of HTML and Microsoft Excel format.

V. IMPLEMENTATION

We implemented an anomalous traffic detection system and deployed on the aggregation points of enterprise network. We give anomalous traffic detection as a value-added service to our enterprise customers. From Figure 8 to Figure 10 are sampling reports of this ATDS. Figure 8 shows four weeks traffic view of some customer of same hour same minute.

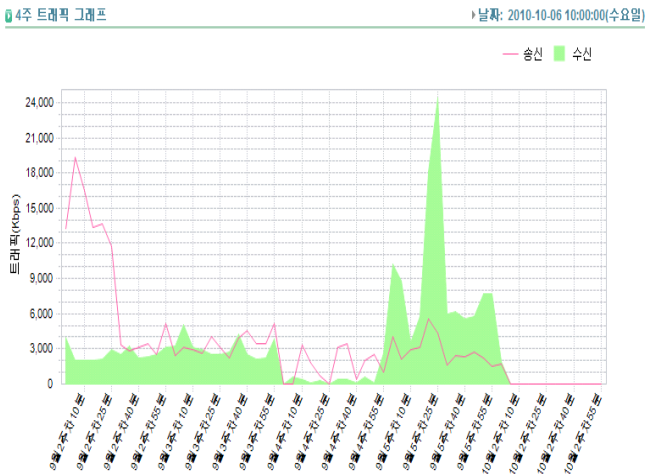


Figure 8. Four weeks Traffic View

In Figure 8, green area stands for download traffic volume from internet to customer network and red line stands for upload traffic volume, vice versa. Traffic volume is rising more than usually high on the right side of Figure 8. At that time, we can analyse the peak traffic using flow and security events.

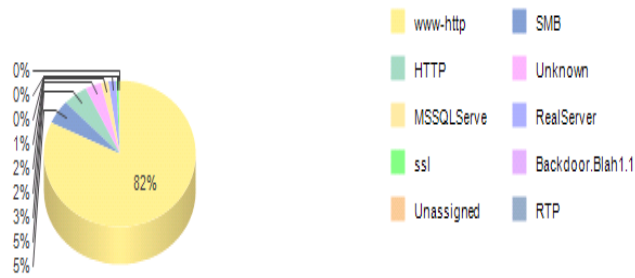


Figure 9. Service categories of traffic

In Figure 9, we analyse the traffic using flow such as what service user generate in detail. By above graph, the portion of web usage is over 80% and others are about 20%.

We also examine thoroughly if security events have a critical threat or not.

In Figure 10, we find out the characteristic of the traffic and suspect’s IP lists and infected IP lists of traffic. By the graph of Figure 10, we may estimate some problem in internal server of last one x.x.x.x and have to observe it closely.

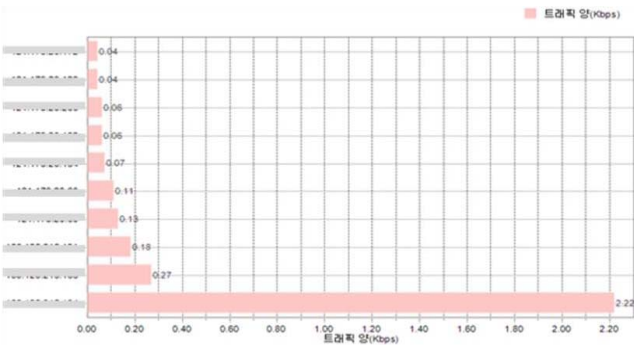


Figure 10. Infected IP Lists of Customer Network

VI. SUMMARY AND FURTHER WORK

In this paper, we developed an anomalous traffic detection system and deployed on the aggregation points of enterprise network. This approach is exploring an anomalous traffic detection based on time series traffic modeling and analyzing traffic by time of day, day of week, and special days. To improve the accuracy of detection, we analyse flow statistics and security events. Especially, ATDS provides traffic view of same time same minute during four weeks so that it detects traffic anomaly more accurate. ATDS gives practical results to our customers but, we have some false positive detection. After that, we need to study more to find better detection method to improve an accuracy of determining an anomaly as a further study.

REFERENCE

- [1] Foster, J.C: IDS: Signature versus anomaly detection (2005) <http://searchsecurity.techtarget.com>.
- [2] Thottan, M., Ji, C.: Anomaly detection in IP networks. Signal Processing, IEEE Transactions on 51 (2003) 2191-2204
- [3] Barford, P., Plonka, D.: Characteristics of network traffic flow anomalies. In: Proceedings of the 1st ACM SIGCOMM Workshop on Internet Measurement, ACM (2001) 69-73
- [4] Brutlag, J.D.: Aberrant behaviour detection in time series for network monitoring. In: LISA '00: Proceedings of the 14th USENIX conference on System administration, USENIX Association (2000) 139-146
- [5] Mahoney, M.V.: Network traffic anomaly detection based on packet bytes. In: SAC '03: Proceedings of the 2003 ACM symposium on Applied computing, ACM (2003) 346-350
- [6] Anderson, D., Frivold, T., Valdes, A.: Next-generation intrusion-detection expert system (NIDES). Technical Report SRI-CSL-95-07, Computer Science Laboratory, SRI International (1995)