Cloud Based Encrypted Traffic Analysis System using Netflow Information

Jungtae Kim, Jong-Hyun Kim and Ikkyun Kim
Information Security Research Division
Electronics & Telecommunications Research Institute
Daejeon, Republic of Korea
e-mail: {jungtae_kim, jhk, ikkim21}@etri.re.kr

Koohong Kang
[2]Dept. of Information and Communications Engineering
Seowon University
Cheongju, Republic of Korea
e-mail: khkang@seowon.ac.kr

Abstract— The paper proposes an encrypted traffic analysis system in cloud network environment. In cloud computing, various services are driven by Virtual Machines (VM), and the most of common application are currently using an encryption methods for the public communication. We propose a method for generating netflow and session information for each VM in various cloud based machines and analyzing encrypted traffic, such as SSL / TLS sessions. The proposed traffic analysis system further helps to detect a web-based HTTPS attack traffic or DDoS traffic by analyzing characteristics of the corresponding encrypted traffics in real time.

Keywords-HTTP Get Flooding; Netflow; DDoS Attack; .

I. INTRODUCTION

Conventional network traffic analysis methods [1] analyzed packet headers and payloads based on IP packets and checked whether traffic is abnormal based on a specific pattern or a signature provided by third parties. However, in a cloud server environment, various virtual machines (VMs) on a single server will provide each OS and service. Therefore, each VM is allocated a private IP to communicate internally and externally. In terms of Open Virtual Switch (OVS), which manages communication between servers, all communication is performed via VLAN, which involves a problem to identify and analyze the flow and session information in detail. In order to overcome the limitations, our paper introduces with a unique method for analyzing traffic encrypted with Secure Sockets Layer (SSL) / Transport Layer Security (TLS) based on the 5-Step configuration method, which further helps to analyze the characteristics of the traffic in real time for detecting web based Hypertext Transfer Protocol Secure (HTTPS) DDoS attacks. The paper is organized with a Literature Review in Section 2, an overview and experimental results on the Proposed Encrypted Traffic Analysis System in Section 3. Finally, Conclusion and Future Works are discussed in the Section 4.

II. LITERATURE REVIEW

As the encrypted traffic is not possible to identify its payload or content information, the network level behavior analysis with advanced netflow information is the only approach, which helps to achieve the goal. Netflow is a feature that was introduced on Cisco routers that provides the ability to collect IP network traffic as it enters or exits an interface [2]. The major advantage of utilizing the flow data is that it helps to analyze the both unencrypted and encrypted network traffics with basic parameters including: source and destination IP address, source and destination port, layer 3 protocol type, byte, packet, etc. [3]

III. ENCRYPTED TRAFFIC ANALYSIS SYSTEM

As shown in Figure 1, we propose a traffic analysis system with the cloud based flow-generating routers or virtual switches.

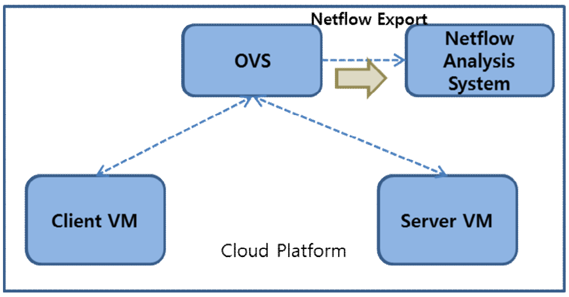


Figure 1. Cloud-based Netflow Collection System.

As shown in Figure 2, the encrypted traffic analysis system involves 5 staged methods. The analysis system does not only examine and classify the encrypted traffic patterns, but also the general traffic are classified.

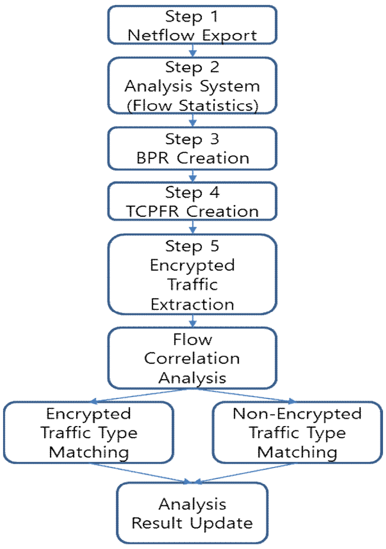


Figure 2. Sequence Flow Chart of the Encryption Traffic Analysis System.

Copyright (c) IARIA, 2018. ISBN: 978-1-61208-674-3

- Step 1 : configure netflow export through OVS in the cloud environment and store related flow information including srcaddr/port, dstaddr/port, dPkts, dOctets, first/last time (at start/last packet of flow), tcp_flags (URG/ACK/PSH/RST/SYN/FIN), prot and tos.

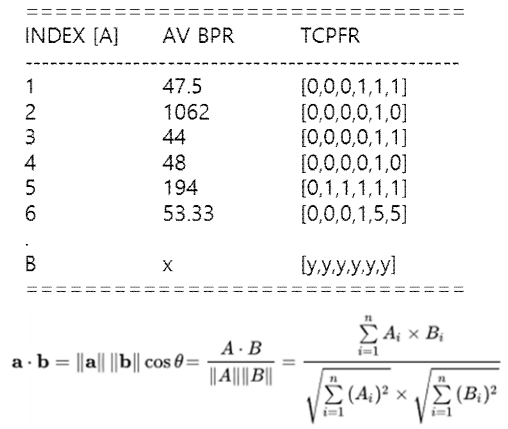
- Step 2 & 3 : extract and calculate the Byte per Packet Ratio (BPR) = dOctets / dPkts per flow in 2-dimensional array, which are initiated and terminated at the relevant time. At this step, the total number of packets and the bytes per flow record are stored for each encrypted session. In other words, if a specific session lasts for 2 minutes, all flow records of the session corresponding to the 2 minutes are searched and collected and statistical values are extracted.

- Step 4 : extract and calculate the average TCP flag information (Average Number of Flags = Total Number of Flags / Total Number of Flows) in 2-dimensional array for each session [URG:0, ACK:0, PSH:0, RST:66, SYN:66, FIN:66]= [0,0,0,1,1,1]. At this stage, the threshold parameters need to be set by analyzing the results of various network level attack tools. For example, as shown in the Table I, the HTTPs based Get Flooding attacks has average BPR of 47.5, which only involves a short repetitive TCP handshake between client and web server for the authentication and key exchange.

TABLE I. COLLECTED INFORMATION OF THE ENCRYPTED TRAFFIC ANALYSIS SYSTEM.

INDEX	1
PROTOCOL	TCP
SOURCE	192.168.120.21: random
DESTINATION	1.245.4.48: 443
BEGIN TIME	2017-02-01 11:20:49.771
END TIME	2017-02-01 11:20:50.000
TCP FLAG	SYN/FIN/RST
PACKETS / OCTETS	4 / 205, 5 / 245, 6 / 285, 7 / 333
Total Packet/Total OCTETS	264 / 13,530
Total Flows	66
AV Packet / Byte	6 / 285
BPR	[51.25, 49 ….. 47.5, 47.5714]
TCPFR	[URG:0, ACK:0, PSH:0, RST:66, SYN:66, FIN:66] = [0,0,0,1,1,1]
Threshold	AV BPR (47.5) / EndTime-BeginTime (1min) = 47.5
TRAFFIC TYPE	DDOS-HTTPs Get Flooding

- Step 5 : Based on the pre-collected attack traffic analysis information, we determine the encrypted traffic types by calculating the cosine similarity between attack and normal sessions as shown in the Figure 3. As the various encrypted sessions are easily distinguished through dstport (TCP / UDP destination port number) in the flow record such as web communication HTTPS 443, email IMAP 993, POP 995, SMTP 465, SSH/SecureFTP 22, the attack traffic type can be identifiable with the collected flow information.



```
==============================
INDEX [A]     AV BPR       TCPFR
------------------------------------------------
1             47.5         [0,0,0,1,1,1]
2             1062         [0,0,0,0,1,0]
3             44           [0,0,0,0,1,1]
4             48           [0,0,0,0,1,0]
5             194          [0,1,1,1,1,1]
6             53.33        [0,0,0,1,5,5]
.
B             x            [y,y,y,y,y,y]
==============================
```

$$\mathbf{a} \cdot \mathbf{b} = \|\mathbf{a}\| \, \|\mathbf{b}\| \cos\theta = \frac{A \cdot B}{\|A\|\|B\|} = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n} (A_i)^2} \times \sqrt{\sum_{i=1}^{n} (B_i)^2}}$$

Figure 3. Similarity Calculation between the Attack vs Encrypted Traffic with the Average BPR and TCPFR.

The AV BRP and TCPFR information for the encrypted session B is analyzed with 7-dimensional vector with other encrypted attack types (Index 1~N), and the cosine similarity of the analyzed unknown attack session is calculated between 0 for independent to 1 for identical attack session in the collected attack profile information.

IV. CONCLUSION AND FUTURE WORK

We propose an analysis method of the encrypted attacks traffics based on the netflow flow data (Cflow, Jflow, Netflow) provided from existing network devices such as routers or switches, but also to traffic analysis using flow data provided by OVS in the cloud network environment. The proposed encrypted traffic analysis system utilizes the BPR and TCPFR for each flow generated and terminated at the corresponding time to analyze SSL/TLS encrypted traffic to detect a web-based HTTPS attack or encrypted DDoS traffic by analyzing characteristics of the corresponding encrypted traffics in real time.

ACKNOWLEDGMENT

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIP) (No.R-20160222-002755, Cloud based Security Intelligence Technology Development for the Customized Security Service Provisioning).

REFERENCES

[1] P. Velan, M. Cermak, P. Celeda and M. Drasar, “A Survey of Methods for Encrypted Traffic Classification and Analysis,” International Journal of Network Management, 2014 [accessed Aug 2018]

[2] Cisco IOS NetFlow, Cisco Systems, Inc. https://www.cisco.com/c/en/us/products/ios-nx-os-software/ios-netflow/index.html [accessed Aug 2018]

[3] NetFlow Export Datagram Format, Cisco Systems, Inc. http://www.cisco.com/c/en/us/td/docs/net_mgmt/netflow_collection_engine/3-6/user/guide/format.html [accessed Aug 2018]