Suspicious Traffic Detection Based on Edge Gateway Sampling Method

Sinh-Ngoc Nguyen, Jintae Choi, Kyungbaek Kim
Department of Computer Engineering
Chonnam National University
Gwangju City, South Korea
sinhngoc.nguyen@gmail.com, jefron1100@gmail.com, kyungbaekkim@jnu.ac.kr

***Abstract*—Packet sampling is commonly deployed in all of Intrusion Detection System (IDS) to block the resources consumed from DDoS attack in the network. The IDS usually places sample collector either at distributed points in the network or next to the victim. The sampling methods use the threshold of traffic to detect the attack. It will stop an attack if having a matching with the threshold which is defined in the rule of IDS. We assume that there are lots of suspicious traffics with small volume going to the same destination. The IDS with distributed sampling method cannot detect these traffics. Because of small volume of suspicious traffic, it is not large enough to match the threshold in the rule. Although suspicious traffics have small volume, lots of the traffics generate a large volume at the same destination of victim. To handle this problem, placing the sample collector next to the victim and implementing the destination rule in IDS are proposed, the destination rule can detect multiple traffics that have the same destination IP. However, it still gets the problem that it lets the suspicious traffic going through the network, which generates a large number of unnecessary traffics and causes to the low performance of network. In this paper, we propose a sampling based approach that samples the traffic at the edge gateway in a Software Defined Networking (SDN) based network. It could detect the DDoS attack earlier before reaching to the complex core network, and it could determine which domain the DDoS attack comes from.**

Keywords—DDoS Attack; Edge Gateway; Sampling Method; SDN

I. INTRODUCTION

With the growth in Internet traffic and the growing variety of Internet application, network security is a key element of the system. It helps a system to face the attack from many other malicious sources. To deal with the attack, we have to implement detecting and preventing techniques to protect the system. As many previous studies have indicated, several detection approaches have been proposed such as statistical, soft-computing, clustering, knowledge-based. These approaches can also be classified as supervised or unsupervised [1].

In the case of large-scale network showed in Fig 1, it includes a core network that connects many sub-networks to many other domains. The IDS is a center security of the network, it can be placed inside or in-line of the network [2]. The sampling based method is deployed with distributed sampling point to collect the data going through the network. Then the sample traffic in each switch will be forwarded into IDS for analysis. However, the increase of network traffic and the expansion in the network scale, it is hard to determine the detection point where IDS is placed and where the data packet is captured in the network. Also, because of the hardware limitation in IDS such as CPU power, memory access speed, and storage capacity, the high cost of deploying multiple IDSs to inspect all the data packet of the network is caused [3].

Furthermore, distributed sampling in the network may lead to a problem that missing the suspicious traffic of attacks. If there are many suspicious traffics from many different domains with small volume of traffic that going to the similar destination, the IDS with distributed sampling based method cannot detect this traffic as well. Because a small volume of suspicious traffic is not large enough comparing to the threshold of the rule in IDS, this kind of traffic will be passed through all the nodes in the network. It leads to the problem that many of small traffics with the same destination can cause to large volume traffic that floods the victim.

To handle multiple suspicious traffics that have the same destination as above problem, the IDS and sampling collector are placed next to the destination point which is showed at Fig 2. This deployment is usually used to monitor the network behavior and detect the malicious traffic which toward the cloud server. It can detect the suspicious traffic with small volume by using destination rule which is implemented in the IDS. However, this deployment gets the problem that the suspicious traffics still exist in the network. It makes the network be slow with many traffics going through it. Furthermore, it cannot indicate exactly which domain the suspicious traffics come from.

Recently, SDN is an emerging network architecture that enhances the performance of IDS in inspection the attack traffic and management the large-scale network as well as the cloud network. It has a centralization view to the network, called SDN controller which indicates the network policy to face the attack in the network. SDN controller uses OpenFlow protocol, which enables the communication in the network between controller and switches. SDN is a promising solution to enhance the network security and network management in the system. By using SDN-based network, we can easily sample data traffic from multiple switches and forward them to the IDSs for analysis.

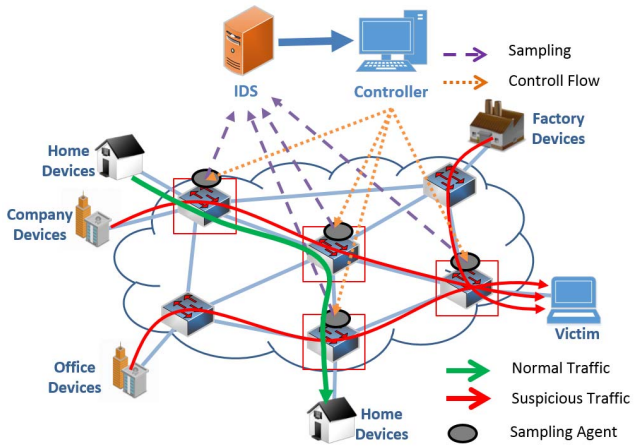

Fig 1. IDS for Large-Scale Network

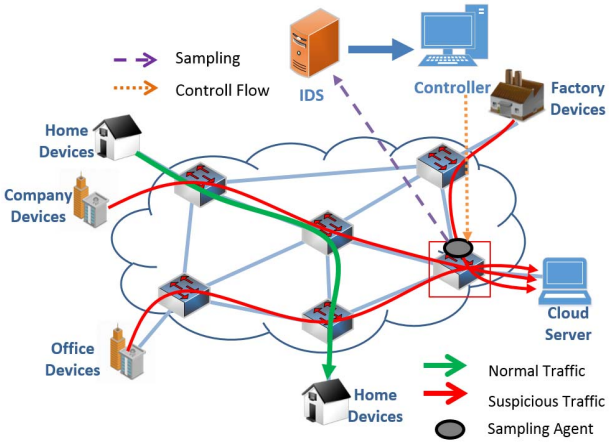

Fig 2.IDS for Cloud Server

In this paper, we propose the IDS with sampling based method at the edge gateway to detect the DDoS attack as well as the suspicious traffic in SDN based network. It can detect attack with small volume of traffic by implementing the appropriate sampling rate of each domain. And it blocks the attack at the edge of network to prevent the unnecessary traffics going through the complex network. Also, we provide the algorithm to determine edge gateway in the network, and the method to evaluate the sampling rate of each domain.

II. Related Work

There have been some important researches on intrusion detection technology. Most of them focuses on how to process a large amount of traffic samples efficiently. They proposed various traffic characteristics analysis such as the flow statistics, flow size, and flow entropy [3,4,6,7]. Ryoichi Kawahara et al. [3] use flows statistic obtained through packet sampling to detect network anomalies. By the source IP addresses, they partition the monitored traffic into several groups and analyze them individually. This method increases the detectability of such anomalies. Androulidakis et al. [4] proposed a flow-size sampling based techniques to address the abnormal detection issue. The authors analyze the network traffic and characterization of dynamic statistically properties in order to accurately and timely detect network anomalies.

In [6], Mai at al. apply several algorithms to detect the *portscans* in sampled packet traces. The authors use TRWSYN algorithm to perform traffic analysis, TAPS algorithm to tracks the connection pattern of scanners, and entropy-based traffic profiling. The results indicate that portscan algorithm can be enhanced to be more robust to sampling. Kacha et al. propose a new pattern matching technique [7] to improve the performance of Snort IDS. The authors modify Snort signatures to provide a faster packet detections and minimize the resource consumption.

Recently, SDN technology is an emerge architecture for network security. There have been many researches using SDN technology to enhance the performance of IDS in the network [5,8,9,10]. In [5] Giotis et al. extend diversity functionalities of network with an efficient and scalable mechanism to detect the anomalies traffic and mitigate in OpenFlow SDN. In [8] Ha et al. propose a traffic sampling rate at network switches for fully utilizing the detection capability of malicious traffic while the volume of sampled traffic is kept below the maximum processing capability of IDS. The authors consider an IDS to prevent the malicious data from SDN-based network.

Zhang et al. propose a novel method that performs adaptive zooming in the aggregation of flows to be measured [9]. This method is used to balance the monitoring overhead and the anomaly detection accuracy. Also, the authors propose a prediction based algorithm that dynamically changes the granularity of measurement along both the spatial and the temporal dimensions. The authors in [10] use SDN to monitor the network traffic as compared to the previous existing networks. Also, they use adaptive monitoring function in SDN to enhance the efficient of anomalies detection.

In most of previous researches, they focus on sampling based methods to detect suspicious traffic which is observed from single point or distributed point in either traditional or SDN based network. In this paper, we propose an edge gateway sampling based method that samples the traffic at the edge gateway on the SDN-based network. In consideration of the IDS detection ability, we provide algorithm to determine the edge gateway and a method to estimate the sampling rate of a domain. Then we setup a testbed to evaluate our proposing method and show the result of detection ability.

III. Proposed Approach

A. Edge Gateway Sampling Based Method

We consider the IDS with distributed sampling based method in SDN based network which is shown in Fig 1. It composes several sample collectors placed on distributed switches of network. Sample traffic will be forwarded into IDS for analysis. The controller gets the result of detection from IDS and generates the flows which include the network policies. Then they are applied to the switch through OpenFlow by controller to block the attack traffic.

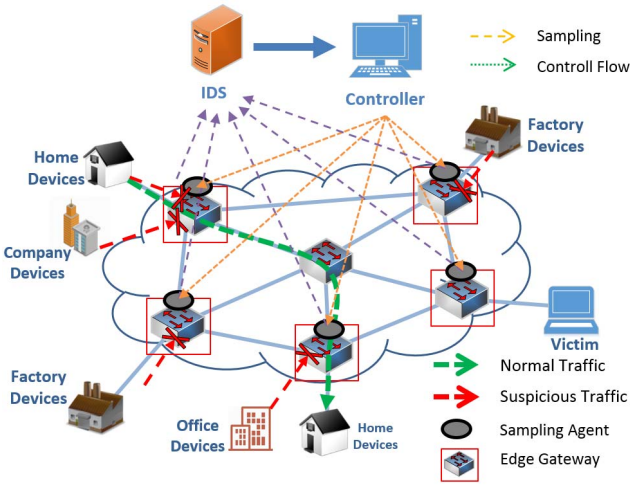

Fig 3. Edge Gateway Sampling

Assume that there are lots of suspicious traffics with small volume going to the same destination in the network. With the setting in the Fig 1, the IDS cannot detect these traffics with distributed sampling based method. Because the sampling based method cannot detect the attack traffics whose volume is small. Even though the suspicious traffics have small volumes at the source, lots of them having the same destination generate large traffics at the destination. To solve this problem, we can apply the setting in Fig 2. The IDS is placed in front of the Cloud Server and the destination rules are implemented to detect the attacks that have the same destination with small volume. However, there is still a problem here: the suspicious traffics generate a large amount of unnecessary traffics in the core network before going to the destination. Worse yet, the setting in Fig 2 cannot indicate exactly where the domain attack is come from.

In this paper, we propose an IDS with Edge Gateway sampling based method to detect DDoS attacks that have the small volume of traffics. In this method, we place the sample collector at the Edge Gateway of network which is shown in Fig 3. We provide the rule with sampling rate for each Edge Gateway depending on the traffic consumption of each domain. Then the IDS can detect the attack with small volume of traffic. Controller generates a new flow including the network policies and applies this policy to each Edge Gateway to block the suspicious traffics before going to the network. Also, our methods can remove the unnecessary traffic, improve the performance of transferring data in the network and it can detect where domain of the attack comes from.

In order to setup the rule on IDS with Edge Gateway sampling base method, we frequently consider the sampling rate, which is the total packets of each domain sending to the network over a period of time, at each domain.

$$SRate = \sum_{i=1}^{n} \frac{f_i . P}{t} + b \qquad (1)$$

Where ***SRate*** is the total packets of a domain that send to the network in a period time, f_i is the number of flows generated by device ***i*** in that domain, ***P*** is the number of packet that device *i* send to the network, ***t*** is the frequently time that we observe the packet, ***b*** is the bias of the function to adjust the accuracy of ***SRate***. We use the equation (1) to evaluate the sampling rate of each domain in frequently. Then we can indicate the number of packet sending to the network in a period time of each domain. That is the threshold of the rule in IDS.

Algorithm 1: Edge Node detection
Input: a network topology.
Output: a set of edge node.

```
H = ∅;  // set of hosts in network
S = ∅; // set of switches in network
E = ∅; // set of edge node in network
foreach node in topology:
    name = getNameNode(node)
    if name.contain("Host"):
        H.add(node)
    elseif name.contain("openflow"):
        S.add(node)
foreach link in topology:
    src = getSrc(link)
    dst = getDst(link)
    if S.exist(src) && H.exist (dst):
        if not E.exist(src):
            E.add(src)
return E
```

B. Finding Edge Gateway

In this work, we consider a directed graph $G_0 = \{V_0, E_0\}$, where V_0 is the set of nodes in the network, and E_0 is the set of edges on SDN based network. Set of nodes include both switches and hosts, each node corresponds to a switch, each host is an end-user device that connects to the network, and each edge is a link connecting between device and switch or two switches. In this network, the nodes include edge nodes placed at edge of network, which connects to host and other nodes; and core nodes placed inside of network, which connects to other nodes.

In order to determine whether a node is an edge node or not, we have to traverse all the nodes of the graph and check the existing of an edge between it and host. If a node that does not have any connection to host, then it is the core node. And if a node that has connections to host and other nodes, that is edge node. We propose the Algorithm 1 to detect the edge node in the network.

Algorithm 1 indicates how to get the edge nodes in the SDN based network. By using the information from rest API of OpenDayLight, we can easy to traverse all nodes in the topology to get the switches and hosts, then we check each node that have the connection to hosts and other nodes. We add it into E which is the edge node set.

C. Evaluation Setting

To evaluate the detection ability of our proposed approach, we setup a testbed with the settings which are showed at Fig 3. It includes a network with 6 switches connected together by

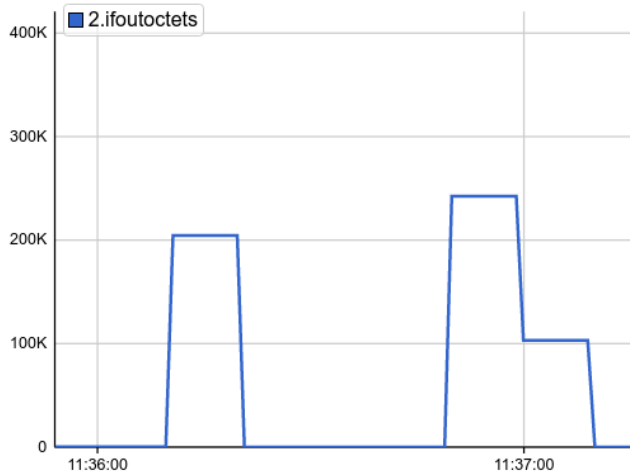

Fig 4. Result of Detection and Prevention

using OpenvSwitch. And the end-user devices (e.g cellphone, laptop, etc.) would be connected to the network through Edge Gateway. We deploy an IDS with the rule including the sampling rate calculated at above. The sample traffic at Edge Gateway will be forwarded to IDS by another network for analysis. Then the result of detection from IDS will be updated in database. We use OpenDayLight [11] as the controller to read the result from IDS and applies the new policies into the network to block the attack.

OpenDayLight controller would provide an overview on the network. And we implement *Algorithm 1* in our system. It is easy to traverse on all of the edges and nodes to find which node is the edge gateway in the network. In our testbed, we found 5 edge gateways that connect to other switches and host.

After finding the edge gateways, we setup the rule in IDS to analyze the sample traffic which is collected from edge gateways. By using equation (1), we can evaluate the sampling rate of each domain frequently, which is the total packet sending to the network in a period of time. This is the information of threshold parameter in the Snort rule inside the IDS. If there is any suspicious traffic that has volume larger than the sampling rate threshold in the rule, it will be reported as the attack.

Finally, the attack traffic is generated by Hping3 tool. It supports any protocol such as TCP, UDP, ICMP …, and several useful functions. Also, it can specify the interval time to send a packet to victim as well as the speed of sending packets. We install Hping3 on several devices and make the attacks to victim to test the detection and prevention ability on our system. The result will be showed in next section.

D. Result

The result of our experiment is shown at Fig 4, which indicates the detection and prevention ability of our method. When there is an attack, the traffic will increase high, then the IDS would detect the attack and the policies to block that traffic at the edge gateway in the network is applied by the controller. Then the traffic will go down. After a period of time, the block policies will be expired, and the attack traffic would go high again.

IV. CONCLUSION

In this paper, we proposed an edge gateway sampling based method to detect the suspicious traffic in SDN based network. The proposed method can detect the suspicious with small volume of traffics on IDS by estimate the appropriate sampling rate at each domain and using proper destination rule. With the useful functions from SDN technology, our proposed method determines and measures the suspicious traffics at edge gateways according to sampling rate. The experiment result indicated the detection and prevention ability of our proposed method.

ACK NOWLEDGMENT

This research was supported by the MSIP(Ministry of Science, ICT and Future Planning), Korea, under the ITRC(Information Technology Research Center) support program (IITP-2017-2016-0-00314) supervised by the IITP(Institute for Information & communications Technology Promotion). This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Science, ICT & Future Planning(NRF-2017R1A2B4012559).

REFERENCES

[1] Gogoi, Prasanta, et al. "A survey of outlier detection methods in network anomaly identification." The Computer Journal (2011): bxr026.

[2] Shin, Seungwon, and Guofei Gu. "CloudWatcher: Network security monitoring using OpenFlow in dynamic cloud networks (or: How to provide security monitoring as a service in clouds?)." Network Protocols (ICNP), 2012 20th IEEE International Conference on. IEEE, 2012.

[3] Kawahara, Ryoichi, et al. "A study on detecting network anomalies using sampled flow statistics." Applications and the Internet Workshops, 2007. SAINT Workshops 2007. International Symposium on. IEEE, 2007.

[4] Androulidakis, Georgios, and Symeon Papavassiliou. "Improving network anomaly detection via selective flow-based sampling." IET communications 2.3 (2008): 399-409.

[5] Giotis, Kostas, et al. "Combining OpenFlow and sFlow for an effective and scalable anomaly detection and mitigation mechanism on SDN environments." Computer Networks 62 (2014): 122-136.

[6] Mai, Jianning, et al. "Impact of packet sampling on portscan detection." IEEE Journal on Selected Areas in Communications 24.12 (2006): 2285-2298.

[7] Kacha, Chintan C., Kirtee A. Shevade, and Kuldeep S. Raghuwanshi. "Improved Snort intrusion detection system using modified pattern matching technique." International Journal of Emerging Technology and Advanced Engineerings (IJETAE) 3.7 (2013): 81-88.

[8] Ha, Taejin, et al. "Suspicious traffic sampling for intrusion detection in software-defined networks." Computer Networks 109 (2016): 172-182.

[9] Zhang, Ying. "An adaptive flow counting method for anomaly detection in SDN." Proceedings of the ninth ACM conference on Emerging networking experiments and technologies. ACM, 2013.

[10] Garg, G., and R. Garg. "Efficient anomaly detection using adaptive monitoring in SDN." International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE) 5.6 (2015): 498-501.