IOP Conference Series: Materials Science and Engineering

PAPER • OPEN ACCESS

Using SVM to Detect DDoS Attack in SDN Network

To cite this article: Dong Li *et al* 2018 *IOP Conf. Ser.: Mater. Sci. Eng.* **466** 012003

Using SVM to Detect DDoS Attack in SDN Network

Dong Li, Chang Yu, Qizhao Zhou and Junqing Yu

Network and Computing Center, Huazhong University of Science and Technology, Wuhan, Hubei,China
Email: lidong@hust.edu.cn

Abstract. Software Defined Network(SDN) controller has the global view of the network, but it is vulnerable to DDoS attack. This paper proposes a new model to detect DDoS attack in SDN based on SVM . Firstly The model extracts several key features from the packet-in messages and measures the distribution of each feature by using entropy, then uses trained Support Vector Machine(SVM) algorithm to detect the DDoS attack. Experiments shows that this method can detect security events with high efficiency and mitigate the DDoS attack in real-time.

1. Introduction

Software-defined network separates a network's control plane from data plane, it becomes simple for network management. This makes SDN deployable in many network environments, especially in multi-tenant data center network.

But the novel network architecture faces some new serious security concerns. In traditional network, Distributed Denial of Service (DDoS) is easy to launch and hard to defend, also in SDN network it can overload the centralized controllers and break down the entire network[1].

Nowadays, OpenFlow [2] protocol are mostly used in SDN network to make switches communicate with an SDN controller. In OpenFlow, the switches uses packet-in messages to report to controller to obtain new flow rules for any data flow they do not know how to handle. So the DDoS attack may produce a large number of new data flows, then the switches will generate many packet-in messages to send to the controller, which will exhaust the resources of controller, such as CPU and memory, and it lead to a failure of the controller

In SDN it is more difficult to mitigate the DDoS attack Because the switches handle the packets only by flow entry received from the controller. they are more stupid than traditional devices. They cannot detect the malicious flows. what's more, the controller cannot judge whether it is under a DDOS attack based only on the number of incoming queries, for example, a flooding of queries may result from burst benign flows. especially The DDoS attack is often reflection-based in which the attackers may generate the flows based on any protocol (TCP, UDP, et al.) to trigger the attack against a controller. it is different from those of the known DDoS attacks, such as ICMP flood, TCP SYN flood, and HTTP flood. Correspondingly, existing detection and defense mechanisms are not fully applicable to this attack.

Some researches[3-5] have pointed out that the SDN controller is a vulnerable target of DDoS attacks. they put forward various methods against DDoS. For example, establishing a packet-in filtering mechanism to protect the SDN control plane[6]. it records the values of packet header fields and filters out packets that have the same values as the recorded ones. but it doesn't work if the attackers will fake the packets which have different values from the recorded ones. In [7] an detection method based on the entropy variation of the data flows' destination IP addresses is proposed. It assumes that the destination

IP addresses are almost evenly distributed in the normal flows, but attacker also can fake normal traffic to overload the controllers.

This paper aims to establish a novel mechanism to protect SDN network. Based on machine learning framework, firstly it collects traffic data from the packet-in messages, and extracts some key features' value, such as srcIP(source IP address), srcPort(source port), desIP(destination IP address), desPort(destination port) etc. Then using entropy to measure the distribution of these features. Via training the model by normal and abnormal traffic data and comparing to some other machine learning algorithms, experiment shows SVM is a better framework to detect the DDoS attack in SDN network.

2. Problem description

In this section, we describe the DDoS detection problem in the SDN network. Fig. 1 illustrates the logical view of the network model. A DDoS attack could come from any port of any switch, therefore the goal of DDoS detection is to detect the attack and locate the switch and port.

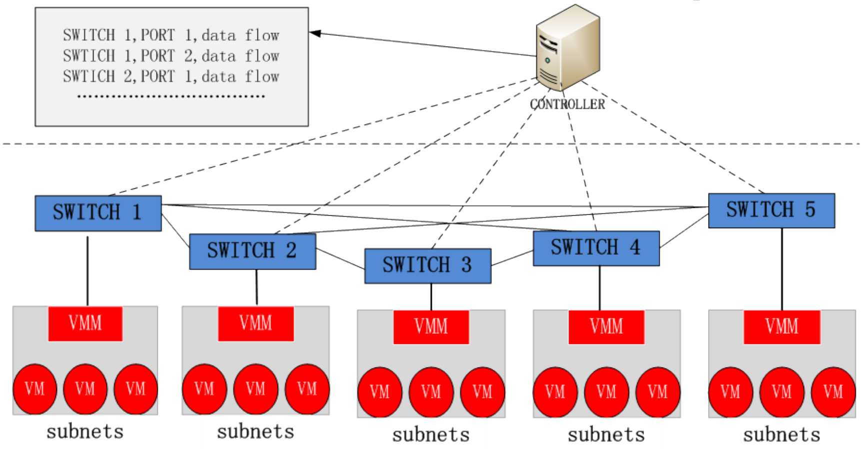


Figure 1. SDN Network Description

In the detection, suppose that each switch will ask the controller for new flow rules to forward unmatched packets and the controller is capable of obtaining the statistics information of the incoming flows from the switch via OpenFlow protocol. This can be achieved in SDN network. First, OpenFlow itself can monitor per-port and per-rule byte and packet counters. Second, sFlow can estimate the byte and packet counts of flows by fetching real-time packet samples from switches for flow statistics.

In addition, it is reasonable to suppose that attackers tend to generate vast spoofed packets to launch an attack in an SDN network by forging IP and port values which will trigger a great number of packet-in messages to congest the controller.

3. DDoS detection model in SDN

In this section, we present a method for detecting DDoS attack in SDN network. Based on training dataset of normal and abnormal traffic. This work is divided into two main stages, the first one consists of extracting robust and discriminative features from the packets such as source address ,destination address, source port, destination port and protocol type, then using entropy to measure the distribution of each feature. In the second stage, we use the resultant features' entropy to train a nonlinear SVM algorithm, based on the trained algorithm we can predict if the traffic of switch port is abnormal or not.

3.1 Entropy

In information theory, entropy is used to measure the randomness of a system, higher entropy means the system with bigger randomness[8].

In normal situation, the packets are produced by some trigger randomly, but when the DDoS attack happens, the distribution of packet features will change vastly. Often there will be a large volume of different source IPs, destination IPs and the IP flow is very small.

For detecting the DDoS attack, we use entropy to measure the randomness of packet feature. In one time window, we calculate the entropy of packet features of one switch port. When DDoS happens the entropy of this switch port will decrease dramatically.

Here we describe how to calculate the entropy. Suppose a random variable X has N different values and their probability is p_1, p_2 ... p_n, then the entropy of X can be calculated by formula [1]:

$$H(X) = -\sum_{i=1}^{n} p_i * \log p_i \tag{1}$$

H(X) can measure the uncertainty of X, bigger H(X) means higher uncertainty. In our model, the packet features <srcIP,srcPort,desIP,desPort> are all random variables , so we can calculate the entropy of them. For example, in one time window the number of source IPs of packets received by a switch port as follows: IP1(10),IP2(15),IP3(25), then the entropy of source IP is calculated as:

$$H(\mathrm{srcIP}) = -\left(\frac{10}{50}\log\frac{10}{50} + \frac{15}{50}\log\frac{15}{50} + \frac{25}{50}\log\frac{25}{50}\right) = 0.45$$

3.2 SVM

Support Vector Machine (SVM) proposed by [9] is considered as a statistical learning method for classification and regression. Afterward, SVM has been adapted to non-linear problems with using kernel methods[10]. The kernel function is defined as the following :

$$k(X, X)' = \Phi(X) \cdot \Phi(X') \tag{2}$$

Φ(x) is defined for solving non-linear classification problem and project the original input data χ to new feature space H where the classification problem has a linear solution. In our case, we use the RBF kernel function .

$$f(X) = k(||X - X_c||) \tag{3}$$

$X_c$is the center of the vector space. The objective of this learning method is to find a hyper plane separator to classify the input data which can be described mathematically as the following :

$$y(X) = \mathrm{sgn}(f(X)) \tag{4}$$

Statistical learning theory states that the optimal classifier can be found by maximizing the margin[11]. This can be expressed as a minimization problem:

$$\mathrm{Min}\tfrac{1}{2}||W||^2 \tag{5}$$

subject to :

$$Yi * (W.Xi + b) \geq 1, i = \{1, \ldots, n\} \tag{6}$$

where n is the size of input training data and Yi is data label (-1 or +1). In 1-class SVM framework, the data from only one class are available which matches our problem framework as we use only the normal event examples from the observed scene. The objective of one-class SVM, is to define a region in the space X which contains most of the data. This could be achieved by looking for a hyper plane in the feature space, and then try to maximize its distance from the origin, while only a small set of data is located between the hyper plane and the origin [11].

The algorithm is composed of two stages, the first one is the features extraction, and the second step is the classification. Fig. 2 summarizes the proposed algorithm:

Step 1 – Initialization: The feature are extracted from all the training packets set and the entropy will be used to measure the distribution of each feature. Then, the calculated feature entropy will be used in order to train nonlinear one-class SVM.

Step 2 – DDoS attack detection: For each new test packets, we extract features and calculate the entropy which will be given to the trained SVM model in order to decide if is normal or abnormal. If the result is abnormal, it means that DDoS attack happens.

```
Initialization:
N:number of packets set
For i=1:N
          #P: packets set
          X=Feature-extration(Pi)
          Feature_Entropy=Entropy(X)
          Features_Train=[Fearure_Train;Feature_Entropy]
End
      #size(Features_train)=N vectors of 5 dimensions
Model=Train_SVM(Features_Train)
```

```
DDoS Attack Detection:
M: number  of switch ports in SDN network
Pi: packets set of switch port i
For each  switch port i
        X=Feature-extraction(Pi)
       Feature_Test=Entropy(X)
       #size(Feature_Test)= |Pi| vectors of 5 dimensions
       Score(i)=Predict(Model,Feature_Test)
        For j=1: M
                   if score (j)<threshold
                           Patch j is Abnormal
                    End
        End
End
```

Figure 2. Algorithm of abnormal detection by SVM

4. Experiment and evaluation

For evaluating our method, we established an experimental platform based on MININET, the controller is Floodlight and the server has 64GB memory and 32 core CPU. Network topology is shown as Figure 3.

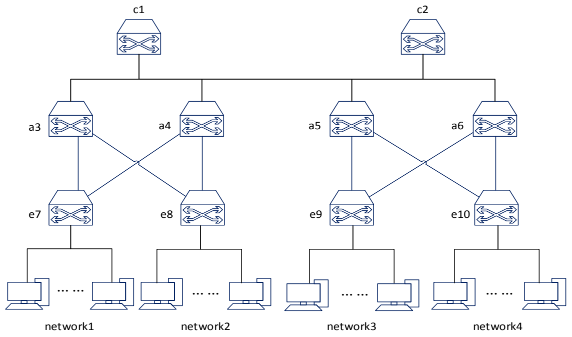


Figure 3 Experiment network topology

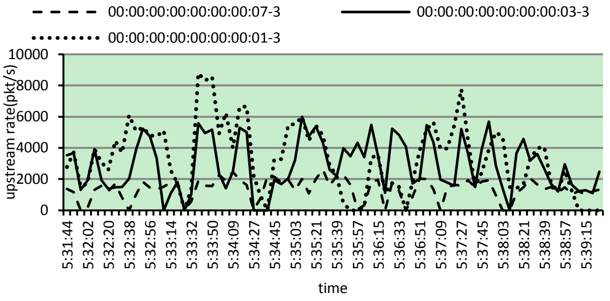


Figure 4. Background traffic

In our simulated network, the network topology adopts three-layer structure: Core layer, Convergence layer and Access layer. In Figure 3, C1 and C2 belong to core layer, a3 to a6 belong to Convergence layer and e7 to e10 belong to access layer. Otherwise there are 50 hosts in our experimental platform .

Background traffic

For simulating real network environment, normal traffic should be triggered as background traffic. In our experiment it is produced by the traffic generator D-ITG periodically，and the traffic ratio is: tcp:udp:icmp=85:10:5. in every period，hosts in network will establish connections randomly, and the packet sending speed is about 1000pkt/s. The speed of packets forwarding for switches in access layer, convergence layer and core layer are described in Figure 4.

For validating the effectiveness of entropy in measuring the abnormality of DDoS attack packet features, we simulate a DDoS attack under background traffic. Figure 5 shows the packet feature entropy will change obviously when DDoS happens.

4.1 Model training

For evaluating our proposed method ,we simulate several kinds of real DDoS attack in history. detailed description is shown in table 1

Table 1 Descriptions of DDoS attacks

Type	Description
A	100% udp flood
B	83% udp flood,14% icmp flood, 2% syn flood
C	40% udp flood,52% syn flood,6% icmp flood
D	62% udp flood, 20% syn flood,18% icmp flood

In training stage, the normal traffic is generated by the hosts in network. The software hping3 is used to simulate these DDoS attack with spoofed source address and destination address illustrated in table 1 and the duration of attack is 30s.

After the traffic is generated in the network, we can collect the data from the controller and filter the nonstandard traffic data， Once the time window is determined and the entropy of <srcIP,srcPort,dstIP,dstPort,proType> will be calculated to be a 5-dimention vector. These vectors are the sample of SVM model. the sample is divided into two groups: one of triggered by normal traffic and the other is triggered by DDoS attack traffic. What's more， DARPA1999 data set also is used to train our model. In training stage, the parameters of SVM will be fixed and used to analyze the real testing data.

4.2 Evaluation

For scoring the capability of DDoS attack detection, we use three indexes：PR(Precise Rate)、RR(Recall rate) and F1 score. The definition of these three indexes is formulated by formula (7),(8),(9).

$$PR = TP/(TP + FP) \tag{7}$$

$$RR = TP/(TP + FN) \tag{8}$$

$$F = (2 * PR * RR)/(PR + RR) \tag{9}$$

The meaning of TP,FP,TN is shown in table 2. PR can describe the accuracy of DDoS detection, RR can describe the integrity of DDoS detection and F can describe the balance between accuracy and integrity. The goal of our model is to get greater value of PR, RR and F.

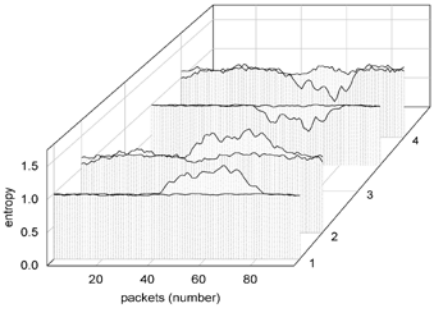
Table 2 Indexes of evaluation

	True	False
Positives alerts	TP	FP
Negative alerts	FN	TN

In our experiment, we simulated three DDoS attacks. For different attack, the ratio of malicious traffic is different, as shown in Figure 7, from the results we can see that our method can detect the DDoS attack efficiently and it works better to detect large-scale DDoS attack.

Also we use some other machine learning algorithms, such as decision tree, naive bayes, KNN, Random Forrest, to analyze the traffic and detect DDoS attack, as illustrated in Figure 8, we can see that SVM is the best one and can detect DDoS attack with higher accuracy and shorter time.

For testing the capability of mitigating DDoS attack, we make use of SDN global view to locate the switch port which is launching the DDoS attack, and coordinate with the controller to update the flow table of the switch to mitigate DDoS attack. figure 8 illustrates that after updating flow rules of the switch, the entropy of the switch port becomes normal immediately.



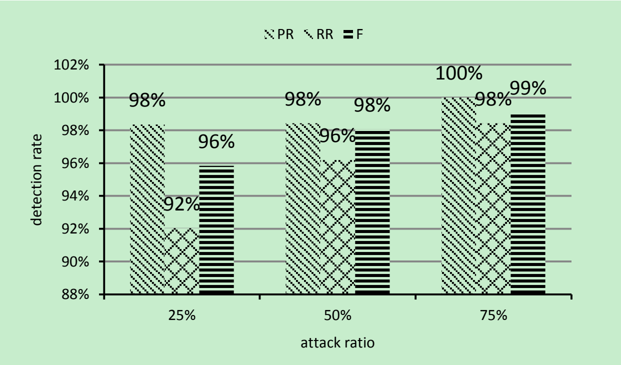
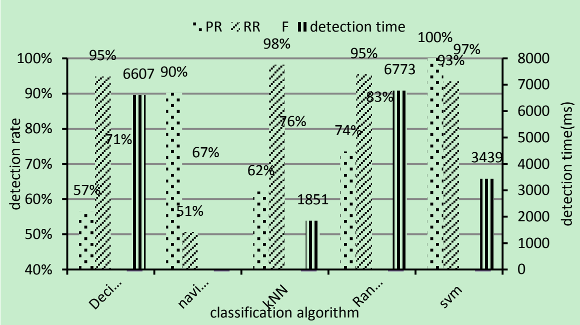


Figure 5. Entropy of <srcIP,srcPort,desIP,desPort> **Figure 6.** Comparison of DDoS attack detection



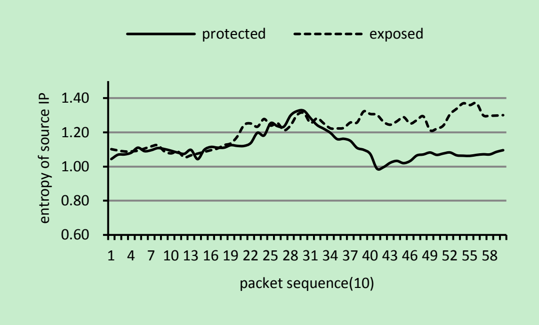


Figure 7. Comparison of different algorithms **Figure 8**. Entropy of switch port under DDoS attack mitigation

5. Acknowledgement

This work was Supported by National Key R&D Program of China (2017YFB0801703).

6. References

[1] Dacier M C, König H, Cwalinski R, et al. Security challenges and opportunities of software-defined networking. IEEE Security & Privacy, 2017, 15(2): 96-100.

[2] Open Networking Foundation, OpenFlow Switch version 1.5.1. https://www.opennetworking.org, 2015.

[3] d.F.Yu, Distributed denial of service attacks in software-defined networking with cloud computing, IEEE Communications Magazine, 2015, vol. 53, no. 4, pp. 52-59.

[4] K.Benton, L.J.Camp, and C.Small, Openflow vulnerability ACM SIGCOMM workshop on hot topics in software defined networking, 2013, pp. 151-152.

[5] P Dong ,X Du, H Zhang ,et al. A detection method for a novel DDoS attack against SDN controllers by vast new low~traffic flows. In: IEEE International Conference on Communications. 2016: 1~6.

[6] D. Kotani and Y. Okabe, A packet-in message filtering mechanism for protection of control plane in openflow networks, In Proceedings of the on architectures for networking and communications systems, 2014, USA, pp. 29-40.

[7] S. M. Mousavi，Ear1y Detection of DDoS Attacks in Software Defined Networks Controller，2014. Master dissertation， Carleton University Ottawa.

[8] Jun JH, Lee D, Ahn CW, Kim SH. DDoS Attack Detection Using Flow Entropy and Packet Sampling on Huge Networks. In: The Thirteenth International Conference on Networks. 2014: 185~190.

[9] Vladimir Vapnik. Pattern recognition using generalized portrait method. Automation and remote control, 24:774–780, 1963.

[10] Tian Wang and Hichem Snoussi. Histograms of optical flow orientation for visual abnormal events detection. In Advanced Video and Signal-Based Surveillance (AVSS),2012 IEEE Ninth International Conference on, pages 13 – 18. IEEE, 2012.

[11] Claudio Piciarelli, Christian Micheloni, and Gian Luca Foresti. Trajectory-based anomalous event detection. IEEE Transactions on Circuits and Systems for video Technology, 18(11):1544 – 1554, 2008.