www.revistageintec.net
ISSN: 2237-0722
revista GEINTEC Gestão, Inovação e Tecnologias

Detecting Ddos Attack Using Adaptive Boosting with Software Defined Network in Cloud Computing Environment

Sisay Wayu Tufa[1]; Mesay Mengstie[2]; Haftom Gebregziabher[3]; B. Ravindra Babu[4]
[1]Research Scholar, Ethiopian Technical University (ETU), Addis Ababa, Ethiopia.
[1]sisw24@gmail.com
[2]Research Scholar, Ethiopian Technical University (ETU), Addis Ababa, Ethiopia.
[2]ftveti-elt@outlook.com
[3]Research Scholar, Ethiopian Technical University (ETU), Addis Ababa, Ethiopia.
[3]habtilo@gmail.com
[4]Research Scholar, Ethiopian Technical University (ETU), Addis Ababa, Ethiopia.
[4]ravindrababu4u@yahoo.com

Abstract
Cloud computing is most widely used platform in past decade for computational operations in the computers and which offers the cost effective system with the measurable results. Cloud computing and software defined network (SDN) combination gives the better environment which reduces the difficulties with the cloud network and improves the dynamism, programmability, scalability and manageability of the cloud. Several weaknesses are attacked on one side by changing the SDN pattern into the centralized architecture namely as Distribute denial of service (DDoS) attacks. So these interrupts are detected and then prevented in SDN with the technology. Detecting of DDoS attack by using anomaly-based adaptive boosting in SDN cloud environment is presented in this paper. By coordinating the SDN features with the adaptive boosting algorithm, the cloud environment of SDN system DDoS attacks are noticed and prevented. Decision stump is used to generate the learning data and serves huge data to the servers within less time which is used as the prediction data. Predictions are can be processed by the formation of learning data. The experimental test results show that the adaptive boosting algorithm with SDN features has effectiveness in detecting attacks with high accuracy and in the SDN the DDoS attacks are detected even in the low communication.

Key-words: SDN, DDoS, Cloud Computing, Adaptive Boosting Algorithm, Attack Detection.

1. Introduction

From the central controller Software defined networking (SDN) involves certain operations as management, configuration and controlling [1]. SDN basis are dividing the data planes and control

technique. From the modern data centres which maintains the large data in every second with the servers of high volumes, main driver of the software defined networks are raised. The elements required in this method are high cost and the process is taking more time while configuration by the manual. By the management and central control, this problem is solved in the Software defined networks [2]. But the Software defined networks is having a weak organization in several threats in which the Distributed Denial of Service attacks (DDoS) attacks are the dominated [3]. These attacks are encouraged in the environment of SDN and by using centralized controller these attacks are disturbed by merging. To improve the scalability of data and performance requirements by taking the different new approaches in order to control the DDoS attacks which are raised their frequency and strength day to day and these new approaches gives the best protection of the system from the disturbing attacks [4].

In SDN controller the traffic flow can be disturbed and sometimes the unavailability of network occurred in the packets flow by the DDoS attacks. The system is mainly disrupted by these attacks and then the network flow pattern is changed accordingly [5]. The efficiency of the overall system is reduced by these attacks which causes the main problem in the internet and network [6]. The packets get delayed in the traffic so packets are overflowed and the appropriate route of the packets is difficult to predict. By this cause to loss the hidden packets because of it's over flow, and it is difficult for locating the packet pattern and routing. By considering these DDOS attacks the research has to be done very carefully for the packet successful flow and network smooth flow [7]. In this paper the DDoS attacks are detected by using the approach based on Information Distance in the cloud environment of SDN. Then the framework of adaptive boosting algorithm with SDN features is employed for attack detection systems on DDoS. At last, by using the experiments with the simulation results adaptive boosting approach effectiveness is demonstrated. Decision stumps are generates the huge learning data in which the attacks are detected effectively in account of SDN features by using the framework of adaptive boosting algorithm. So by using this frame work the attacks are detected and prevented with accurate results.

2. DDOS Attacks in Software Defined Networks

One of the most current interruptions in security systems is Distributed Denial of security Service attacks those are having the main aim to disturbing the access path of users to other network source or server. This attack is caused by merging the server to host therefore the resources such as traffic, CPU and memory are vanished in host. So this problem is raised when the data transferring between server to authentic users. Other attacks are overcome simply by rebooting but this flooding or

merging model is difficult. Different types of DDoS attacks are present in the cloud as Protocol based attacks, application based attacks and Volume based attacks [8].

TCP SYN (transmission control protocol synchronization) attack can adopt by the any type of service system or server which uses the TCP protocol in their underlying transport layer [9]. By using the TCP three way handshake mechanisms this TCP SYN attack is executed. The corresponding final ACK (acknowledgement) is not sending to the server by the client is makes the connection to interacting the attackers in the system. These connections are maintained with some memory in the server of web. Several connection requests with no final ACK are sending by the attackers to the server [10]. So that the overflow is occurred in the SYN-queue which makes the genuine hosts are disconnected. TCP SYN flooding attack is possibly targeted in the TCP socket, and any network service that is indeed to listened [11]. The Ping flood attack is other protocol attack. Focusing on depletion of resource, this Ping flood attack is the DDoS simplest form. The victim server can be requested by the number of IC mapping requests from the botnet Hosts. So the server buffer is overcome by the heavy traffic which makes the damaging of the system [12].

The Slow HTTP (hyper text transfer protocol) attack is one of the attacks in the system. The HTTP protocol working methods are activities and it requires fully completed HTTP request before it executed [13]. The Slow HTTP attack is an application oriented attack. So many HTTP requests which are incomplete (without carrier return line feed) are sending by the botnet client to the server web [14]. Request can be completed after waiting the victim server. So the server buffer is overcome the requests from clients and extra requests cannot be recognized in this time then there is a cause of dis-connectivity happens between client and server [15]. As a result knowledge about types of attacks and DDoS attack gives the better experience which is used to construction of a best system. It gives the idea about network security.

3. Framework Detecting Ddos Attacks Using Adaptive Boosting Algorithm with Sdn

For adaptive boosting approach, botnet client is sending the attacks along with data with no IP address. Two main reasons are present in this assumption. First one is, in the edge routers filtering methods are integrated in current networks and these are eliminates the bluffed source IP addresses inside the host network. Second reason, there is a large number of active bots in advanced botnets. So spoofing source IP addresses of requirements is removed. In general hundreds or thousands of average bots are present in the botnet. By generating a huge number of requests in bot then the highest request rate is achieved in DDoS attacks. Number of generating requests in the bot is equals to number of

generating requests in the legitimate clients. So the overall request rate improved and it depends on increasing DDoS attacks in single client.

3.1. Generalized Entropy

The probability of the finite discrete distribution is as,

$$P = \{p_1, p_2, p_3, \cdots, p_n\}$$

Where, $p_k \geq 0, (k = 1,2,3, \cdots, n)and\sum_{k=1}^{n} p_k = 1$. The distribution information entropy P is defined as the outcomes of the experiments which are amount of uncertainty for which are having the probabilities$p_1, p_2, p_3, \cdots, p_n$. Information entropy is defined by the Shannon as:

$$H(p) = \sum_{k=1}^{n} p_k \, log_2 \frac{1}{p_k} \qquad (1)$$

The random variable information is calculated by the information entropy in information theory. So the decrement in the entropy gives the increment in the variable certainty of information and variable randomness is increases with increment in the entropy. In the probability distributions the divergence and distance are also measured with the basics which are given by the entropy. The generalized entropy of order α is defined by the Rényi as:

$$H_\alpha(p) = \frac{1}{1-\alpha} log_2(\sum_{k=1}^{n} p_k^\alpha), \; \alpha > 0 \text{ and } \alpha \neq 1 \qquad (2)$$

Entropy defined by the Shannon's is the limiting case of $H_\alpha(p)$ for α→1.

Information distance: Let's take two probabilities of finite discrete distributions as $P = \{p_1, p_2, p_3, \cdots, p_n\}$ and

$$Q = \{q_1, q_2, q_3, \cdots, q_n\}$$

Where, $0 \leq p_k \leq 1$and $0 \leq q_k \leq 1 (k = 1,2,3, \cdots, n)$ and

$\sum_{k=1}^{n} p_k = \sum_{k=1}^{n} q_k = 1$. Between the two divergent distributions P and $Q$the divergence is measured by the information divergence. Information divergence of order α is defined by the Rényi.

$$D_\alpha(P||Q) = \frac{1}{1-\alpha} log_2 \left(\sum_{k=1}^{n} \frac{p_k^\alpha}{q_k^{\alpha-1}} \right),$$

$$\alpha \geq 0 \text{and} \alpha \neq 1 \qquad (3)$$

The defined information divergence in the equation 3 is asymmetric. i.e. $D_\alpha(P||Q) \neq D_\alpha(Q||P)$. This asymmetric property is to be overcome in finding the distance metrics. Information distance metric which is symmetric is $D_\alpha(P||Q)$ as follows,

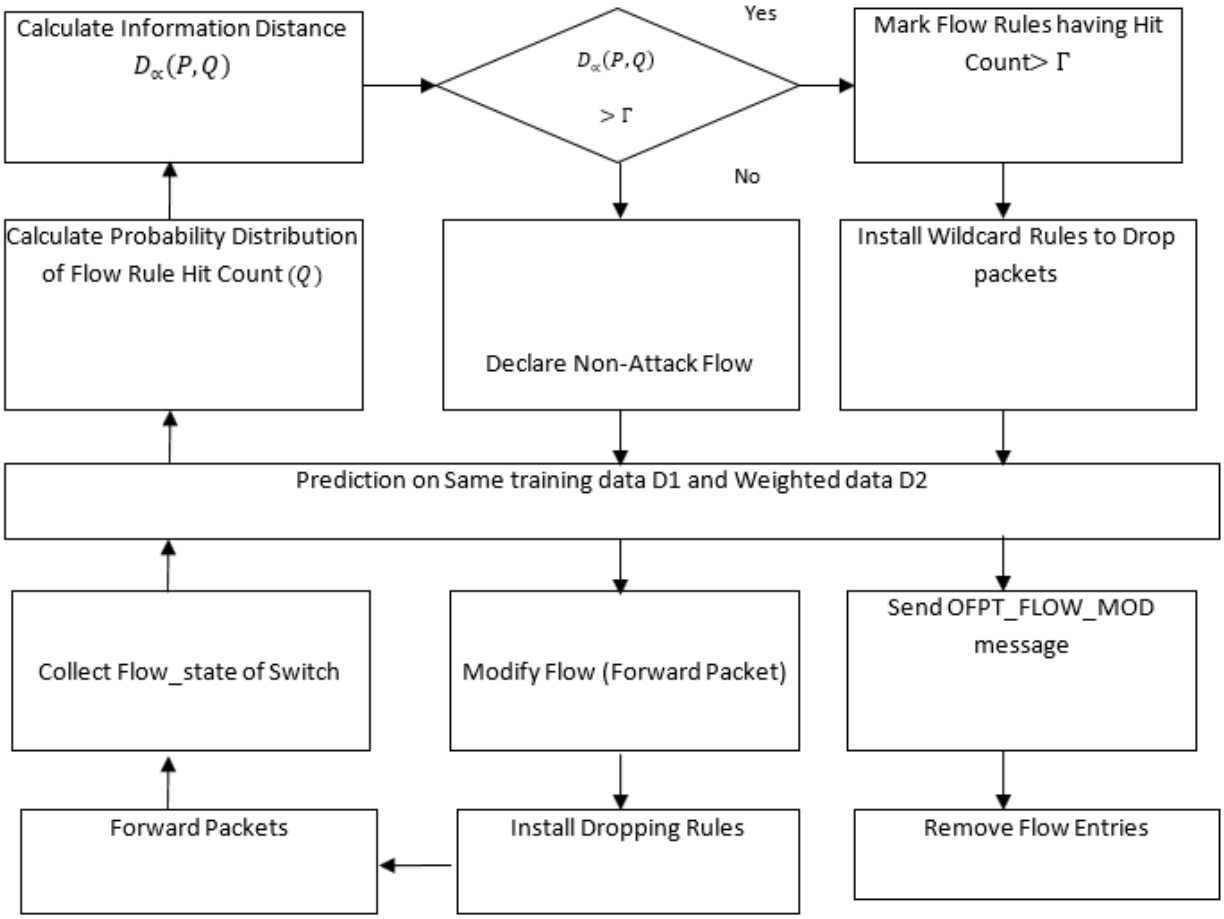
$$D_\alpha(P||Q) = D_\alpha(P||Q) + D_\alpha(Q||P)$$

$$= \frac{1}{1-\alpha} log_2 \left(\sum_{k=1}^{n} \frac{p_k^{\alpha}}{q_k^{\alpha-1}} \times \sum_{k=1}^{n} \frac{q_k^{\alpha}}{p_k^{\alpha-1}} \right) \quad (4)$$

3.2. DDoS attack Detection Approach

For packet forwarding a flow table is maintained by the each switch in the SDN based networking. Flow rules in the flow table switches are matching to the incoming flow. If the matching is finds then there is flow of packets in the forward direction as action mentioned. In the flow entry after packet forwarding the counter filed is also increased. The attack of flow entry in the counter field has high value by comparing it with the value of counter field flow rules of non-attack flows in case of DDoS attacks. In environment of cloud computing of SDN based, the DDoS attacks are detected by using the mentioned characteristics of the attack and non-attack flows. The DDOS attacks detection framework using adaptive boosting algorithm with SDN is depicted in figure 1.

Fig. 1 - Ddos Attacks Detection Framework using Adaptive Boosting Algorithm with Sdn



Received: 14.06.2021 – Accepted: 16.07.2021

In the period of non-attack, flow rule hit count of probability distribution $(P)$is calculated. Information distance $D_\alpha(P||Q)$ and flow rule hit count probability distribution (Q) and are calculated again when the incoming flow rate is exceeds the value of specified threshold T. The information distance is calculated by using the formula which is mentioned in the equation (4). High hit count value is obtained in the DDoS attacks because of high occurrence of attack flow probability. So the probability distributions are different for the distributions P as well as Q. The information distance $D_\alpha(P||Q)$ is increased with the increment in the probability distributions (P) and (Q)divergence between them. Therefore the value of $D_\alpha(P||Q)$ is small when the non-attack flow as the incoming flow, and the value of $D_\alpha(P||Q)$ is high when an attack flow as the incoming flow. The incoming flow is as DDoS attack flow if the observed$D_\alpha(P||Q)$ value is greater than the value of the specified threshold T. Two measurements are taken into account to minimize the DDoS attack, in that first one is greater than the threshold T hit count flow rule values are eliminated from the flow table. For attack client requests, there is a switch for wildcard flow rules. This is second measurement. So the attacks are eliminated at this switch by using these two measurements. The DDoS attacks in cloud environment SDN-based network are detected by using the adaptive boosting algorithm which as explained below.

3.3. Adaptive Boosting Algorithm

The improper observations are weighted by using the adaptive boosting algorithm and the formula for the decision stump is as follows;

$$f(x)=s(x_k > c) \qquad (5)$$

For vector x, element k value is more than threshold c, and then the prediction value is produced by the function $f(x)$ as 1. If element $k$value is lessthan the threshold value, then the prediction value of function$f(x)$as -1. So two functions are produced by the values either 1 or -1as $x_k > c$ and$x_k \leq c$. To get the final prediction value all predictions are added. Weighting$t = 1 \cdots T$ are applied for boosting iteration for each learning sample in (6) through (9).

Given:$(x_1, y_1), \cdots, (x_m, y_m)$ where $x_i \in X, y_i \in Y = \{-1, +1\}$

Initialization$D_1(i) = \frac{1}{n}$, where $n =$ amount of data for $t = 1\ to\ T$:

With distribution D_tTrain base learner

Get a weak hypothesis h_t: X → {-1,+1} with error

$$\epsilon_t = Pr_{i\sim D_t}[h_t(x_i) \neq y_i] \qquad (6)$$

$$\propto_t = \frac{1}{2}\lambda\mathcal{V}(\frac{1-\epsilon_t}{\epsilon_t}) \qquad (7)$$

Received: 14.06.2021 – Accepted: 16.07.2021

Update:

$$D_{t+1}(i)=\frac{D_t(i)}{Z_t}\times\begin{cases}e^{-\propto_t}, jikah_t(x_i)=y_i\\ e^{\propto_t}, jikah_t(x_i)\neq y_i\end{cases} \quad (8)$$

$$=\frac{D_t(i)exp\left(-\propto_t y_i h_t(x_i)\right)}{Z_t}$$

The normalization factor functionZ_t,D_{t+1} will be distributed by choosing this Z_t and the output of final equation is as below:

$$H(x)=sign(\sum_{t=0}^{T}\propto_t h_t(x)) \quad (9)$$

By adjusting the weak learners and modelling iterations, this approach of Adaptive boosting is processed. For different regions appropriate models of different predictions are accepted and for the heterogeneous data Decision Stump Model is suitable. By using the Decision Stump threshold value T is calculated as:

$$T=data\ (x_1\ or\ x_2)*index\ data+step(x_1\ or\ x_2) \quad (10)$$

Prediction can be obtained by comparing the data value at the index n with the calculated threshold value T. Adaptive boosting uses the prediction value which is learning the data to do iteration of next level. There are 22 thresholds formed for indices - 1 through 5 with a total of 5 data is calculated for the each data of threshold value. Decision stump is used by the predictions which are calculated in two ways in the next step as:

$$Gt(greater\ than)=if\ label\ 1\ than\ (X>T,1,-1) \quad (11)$$

$$Lt(less\ than)=if\ label-1\ than\ (X\leq T,1,-1) \quad (12)$$

4. Results

The adaptive boosting approach of DDoS attack detection performance is evaluated in this section. By using the Mininet (version 2.3.0) tool, effectiveness of adaptive boosting approach of DDoS attack is detected. Virtual network system in the switches, links, controllers, and hosts can be created by the network emulator of Mininet which is open source. For the experiments, the SDN controller POX (version 0.2.0) is used. It is a platform for Open flow and SDN controllers as Python-based open source development. Open flow 1.0 controller is used to implement the adaptive boosting approach. On machine running Ubuntu Linux 16.04, the experiments are executed. By using mesh topology10 Open flow Switches connected with SDN test bed for the experiments. At the beginning same weights (W) are given to the all data. The adaptive boosting approach performance can be evaluated by calculating the information distance of non-attack traffic with the attack traffic and legitimate traffic.

The pre-processed data is used as the learning data which is extracted from the statistical flow by using hping3 tools in the experimental results. The data is divided into two parts as labelled as as 1 for non-attacks and -1 for attacks, so the normalization of data is calculated after pre-processing. With the two predictive labels (1, -1) decision stump method is worked. The reasonable data is represented by the positive symbol (+) and the data of more than 65536 bytes is represented by the negative (-).

Learning data can be taken from the 5 data and it converts into one data as greatest number of errors, namely third data ($\alpha 2 = 0.47$) by applying the equation (11) becomes a reference to look for predictions greater than (Gt). Similarly some errors are present in the data that are fourth data ($\alpha 1 = 1.09$) by applying the equation (12) and a reference to look for predictions less than (Lt). The prediction results of errors Gt and Ltthreshold values can be seen in Table 1.

Table 1 - Prediction Labels of Lt and Gt Functions

Data	Number of packets (x1)	Label	Lt prediction	Gt prediction
1	31	1	-1	1
2	26	1	-1	1
3	19	1	-1	-1
4	292777	-1	-1	1
5	1106210	-1	-1	1

A result from the first and second iterations, first iteration index threshold value (110638.1) for functions less than (Lt), and second iteration index threshold value is 19. By using equations (7) and (8) the prediction of attack and non-attack (h1, h2) on the Lt and Gt functions is calculated. In the first iteration of the function (Lt), for any data (not attack) the prediction value is positive and in the second iteration of the function (Gt), for any data (attack) the prediction is negative. Adaptive boosting algorithm final process is the voting process from a set of iterations and this voting process can be done with equation (5). From the results of iteration, final prediction is processed from the dataset. This process is the final prediction of datasets from the iteration results.

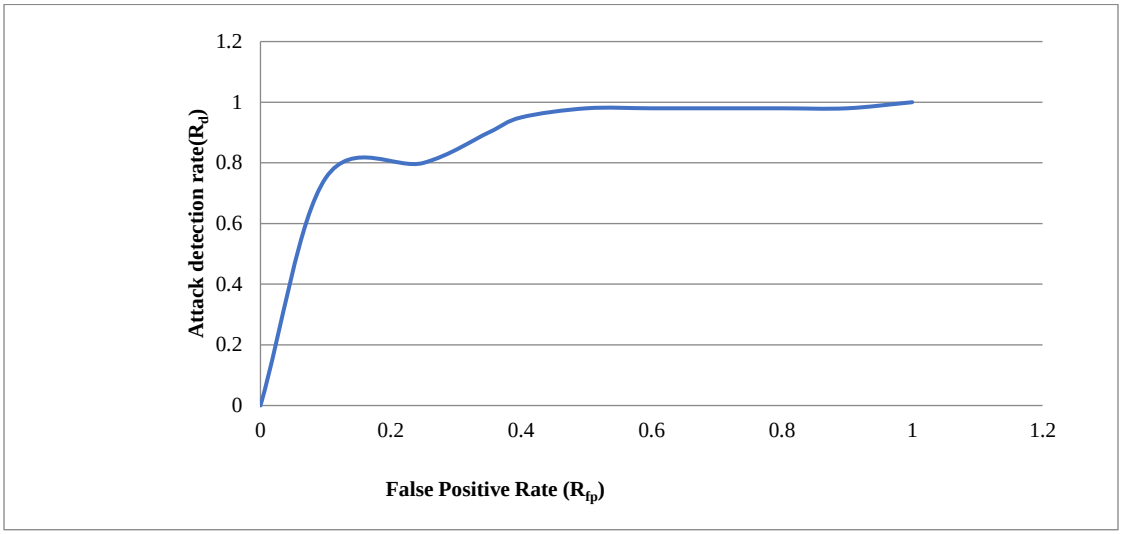
Table 2 - Prediction of Error Greater than (Gt) and Less than (Lt)

Data		1	2	3	4	5
Number of packets (x1)		31	26	19	292777	1106210
Lt	Prediction	-1	-1	-1	-1	-1
	Error	0,1	0,1	0,1	0	0
	Total Error	0,3				
Gt	Prediction	1	1	-1	1	1
	Error	0	0	0,1	0,1	0
	Total Error	0,2				

Received: 14.06.2021 – Accepted: 16.07.2021

The ratio detected attack flows to the processed flows in that which is called attack detection rate. The ratio of attack-free flows in the attack flows to the processed attack-free flows is called as the false positive rate. The performance parameters such as attack detection rate(R_d) and false positive rate $(R_{fp})$are evaluated in this paper to measure the accuracy and effectiveness adaptive boosting as shown in figure (2).

Fig. 2 - Accuracy Measure as Attack Detection Rate Vs False Positive Rate



5. Conclusion

In this paper, cloud computing environment based software defined network system DDoS attacks are noticed and prevented. For attack flow hit count flow rule is more compare to the non-attack flow hit count in the DDoS attacks. So the information distance is calculated to differentiate the attack and non-attack flows by using this approach. The attacks are removed from the system by using the wildcard rule with the removing of flow table switches which hold back the frequent requests from the sources of attacks, when the detection of DDoS attacks. Learning data can be gathered by doing several iterations of adaptive boosting algorithm for the predictive voting process. Small learning data can cause errors in each iteration. This error can be detected and eliminated by doing number of iterations with huge amount of learning data. The DDoS attacks are detected by this approach in SDN based cloud environment effectively and the experimental results are obtained with great accuracy. By using the Adaptive boosting method the attacks are detected and prevented in this SDN's (software-defined networks).

References

Victor C.M. Leung, Zhu Han, Xiuhua Li, Xiaofei Wang, Sangheon Pack, "STCS: Spatial-Temporal Collaborative Sampling in Flow-Aware Software defined Networks", *IEEE Journal on Selected Areas in Comm., 38*(6), 2020

Lefteris Mamatas, Tryfon Theodorou, "A Versatile Out-of-Band Software-Defined Networking Solution for the Internet of Things", *IEEE Access,* 8, 2020

Ismael Amezcua Valdovinos, Jesús Arturo Pérez-Díaz, Dakai Zhu, Kim-Kwang Raymond Choo, "A Flexible SDN-Based Architecture for Identifying and Mitigating Low-Rate DDoS Attacks Using Machine Learning", *IEEE Access,* 2020.

Yuchuan Deng, Yue Pan, Liang Tan, Jing Wu, Hao Jiang, Jianguo Zhou, "A New Framework for DDoS Attack Detection and Defence in SDN Environment", *IEEE Access,* 8, 2020

Trung V. Phan, Tri Gia Nguyen, Nhu-Ngoc Dao, Truong Thu Huong, Nguyen Huu Thanh, Thomas Bauschert, "DeepGuard: Efficient Anomaly Detection in SDN With Fine-Grained Traffic Flow Monitoring", *IEEE Trans. on Network and Service Manag., 17*(3), 2020

Olivier Festor, Hoang-Long Mai, Tan Nguyen, Guillaume Doyen,Rémi Cogranne, Luong Nguyen,Wissam Mallouli, Edgardo Montes De Oca, Moustapha El Aoun, "Reliable Detection of Interest Flooding Attack in Real Deployment of Named Data Networking", *IEEE Trans. on Information Forensics and Sec., 14*(9), 2019

Khushnood Abbas, Shi Dong, Raj Jain, "A Survey on Distributed Denial of Service (DDoS) Attacks in SDN and Cloud Computing Environments", *IEEE Access, 7,* 2019

Truong Thu Huong, Van Tuyen Dang, Pham Ngoc Nam, Nguyen Huu Thanh, Alan Marshall, Nguyen Ngoc Thanh, Steven Furnell, "SDN-Based SYN Proxy—A Solution to Enhance Performance of Attack Mitigation Under TCP SYN Flood" *The Computer Journal, 62*(4), 2019.

Wataru Kurihara, Ryosuke Nagai, Toshio Hirotsu, Shun Higuchi, "Design and Implementation of an OpenFlow-Based TCP SYN Flood Mitigation", *2018 6th IEEE International Conf. on Mobile Cloud Computing, Services, and Engg. (MobileCloud),* 2018.

Reza Javidan, Reza Mohammadi, Mauro Conti, "SLICOTS: An SDN-Based Lightweight Countermeasure for TCP SYN Flooding Attacks", *IEEE Trans. on Network and Service Manag, 14*(2), 2014.

N. Sreenath, K. Geetha, "SYN flooding attack — Identification and analysis", *International Conf. on Information Comm. and Embedded Sys. (ICICES2014),* 2014

Azizah Abdul Manaf, Mohammed A. Saleh, "Optimal specifications for a protective framework against HTTP-based DoS and DDoS attacks", *2014 International Symposium on Biometrics and Secu. Technol. (ISBAST),* 2014.

Chimin Zhou, Min Zhang, Jin Wang, Keping Long, Xiaolong Yang, "HTTP-sCAN: Detecting HTTP-flooding attaCk by modeling multi-features of web browsing behavior from noisy dataset", *2013 19th Asia-Pacific Conf. on Comm. (APCC),* 2013

T. Sivakumar, G. Aghila, Tarun Karnwal, "A comber approach to protect cloud computing against XML DDoS and HTTP DDoS attack", *2012 IEEE Students' Conf. on Electrical, Electronics and Computer Sci.,* Year: 2012.