Sādhanā (2020)45:104
https://doi.org/10.1007/s12046-020-01353-x

SDN-based DDoS Attack Mitigation Scheme using Convolution Recursively Enhanced Self Organizing Maps

PILLUTLA HARIKRISHNA[1,*] and A AMUTHAN[2]

[1]Department of Computer Science and Engineering, Rajeev Gandhi Memorial College of Engineering and Technology, Nandyal, Andhra Pradesh, India
[2]Department of Computer Science and Engineering, Pondicherry Engineering College, Puducherry, India
e-mail: pillutlaharikrishna@yahoo.co.in; amuthan@pec.edu

MS received 6 June 2019; revised 28 November 2019; accepted 31 March 2020

Abstract. In a cloud computing environment, the Distributed Denial of Service (DDoS) attack is considered as the crucial issue that needs to be addressed in ensuring the availability of resources that emerge due to the compromisation of hosts. The process of detecting and preventing DDoS attacks is determined to be predominant when the potential benefits of decoupling data plane from the control plane are facilitated through the Software Defined Networking (SDN) in the cloud environment. The incorporation of SDN in DDoS mitigation also enhances the probability of investigating the data traffic flow using the reactive process of updating forwarding rules, analyzing the network with a global view and centralized control in monitoring for better DDoS mitigation enforcement. In this paper, a Convolution Recursively Enhanced Self Organizing Map and Software Defined Networking-based Mitigation Scheme (CRESOM-SDNMS) is proposed for ensuring the better rate of detection during the process of preventing DDoS attacks in clouds. This proposed CRESOM-SDNMS facilitates a predominant option in resolving the issue of vector quantization with enhanced topology preservation and the superior initialization mechanism during the process of SOM-based categorization of flooded data traffic flows into genuine and malicious. The simulation experiments and results of the proposed CRESOM-SDNMS confirmed a superior classification accuracy of around 21% when compared to the existing systems with minimized False Positive rate of 19% compared to the benchmarked DDoS mitigation schemes of the literature.

Keywords. Software Defined Networking; Convolution Recursively Enhanced Self Organizing Map (CRESOM); DDoS attacks; learning rate.

1. Introduction

In a cloud computing environment, Distributed Denial of Service attack (DDoS) is determined to be the real threat since they are significant in causing maximum disruption that hurdles the availability of the cloud resources [1]. These DDoS attacks are launched in the cloud environment for unintentional and intentional disruption, political and financial gains. The DDoS attack paralyses the cloud services by flooding maximum amount of packets to the servers and links with malicious traffic [2]. This DDoS attack imposes complete service denial and service degradation resulting in maximum losses in the cloud computing environment. The emergence and evolution of Software Defined Networking in the recent decade has widened the option of encountering the DDoS attack in the cloud environment [3]. This possibility in defeating the DDoS attack is possible in the SDN due to its significance in the dynamic enhancement of rules, centralized point of monitoring and global perspective of investigation [4]. SDN is determined as the indispensable entity for cloud and service providers since they have facilitated the possibility of making the network programmable. Further, the storage, bandwidth and computational power have increased considerably as the devices in the applications and data center have continuously risen in its number [5]. This infrastructure has to be maintained, managed and updated in order to achieve the better classification of normal flow from the malicious flow of the data traffic. Thus, the cost and complexity are maximum for maintaining the classical data center framework.

Further, SDN-based machine learning mechanisms using Self Organizing Maps (SOM) were considered to be excellent in appropriate detection and prevention of DDoS attacks in the cloud by investigating the data traffic flow in a reliable manner [6]. This investigation of data traffic using SOM focuses on the process of transforming and visualizing the high dimensional data into a two-

*For correspondence
Published online: 06 May 2020

dimensional grid [7]. The SOM consists of neuron collections that systematically adapt the input through the process of competitive learning in order to create reliable and ordered prototypes for facilitating better categorization of data traffic [8]. The establishment of the created and ordered prototyping characteristics in SOM is mainly used for focusing on the topology preservation of the mapped input data that seems to be highly suitable for cluster analysis [9]. A number of variants of SOM-based DDoS attack mitigation mechanisms were proposed in the literature for enhancing the degree of quantization error such that significant classification accuracy is determined during the process of prevention [10]. However, the rate of learning and topology preservation characteristics remains the major issue that needs to be handled for better discrimination of flows during the process of preventing DDoS attacks in the cloud environment. Thus, a predominant SOM-based DDoS attack mitigation mechanisms are essential for confirming the learning rate and topology preservation characteristics is an excellent way during the process of preventing DDoS attacks in the cloud environment.

In this paper, a Convolution Recursively Enhanced Self Organizing Map and Software Defined Networking-based Mitigation Scheme (CRESOM-SDNMS) is proposed for defeating DDoS attacks in the cloud computing environment. This proposed CRESOM-SDNMS utilizes the benefits of merging and splitting in enhancing the initialization process such that high-density areas of the input space are detected in a reliable and reactive manner. This detection of high- density areas by the proposed CRESOM-SDNMS aids in the potential generation of neurons that results in the new topology in order to limit the adaptation rate of neurons is an excellent way. This proposed CRESOM-SDNMS utilizes the newly created topology for facilitating superior local quantization error compared to the classical SOM-based SDN mitigation mechanisms contributed to the literature. The simulation experiments for the proposed CRESOM-SDNMS have been conducted for determining its predominance under Classification Accuracy, True Positive rate, True Negative rate and False Positive rate evaluated under different data traffic flow rates.

The forthcoming sections of this paper are structured as follows. Section 2 details the comprehensive review on the most recent existing SDN-based DDoS mitigation mechanisms in the cloud environment contributed over the recent years. Section 3 details the implementation steps of the proposed CRESOM-SDNMS with its merits and phenomenal improvement in the process of preventing DDoS attacks through the benefits of CRESOM through SDN. Section 4 highlights the results and investigations of the proposed CRESOM-SDNMS for quantifying its predominance in terms of Classification Accuracy, Precision, Recall and False Positive rate determined under varying intensities of data traffic rates. In section 5 conclusions are provided with major contributions of the proposed CRESOM-SDNMS with the future scope of enhancement.

2. Related work

In this section, the most significant SDN-based DDoS mitigation schemes contributed to a cloud environment over recent years are detailed with their pros and cons.

Initially, a novel mitigation mechanism using a filtering tree is proposed for preventing the HTTP and XML-based DDoS attacks on the cloud computing environment [11]. This filtering tree-based DDoS attack prevention scheme utilized five levels of filters for reducing the impacts of the HTTP and XML-based DDoS attacks. In this filtering tree-based DDoS attack prevention approach, the suspicious packets are analyzed using a puzzle resolver for handling the issues that emerge due to the generated SOAP header-based malicious data packets. This filtering tree scheme first identifies the malicious message initiated IP addresses for sending the puzzle and the sent puzzles are solved for determining the genuine client. An intelligent DDoS mitigation scheme was proposed for combining the network as the portion of the protective framework with the benefits of network traffic filtering, shuffling targets with BGP protocols [12]. This intelligent DDoS mitigation approach explores the various possibilities of flow features that could be phenomenally utilized for detection of malicious data traffic from the monitored data traffic. The classification accuracy of this intelligent DDoS mitigation scheme was proved to be approximately 92%. A high programmable potential-based network monitoring scheme was proposed for encountering and defeating the DDoS attacks in the cloud space by eliminating the limitations of the existing mitigation architecture [13]. This high programmable potential-based network monitoring scheme is capable of facilitating a superior attack detection process that enables adaptive control organization in order to permit rapid and reliable detection of TCP-based attack in the clouds. The results of this high programmable potential-based network monitoring scheme were estimated to be effective and efficient in addressing the security issues introduced by the advent of the network paradigm considered for DDoS detection and prevention.

Then, an OpenFlow-based SDN framework for DDoS attack mitigation was contributed for preventing the issues that emerge due to the emergence of a smooth packet replaying attack [14]. This OpenFlow-based SDN framework controls the intensity of the DDoS attack by utilizing the mutual benefits of data plane and control plane used for investigation. This OpenFlow-based SDN framework is also significant in managing and monitoring the data traffic flows using the advantages of SDN in clouds. A Rapid and accurate SDN-based DDoS attack mitigation scheme was proposed for resolving the issues introduced by the flooding based DDoS attacks using the process of entropy variation [15]. This entropy variation factor utilized in this accurate SDN-based DDoS attack mitigation scheme was determined to be capable of the exact and

distinct categorization of normal and malicious data traffic. The False Positive rate of this entropy variation factor-based DDoS attack mitigation scheme was determined to be reduced by 12% compared to the mitigation frameworks that focused on the OpenFlow-based monitoring process. A novel DDoS detection framework was proposed for preventing compromised clients from accessing the HAP proxy server with the aid of the internet [16]. This detection scheme accepts the packets for monitoring in the first step until the number of connections does not exceed a threshold degree. In the second step, the variable of trust that monotonically increases depending upon the malicious connections is updated in a reactive manner such that the load and scalability are sustained to a maximum level.

A correlation-based DDoS attack mitigation approach was contributed for estimating the association between the arrival rate and inter-arrival time of the cloud clients that are compromised through intentional attacks [17]. The ability of the attacker in exhausting buffer size and bandwidth of the attacked server is explored depending on the unsupervised learning process of self-organizing maps used in the detection. This correlation-based DDoS attack mitigation approach was determined to be capable of investigating the distinct features of the clusters considered for prevention. The benefits of normal probability and mathematical correlation are integrated together for investigating the characteristics of data flow counts such that predominant categorization is facilitated. A Multi-Criteria Security Architecture incorporated SDN-based Mitigation Scheme (MCSA-SDNMS) was contributed for investigating real traffic based on DDoS thresholds and DDoS indicators [18]. This MCSA-SDNMS based mitigation approach utilized an Openflow-based mitigation framework for rapid mitigation such that the traffic volume sends through the interface is phenomenally reduced during its transmission between the control and data plane of SDN. This MCSA-SDNMS based mitigation approach employed the Fuzzy Logic properties for effective detection of DDoS attacks in order to reduce the False Positive rate to a lower degree of 5%. This MCSA-SDNMS based mitigation approach was also determined to be capable of detecting and preventing the intensity to the maximum extent of 97% with reduced attack time. The authors also proposed a Fuzzy Self Organizing Maps enforced SDN-based mitigation scheme (FSOM-SDNMS) for effective mitigation of DDoS attacks in the cloud [19]. This Fuzzy-based SOM detection mechanism is capable of exploring the possible dimensions of the data traffic flow features in a predominant manner. The Classification Accuracy, True Positive rate, True Negative rate of the FSOM-SDNMS approach was estimated to be improved by 19%, 15%, and 23% compared to the MCSA-SDNMS based mitigation approach. Finally, an Improved Self Organizing Maps enforced SDN-based mitigation scheme (ISOM-SDNMS) was contributed for reliable detection of DDoS attacks in the cloud [20]. This Improved Self Organizing Maps utilized the features of a recursive property for the predominant determination of malicious data traffic under the exhaustion of cloud resources. The Classification Accuracy, True Positive rate, True Negative rate of the FSOM-SDNMS approach was also estimated to be improved by 23%, 18% and 20% superior to the existing MCSA-SDNMS approach.

The aforementioned review on the recent SDN-based DDoS mitigation schemes contributed to a cloud environment innovated the possibility of devising a Convolution Recursively Enhanced Self Organizing Map and Software Defined Networking-based Mitigation Scheme for effective Classification Accuracy and reduced False Positive rate.

3. Proposed CRESOM-SDNMS

In this proposed CRESOM-SDNMS, the investigation of data traffic flow starts from the extraction of data flow parameters by the SDN for dynamic updating of rules that aids in the superior detection process. The benefits of the Convolution Recursively Enhanced Self Organizing Map are utilized by the SDN for facilitating excellent analysis of data flow parameters extracted from the data traffic. This process of investigating the data traffic flow is initiated with the collection of 'k' input data vectors $V_{IFP} = [I_1, I_2, \ldots, I_m]$, where $I_i \in R^m$ with $1 \leq i \leq m$, respectively. This input vector $V_{IFP} = [I_1, I_2, \ldots, I_m]$ corresponds to the possible number of data flow parameters considered for investigation. In this investigative process, the SOM is utilized for facilitating an unsupervised learning that consists of two-dimensional neurons associated with specific weights as defined by Kohenen [21]. The assigned weights $\phi = [w_1, w_2, \ldots, w_m]$ for each neuron $n = 1, \ldots, f$ in the SOM network defines the set of mapped data according to Equation (1)

$$S_{MD(n)} = \{I \in V_{IFP} : d(I, w_n) < d(I, w_l)\} \tag{1}$$

Under $l = 1, \ldots, f$ with $l \neq n$ and Euclidean distance $d(I, G)$ computed [26–28] through Equation (2).

$$d(I, G) = |(I - G)| = \sqrt{\left(\sum_{i=1}^{m} I^2 - G^2\right)^2} \tag{2}$$

where $I, G \in R^m$.

Then, each data point $I_i \in R^m$ at a time is given as input to the network in order to compare it with all the possible weight vectors for selecting the nearest weight as the superior matching unit for the considered ith data point. This estimated data point is mapped on to the superior matching neuron $S_{MD(n)} = S_{MD(n)} \cup I_i$ with the objective to solve the problem of vector quantization defined through Equation (3).

$$MinimizeQ = \frac{1}{m}\sum_{i=1}^{nm} \|I_i - w_{SMU}\| \tag{3}$$

where w_{SMU} is the weight corresponding to the superior matching unit for the considered ith data point. Further, the collection of neighborhood weights $N_b = \{w_l : c(b.,l) \le r, l \ne c\}$ around the superior matching unit is updated using the estimated distance existing between the superior matching unit and the closest neuron in the two-dimensional coordinates represented in the topology of the network with 'r' as the pre-assigned radius. Furthermore, $c(b,l) \in N_b$ and the value of $c(b,l)$ lies between 0 and r.

In the SOM, a single data point is chosen in every iteration and its corresponding nearest neuron is determined using Equation (4) with the collection of neighboring neurons (w_n) updated based on Equation (5) with predefined weight vectors, radius, and numbers of maximum iterations with the dimensions of the network.

$$b := \arg\min_{n=1,\ldots f} \left\|I_i - w_j\right\| \tag{4}$$

and

$$w_n = w_n + \kappa(\tau)\beta(\tau)(I_i - w_n) \tag{5}$$

where $\beta(\tau)$ and $\kappa(\tau)$ are the learning rate and neighborhood function at each iteration number τ. This process of updating the collection of neighboring neurons is facilitated until all the possible data points are presented as input to the network. The neighborhood function $\kappa(\tau)$ plays a key role in the utilization of the SOM since it is the exponential function that gets decremented depending on the value of the iteration number τ. Similarly, the learning rate $\beta(\tau)$ also influences the performance of the SOM in a linear manner since it is the decreasing linear function of τ. In addition, the process of initialization and topology preservation plays a major role in ensuring optimal map convergence to a maximum degree [22].

In this proposed CRESOM-SDNMS approach, the mechanism of splitting and merging is included in order to integrate them for facilitating optimal algorithm design during the initialization process of neurons. In this context, the cluster centers to 'm' number of clusters represented by $V_{IFP} = [C_1, C_2, \ldots, C_m]$ is computed for solving the problem described in Equation (6)

$$\text{Minimize } f_{succ} = \sum_{i=1}^{m}\sum_{I\in C_I} \|(I_i - C_i)\|^2 \tag{6}$$

Further, two sets are determined for estimating and categorizing the density of data points around the center point C_i through Equations (7) and (8), respectively.

$$\alpha_x^1(\varepsilon) = \{I \in C_i | d(I, C_i) \le \varepsilon\} \tag{7}$$

and

$$\alpha_y^1(\varepsilon) = \{I \in C_i | \varepsilon < d(I, C_i) \le r_i\} \tag{8}$$

where the value of r_i being the maximum distance estimated between the considered data point of each cluster with its associated cluster center point represented in Equation (9)

$$r_i = \max\{d(I, C_i) | I \in C_i\} \tag{9}$$

Then the two clusters C_X and C_y are considered to be potentially separated when the value of $d(C_x, C_y)$ is greater than $(r_x + r_y)$. It is also obvious that each and every 'mth' number of clusters possess a value of 'ε' that range between 0 and r_i by meeting the constraint $\left|\alpha_x^1(\varepsilon)\right| > \left|\alpha_y^1(\varepsilon)\right|$. Thus, the smaller value of 'ε' ensures and enforces a maximum number of data points to be distributed around the cluster center during the splitting process. Furthermore, the role of 'ε' must be investigated for accurate identification of cluster portions in which most of the data points correlate and reside. The maximum number of points (L_c^i) that is residing within the radius '$\varepsilon_i = \mu r_i$' derived from the collection of 'm' clusters represented through $\psi = [D_1, D_2, \ldots, D_m]$ and its cluster centers $V_{IFP} = [C_1, C_2, \ldots, C_m]$ is determined based on Equation (10).

$$L_c^i = \left|\alpha_c^i(\varepsilon_i)\right| \tag{10}$$

In addition, the identification of data point residing in various portions of the cluster 'C_i' necessitates the transformation such that the data point $(N_v = \delta_1, \delta_2, \ldots, \delta_m) \in R^m$ becomes the cluster center using Equation (11)

$$I^t = I^{t-1} - C_i^t + \delta \text{ with } 1 \le t \le n \tag{11}$$

where the value of δ varying between 0 and 0.1 which needs to be significantly lower for confirming a potential splitting process. After this transformation process, the angle $\theta_{i,I}$ defined between the vectors $I \in C_i$ and N_v is computed based on Equation (12)

$$\theta_{i,I} = \arccos\left(\frac{(I, \widetilde{N_v})}{\delta\sqrt{m\left\|\widetilde{I_j}\right\|}}\right) = \arccos\left(\frac{\sum_{t=1}^{m} I^t}{\sqrt{m\left\|\widetilde{I_j}\right\|}}\right) \tag{12}$$

where $(I, \widetilde{N_v})$ is the term describing the inner product estimated between two vectors I and $\widetilde{N_v}$ respectively. Now, the first set of C_i named $\alpha_x^1(\varepsilon_i)$ is again partitioned based on an angle $\theta_{i,I}$ into two sets as represented in Equations (13) and (14), respectively.

$$\alpha_{x(u)}^1(\varepsilon_i) = \{I \in C_i | \varepsilon_i < d(I, C_i), 0 \le \theta_{i,I} \le \pi/2\} \tag{13}$$

and

$$\alpha_{x(d)}^1(\varepsilon_i) = \{I \in C_i | \varepsilon_i < d(I, C_i), \pi/2 \le \theta_{i,I} \le \pi\} \tag{14}$$

The cardinalities of the aforementioned sets are determined through $L_u^i = \left|\alpha_{x(u)}^1(\varepsilon_i)\right|$ and $L_d^i = \left|\alpha_{x(d)}^1(\varepsilon_i)\right|$ respectively. Moreover, the partitioned sets $\alpha_c^1(\varepsilon_i)$, $\alpha_{x(u)}^1(\varepsilon_i)$ and $\alpha_{x(d)}^1(\varepsilon_i)$ need to satisfy the conditions given in Equations (15), (16) and (17), respectively.

$$L_c^i + L_u^i + L_d^i = |C_i| \tag{15}$$

$$\alpha_c^i(\varepsilon_i) \cup \alpha_{x(u)}^i(\varepsilon_i) \cup \alpha_{x(d)}^i = C_i \tag{16}$$

and

$$\begin{aligned}(\alpha_c^i(\varepsilon_i) \cap \alpha_{x(u)}^i(\varepsilon_i)) &= (\alpha_{x(u)}^i(\varepsilon_i) \cap \alpha_{s(d)}^i(\varepsilon_i)) \\ &= (\alpha_c^i(\varepsilon_i) \cap \alpha_{x(d)}^i(\varepsilon_i)) = \phi\end{aligned} \tag{17}$$

Hence, the process of splitting is applied to the cluster C_i based on the values of L_c^i, L_u^i and L_d^i.

At this juncture, the distribution of data points around the cluster center is determined to be dense when it satisfies the condition highlighted in Equation (18)

$$L_c^i \geq \max[L_u^i, L_d^i] \tag{18}$$

In contrast, the characteristics of the cluster remain unchanged when it satisfies the condition highlighted in Equation (19)

$$L_c^i < \max[L_u^i, L_d^i] \tag{19}$$

Thus, this cluster needs to be partitioned into two cluster subsets as presented in Equations (20) and (21), respectively.

$$\alpha_{(cu)}^1(\varepsilon_i) = \{I \in \alpha_c^i(\varepsilon_i) | d(I, C_i), \leq \varepsilon_i, 0 \leq \theta_{i,I} \leq \pi/2\} \tag{20}$$

and

$$\alpha_{(cd)}^1(\varepsilon_i) = \{I \in \alpha_c^i(\varepsilon_i) | d(I, C_i) \leq \varepsilon_i, \pi/2 \leq \theta_{i,I} \leq \pi\} \tag{21}$$

Furthermore, the original cluster C_i is divided into two new clusters $C_{new(i)}$ and $C_{new1(i)}$ depending on Equations (22) and (23).

$$C_{new(i)} = \{\alpha_{x(u)}^i(\varepsilon_i) \cup \alpha_{(cu)}^i(\varepsilon_i)\} \tag{22}$$

and

$$C_{new1(i)} = \{\alpha_{x(d)}^i(\varepsilon_i) \cup \alpha_{(cd)}^i(\varepsilon_i)\} \tag{23}$$

With centers defined in Equations (24) and (25) respectively.

$$c_{new(i)} = \frac{1}{\left|C_{new(i)}\right|} \sum_{I \in C_{new(i)i}} I \tag{24}$$

and

$$c_{new1(i)} = \frac{1}{\left|C_{new1(i)}\right|} \sum_{I \in C_i new1(i)} I \tag{25}$$

The clusters determined after the process of splitting may not be well separated and hence the clusters that are not well separated are merged into the new cluster. If $C_{q(1)}$ and $C_{q(2)}$ are two clusters that are not well separated, then they are merged based on Equation (26)

$$C_{new}^* = C_{q(1)} \cup C_{q(2)} \tag{26}$$

With center $c_{new}^* = \frac{1}{\left|C_{new}^*\right|} \sum_{I \in C_{new}^*} I$ satisfying the constraint $d(C_{q(1)}, C_{q(2)}) - (r_{q(1)} + r_{q(2)}) < 0$.

In this proposed CRESOM-SDNMS approach, the randomly initialized weights of SOM are modified to prevent the limitation of slower convergence during the process of initialization. The next predominant issue focused on this proposed CRESOM-SDNMS approach is its effectiveness in preserving the topology of the network [23]. In order to preserve the topology of the network, an integer number is defined based on Equation (27)

$$\tilde{r_i} = \left[\frac{d(w_a, w_b)}{(\varepsilon_a + \varepsilon_b)}\right] \tag{27}$$

Thus, the characteristics of CRESOM is enhanced based on splitting and merging-based initializing process and defined integer oriented topology preservation approach. This improvement also has an impact on the learning rate and neighborhood function $\beta(\tau)$ and $\kappa(\tau)$ as defined in Equations (28) and (29), respectively.

$$\beta(\tau) = \lambda\left(\frac{T_p - \tau}{\tau}\right) \tag{28}$$

and

$$\kappa(\tau) = \exp\left(-\frac{r_{i(0)}^2}{2\beta(\tau)^2}\right) \tag{29}$$

Once the learning rate and neighborhood function are enhanced, the convoluted structure of CRESOM is applied for recognizing the malicious flow from the total data traffic flows in the cloud computing environment. This convoluted structure of CRESOM aided in minimizing the function defined in Equation (6) below a threshold as defined in Equation (30)

$$\text{Minimize } f_{succ} = \sum_{i=1}^{m} \sum_{I \in C_I}^{m} \|(I_i - C_i)\|^2 < DTF_{THres} \tag{30}$$

This DTF_{THres} parameter is investigated with the values between 0 and 1 in increments of 0.1 in order to investigate its potential in enforcing the convoluted structure of CRESOM, which concentrates on the reliable detection and prevention process. Hence, the data traffic

flow and its parameters monitored for detecting malicious flow from normal flows are determined in an accurate manner.

4. Simulation results and investigations

In this section, the importance of the proposed CRESOM-SDNMS approach in the DDoS attack prevention in the cloud network is analyzed based on test cases that are distinctly devised with reliable and practical characteristics of monitoring data traffic flow [24, 25]. The superior performance of the proposed CRESOM-SDNMS approach over the compared FSOM-SDNMS, ISOM-SDNMS and MCS-SDNMS schemes are investigated by exploring the five flow features of data traffic as defined in [19]. Similar to the FSOM-SDNMS approach, the proposed CRESOM-SDNMS approach investigates the aforementioned five flow features of data traffic based on the utilization of a flow collection module that is inherently present in the NOX-based cloud data monitoring entity in the SDN. The collected data traffic flow features are investigated through the classification entity of the NOX-based cloud data monitoring component in order to distinguish malicious data traffic from the normal data traffic in the cloud environment during the implementation of the proposed CRESOM-SDNMS approach. Different types of normal and malicious data traffic characteristics are combined for ensuring potential training and testing process in the detection of DDoS attacks in the clouds. The complete data traffic generated and incorporated in the testing process includes 10% of UDP packets, 10% of ICMP packets and 80% of TCP packets.

The Stacheldraht tool is utilized by the proposed CRESOM-SDNMS approach to generate and explore the data traffic essential for preventing DDoS attacks in the cloud environment. For investigating the TCP SYN packets-based flooding attack, 167,832 and 6431 flow count are utilized for testing and training. Similarly, the investigation of UDP packets-based flooding attack utilized 51,227 and 2913 flow count on an average in the process of testing and training. In addition, the ICMP flooding attack is investigated using an average number of 61,564 and 5442 flow count in the process of testing and training. In this experimental analysis of the proposed CRESOM-SDNMS approach, nearly 117,600 flow counts are generated during the training process out of which 51,000 flow counts and 60,000 count flows are aggregated and explored for segregating malicious data traffic from the normal data traffic flows. Moreover, Table 1 presents the number of records in the training and testing dataset for normal and attack traffic. The experimental analysis of the proposed CRESOM-SDNMS approach is conducted based on an improved Intel Quad-Core Xeon processor with 16 GB memory potential in order to facilitate a superior process of training and testing during the enforcement of the implemented approach.

Table 1. Number of records in the training and test dataset for normal and attack traffic.

Traffic type	Training records	Testing Records
Attack		
Normal (N)	49,179	21,076
UDP (U)	2913	51,227
ICMP (I)	5442	61,564
TCP (T)	5442	167,832
TCP & UDP (TU)	4694	2011
UDP & ICMP (UI)	4437	1902
TCP & ICMP (TI)	4739	2031
All (A)	5615	2407

Then, the confusion matrix is used for calculating precision, recall, and F-Measure. This confusion matrix C_{MAT} comprises of $M \times M$ matrix, where M represents the number of classes. It highlights the predicted and actual classes in such a manner that rows and columns are labeled for actual and predicted classes related to all records considered for investigation. Further, the diagonal elements of the matrix portray the True Positive (TP) related to each class, the cumulative sum of the matrix elements in the row except the diagonal element depicts the number of False Positives (FP). The cumulative sum of the matrix elements in the Column except the diagonal element depicts the number of False Positives (FN) (Table 2).

Based on the aforementioned Confusion matrix, the precision, recall, and F-measure with respect to each class 'k' is determined based on Equations (31), (32) and (33).

$$\underset{k}{\mathrm{Pr}} = \frac{TP_k}{TP_k + FP_k} \times 100 \tag{31}$$

$$\mathrm{Re}_k = \frac{TP_k}{TP_k + FN_k} \times 100 \tag{32}$$

$$F\text{-}Measure_k = \frac{2 \times \mathrm{Pr}_k \times \mathrm{Re}_k}{\mathrm{Pr}_k + \mathrm{Re}_k} \times 100 \tag{33}$$

The attacking source topology and the test bed used for implementing the proposed CRESOM-SDNMS approach is explained and highlighted in figures 1 and 2. Figure 1 portrays the complete attacker source topology utilized during the implementation of the proposed CRESOM-SDNMS approach. This utilized attacker source topology consists of four attacker sources in which Attacker 1 utilizes the legitimate IP for attack generation that floods an abnormal number of packets into the cloud environment. Attacker 2 is responsible for generating ICMP, TCP and UDP-based malicious packets for enforcing flooding of data packets in the clouds. Similarly, Attacker 3 generates and introduces IP duplication and DNS spoofing packets

Table 2. The confusion matrix for the eight class classification.

	Actual							
Prediction	N	T	U	I	TU	TI	UI	A
N	99.972	0.981	0.398	0	0.199	0	0.283	0.208
T	0	98.507	0	0	1.492	0.049	0	0.208
U	0	0	99.425	0	4.177	0	1.525	0.332
I	0.009	0	0	100	0	3.939	3.47	0.789
TU	0	0.469	0.177	0	78.319	15.165	0	1.826
TI	0	0	0	0	12.183	76.465	0	5.259
UI	0.009	0	0	0	0	94.585	94.585	7.854
A	0.009	0.043	0	0	3.63	0.158	0.156	83.362

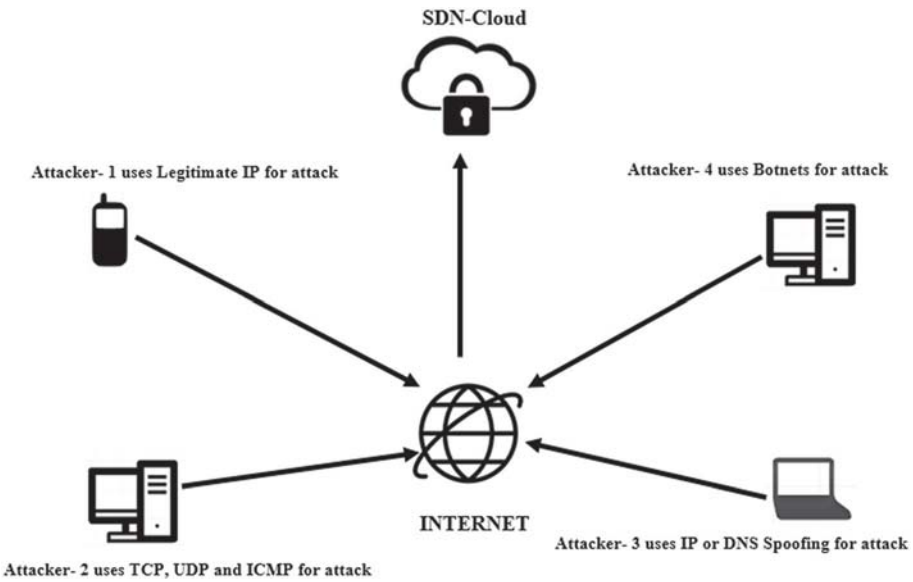

Figure 1. Proposed CRESOM-SDNMS-Attacker source topology for DDoS attack mitigation.

into the cloud environment that increases the rate of data forwarding load over the resources of the clouds. Attacker 4 injects the Botnet kind of malicious behavior in the cloud environment.

Figure 2 highlights the topology of the test bed utilized in the SDN-based cloud network that is considered in the process of deploying the proposed CRESOM-SDNMS approach in order to test its efficacy in detecting and preventing DDoS attacks. This test bed topology comprises of Virtual Firewall Router for integrating the SDN-based cloud infrastructure with the internet. This Virtual Firewall Router existing in the test bed has the potential in allowing the operation of both HTTP and HTTPs protocols. Further, two load balancers are used in the topology for facilitating superior detection and prevention processes. In addition, the two load balancers are incorporated into the application server such that the implicit significance of the VMware is significantly enhanced to the maximum level in the process of mitigation DDoS attacks in the cloud environment.

Initially, the excellence of the proposed CRESOM-SDNMS approach is compared with the existing FSOM-SDNMS, ISOM-SDNMS and MCSA-SDNMS schemes based on the percentage in Classification Accuracy under varying percentages of False Positive rate in the significance of increasing data traffic rates. Figure 3 highlights the predominance of the proposed CRESOM-SDNMS approach evaluated using Classification Accuracy under 100 Mbps of data traffic rate. The results of the proposed CRESOM-SDNMS approach confirmed a superior classification of approximately 12–14%, 15–17% and 19–22% superior to the benchmarked FSOM-SDNMS, ISOM-SDNMS and MCSA-SDNMS schemes. Likewise, figure 4 depicts the superiority of the proposed CRESOM-SDNMS approach to evaluate using Classification Accuracy under

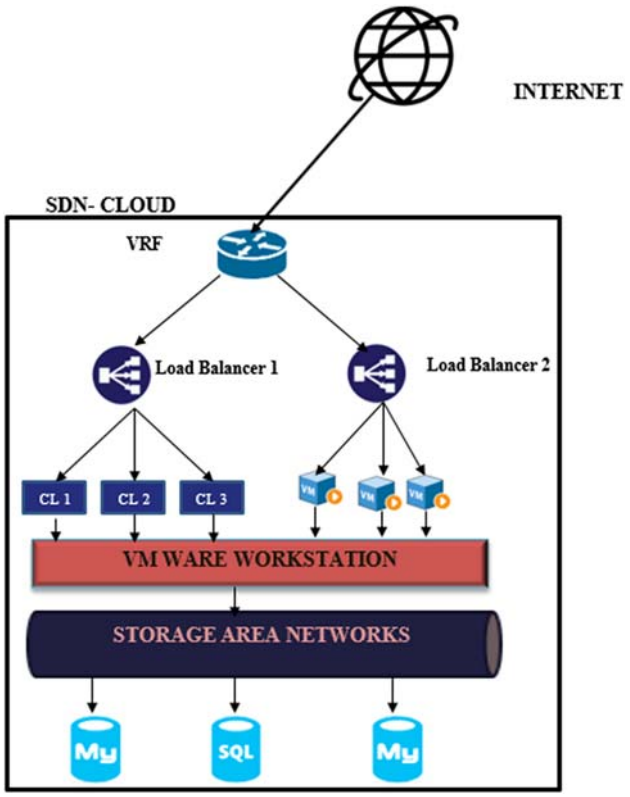

Figure 2. Test-bed topology utilized for implementing the proposed CRESOM-SDNMS.

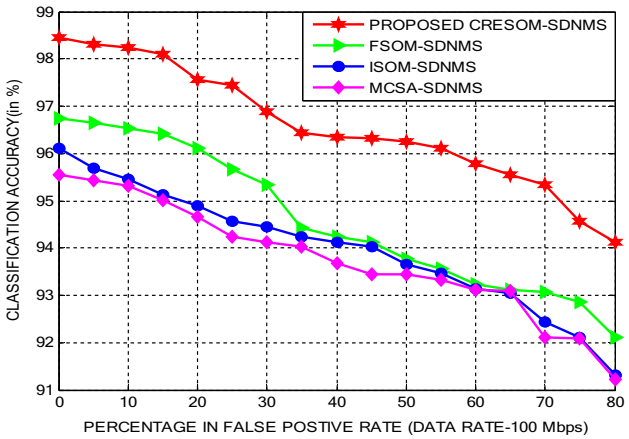

Figure 3. Classification Accuracy-proposed CRESOM-SDNMS-different False Positive rates (Data rate-100 Mbps).

200 Mbps of data traffic rate. The results of the proposed CRESOM-SDNMS approach confirmed a superior classification accuracy rate of approximately 13–15%, 16–18% and 20–24% superior to the baseline DDoS attack mitigation schemes contributed to the cloud environment in the literature. Similarly, figure 5 explores the significance of the proposed CRESOM-SDNMS approach through the Classification Accuracy rate determined under 300 Mbps of data traffic rate. The results of the proposed CRESOM-

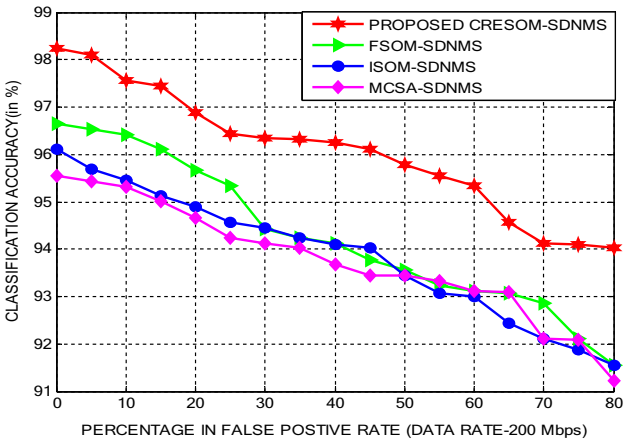

Figure 4. Classification Accuracy-proposed CRESOM-SDNMS-different False Positive rates (Data rate-200 Mbps).

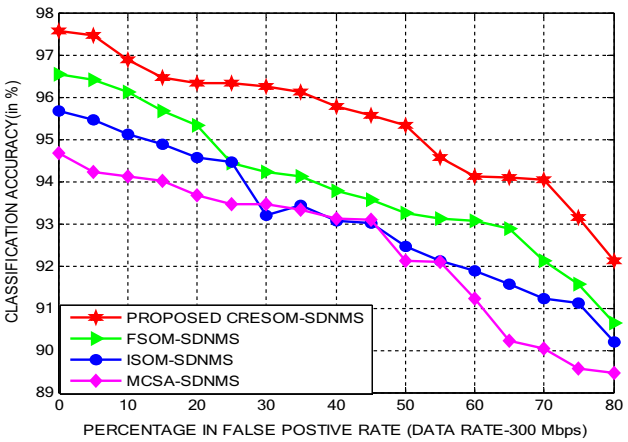

Figure 5. Classification Accuracy-proposed CRESOM-SDNMS-different False Positive rates (Data rate-300 Mbps).

SDNMS approach under 300 Mbps also ensured an outsmarting Classification Accuracy rate of approximately 15–17%, 18–20% and 21–23% superior to the baseline DDoS attack mitigation schemes of the literature.

Further, figures 6 and 7 highlight the performance of the proposed CRESOM-SDNMS approach investigated using Classification Accuracy under data traffic rates of 400 Mbps and 500 Mbps, respectively. The Classification Accuracy of the proposed CRESOM-SDNMS approach under 400 Mbps is nearly 7–9%, 11–14%, and 16–18% compared to the existing benchmarked FSOM-SDNMS, ISOM-SDNMS, and MCSA-SDNMS schemes. The Classification Accuracy of the proposed CRESOM-SDNMS approach under 500 Mbps is also determined to be approximately 6–8%, 10–12%, and 14–16% compared to the existing benchmarked schemes considered for investigation. This phenomenal improvement in the performance of the proposed CRESOM-SDNMS approach investigated using Classification Accuracy is mainly visualized due to the novel initialization that prevents the limitations of

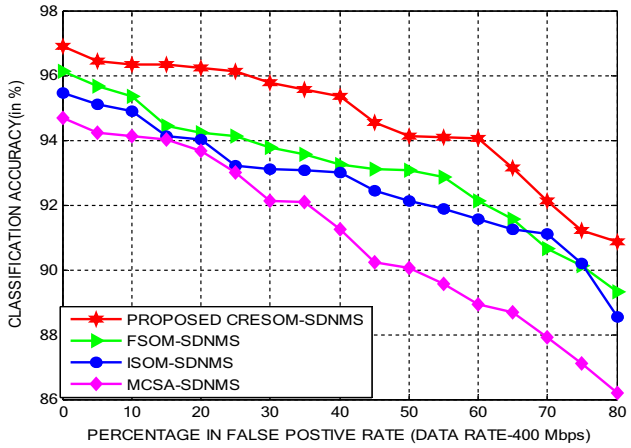

Figure 6. Classification Accuracy-proposed CRESOM-SDNMS-different False Positive rates (Data rate-400 Mbps).

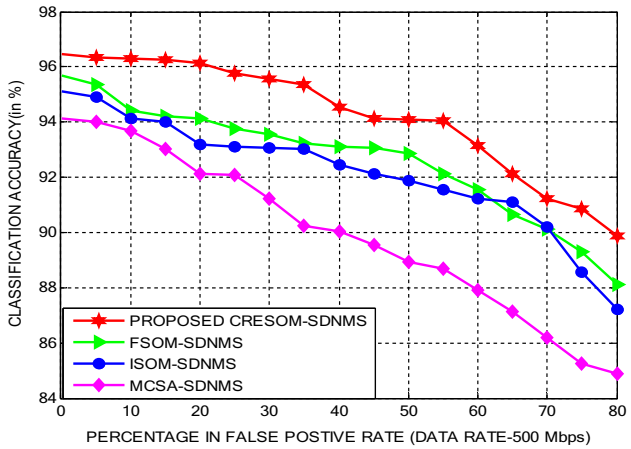

Figure 7. Classification Accuracy-proposed CRESOM-SDNMS-different False Positive rates (Data rate-500 Mbps).

slower convergence and topology preservation in the network.

Furthermore, figures 8, 9 and 10 quantify the False Positive Rate, Precision and Recall value evaluated for the proposed CRESOM-SDNMS approach under the increasing rate of data traffic. The False Positive rate of the proposed CRESOM-SDNMS approach is reduced by 4–6%, 7–9% and 10–12% predominant to the existing benchmarked FSOM-SDNMS, ISOM-SDNMS and MCSA-SDNMS schemes. The minimized False Positive Rate of the proposed CRESOM-SDNMS approach is mainly facilitated due to the utilization of the convoluted characteristics of CRESOM incorporated in the process of distinguishing malicious flow from the normal data flow in the SDN-based cloud computing environment. The Precision of the proposed CRESOM-SDNMS approach is identified to be improved 8–10%, 12–14%, and 16–19% compared to the benchmarked FSOM-SDNMS, ISOM-SDNMS and MCSA-SDNMS schemes. The enhanced precision rate of

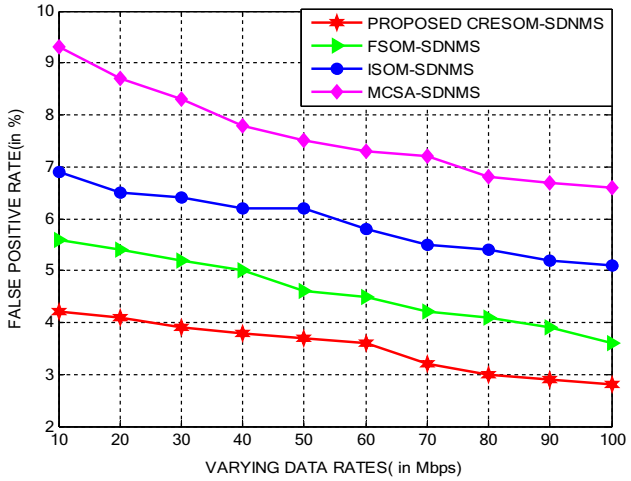

Figure 8. False Positive rate-proposed CRESOM-SDNMS-different data traffic rates.

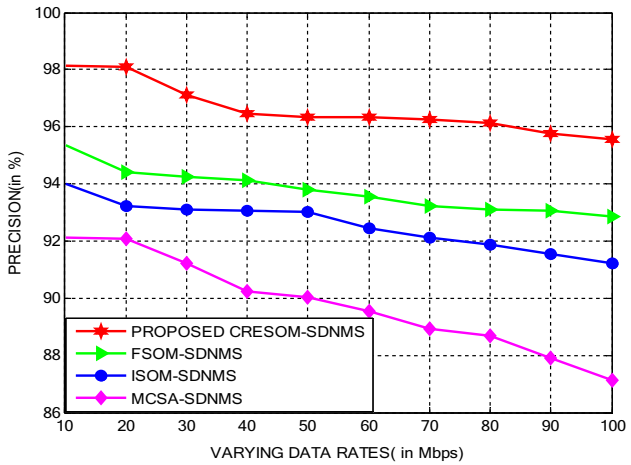

Figure 9. Precision-proposed CRESOM-SDNMS-different data traffic rates.

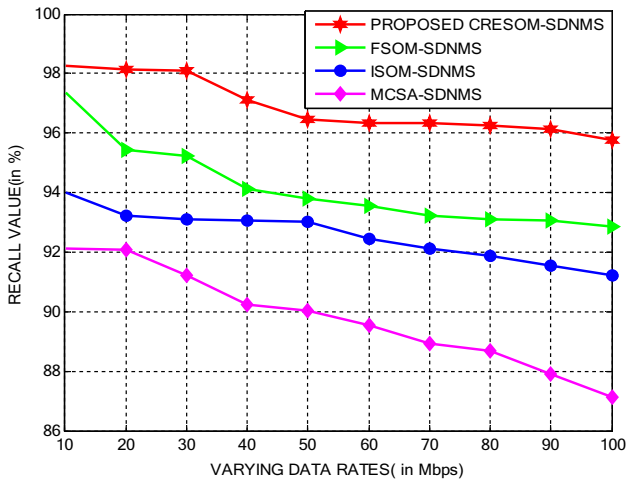

Figure 10. Recall value-proposed CRESOM-SDNMS-different data traffic rates.

the proposed CRESOM-SDNMS approach is mainly due to the incorporation of the merging mechanism that integrates the interrelated data points cluster into a single cluster for preventing slower convergence in the initialization process. In addition, the Recall value of the proposed CRESOM-SDNMS approach was also concluded to be superior over a margin of 10–12%, 14–16%, and 18–21% compared to the benchmarked SDN-based DDoS attack mitigation schemes contributed for improving security in the cloud computing environment.

In addition, figures 11, 12 and 13 show the significance of the proposed CRESOM-SDNMS evaluated using the True Positive rate under the influence of increasing rates of data traffic by varying the degree of False Positive intensity as 10%, 20%, and 30%, respectively. The performance of the proposed CRESOM-SDNMS investigated under False Positive intensity of 10% confirmed that its improvement in the True Positive rate of 9%, 7%, and 4% predominant to the existing FSOM-SDNMS, ISOM-SDNMS and MCSA-SDNMS approaches. Likewise, the performance of the proposed CRESOM-SDNMS quantified under the False Positive intensity of 20% confirmed its enhancement in the True Positive rate of approximately 8%, 6% and 4% excellent to the existing SDN-based DDoS attack mitigation schemes used for comparison. The performance of the proposed CRESOM-SDNMS evaluated under the False Positive intensity of 30% also confirmed the improvement in the True Positive rate of nearly 7%, 5% and 3% excellent to the benchmarked SDN-based DDoS attack mitigation schemes of the literature. This enhancement possibility of the proposed CRESOM-SDNMS is mainly due to the incorporated splitting and merging-based initializing process and defined integer oriented topology preservation approach.

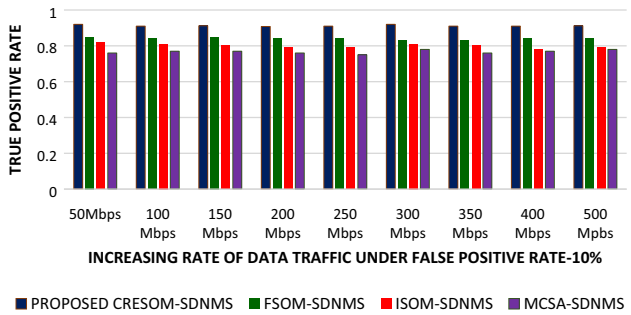

Figure 11. True Positive rate-proposed CRESOM-SDNMS evaluated under different data rates (False Positive rate-10%).

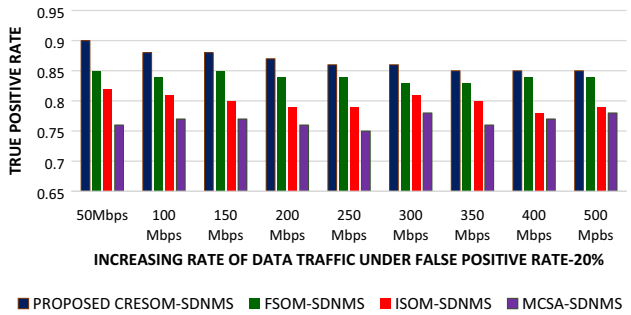

Figure 12. True Positive rate-Proposed CRESOM-SDNMS evaluated under different data rates (False Positive rate-20%).

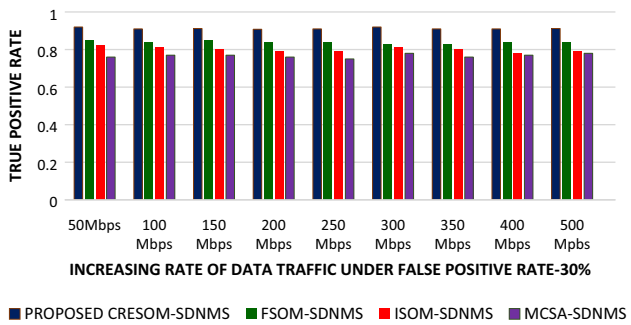

Figure 13. True Positive rate-Proposed CRESOM-SDNMS evaluated under different data rates (False Positive rate-30%).

Table 3. The performance of the proposed model using 2-class classification.

	Actual	
Prediction	Normal	Attack
Normal	99.936	0.342
Attack	0.064	99.658

Finally, the evaluation of the proposed model is also attained through two-class classification based on all categories of DDoS attacks as a comprehensive single attack class in order to compare the potential with the existing works. Table 3 presents the performance of the proposed model using a 2-class classification.

The results presented in table 3, clearly proved that the accuracy of the proposed scheme is 99.86% with F-Measure of 99.87% and 99.79% with respect to normal and attack traffic classes determined from the confusion matrix. Further, the accuracy of the FSOM-SDNMS scheme is 99.78% with F-Measure of 99.75% and 99.74% with respect to normal and attack traffic classes determined from the confusion matrix. The accuracy of the ISOM-SDNMS scheme is 99.72% with F-Measure of 9.65% and 99.71% with respect to normal and attack traffic classes determined from the confusion matrix. In addition, the accuracy of the MCSA-SDNMS scheme is 99.72% with F-Measure of 99.56% and 99.62% with respect to normal and attack traffic classes determined from the confusion matrix.

The results from table 4, proved that the classification time and training time of the proposed are comparatively lower than the benchmarked FSOM-SDNMS, ISOM-SDNMS and MCSA-SDNMS schemes.

Table 4. Classification time and training time of the proposed CRESOM-SDNMS.

Mitigation algorithms	Classification time (s)	Training time (s)
Proposed CRESOM-SDNMS	0.826	518
FSOM-SDNMS	0.872	521
ISOM-SDNMS	0.921	524
MCSA-SDNMS	0.948	532

5. Conclusions

The proposed CRESOM-SDNMS was presented as an excellent DDoS attack mitigation mechanism for cloud environments by incorporating the benefits of CRESOM in SDN for effective discrimination of data traffic flows into genuine and malicious. This proposed CRESOM-SDNMS aided in a better initialization process through splitting and merging methods for ensuring reduced local minimum in the quantization error. This proposed CRESOM-SDNMS utilized the convoluted structure of the recursively enhanced SOM for an appropriate investigation of data traffic flows through the capability of SDN. This proposed CRESOM-SDNMS utilized dynamic updating rules and investigation of data traffic for ensuring reliable detection of the flooding category of DDoS attacks in the cloud. This proposed CRESOM-SDNMS also incorporated a better learning rate and initialization process of better categorization of traffic flows. The results of the proposed CRESOM-SDNMS approach confirmed an excellent improvement in Classification Accuracy by an average of 21% under its investigation with increasing rates of data traffic. The results have also proved that there is a significant improvement of True Positive rate, True Negative rate, Precision, and Recall value for the proposed system when compared with the existing approaches. The False Positive rate of the proposed CRESOM-SDNMS approach was determined to be potentially reduced by 19% compared to the benchmarked SDN-based DDoS attack mitigation approaches considered for analysis. As the future advent of this proposed approach, it is planned to formulate an integrated SDN and Neighborhood function estimated Self Organizing Maps-based DDoS attack mitigation mechanism for facilitating better detection of malicious data traffic flows.

List of symbols

$V_{IFP} = [I_1, I_2, \ldots, I_m]$	The input vector representing data traffic flow
k	Number of input vectors
w_n	Collection of neighboring neurons
$\phi = [w_1, w_2, \ldots, w_m]$	The weights assigned to each neuron
$S_{MD(n)}$	Set of mapped data
$d(I, G)$	Euclidean distance
$I_i \in R^m$	Individual data points
$\widetilde{N_v}$	Neighborhood vector points
$\theta_{i,I}$	Angle estimated between two vectors I and $\widetilde{N_v}$
$(I, \widetilde{N_v})$	Inner product estimated between two vectors I and $\widetilde{N_v}$
Q	Vector quantization function
w_{SMU}	Weight associated with the superior matching unit
N_b	Neighborhood weights
r	Predefined radius
$\beta(\tau)$	Learning rate
$\kappa(\tau)$	Neighborhood function
τ	Iteration number
C_i	Center points
DTF_{THres}	Threshold data traffic flow
$\widetilde{r_i}$	An integer defined for topology preservation
m	Number of clusters
$C_{q(1)}$	Newly constructed first cluster
$C_{q(2)}$	Newly constructed second cluster
$r_{q(1)}$	Radius of newly constructed first cluster
$r_{q(2)}$	Radius of newly constructed second cluster
c^*_{new}	Mean center point of the newly constructed clusters
L^i_c	The cardinality of the set used for partitioning
L^i_u	The upper cardinality number of the set used for partitioning
L^i_d	The lower cardinality number of the set used for partitioning
$C_{new(i)}$	Center point of newly constructed first cluster
$C_{new1(i)}$	Center point of newly constructed second cluster
$\alpha^1_x(\varepsilon)$	First set used for estimating and categorizing density of data points
$\alpha^1_y(\varepsilon)$	Second set used for estimating and categorizing density of data points
ε	Data points correlation parameter
f_{succ}	Successive function of neurons

References

[1] Bhushan K and Gupta B B 2018 Distributed denial of service (DDoS) attack mitigation in software defined network (SDN)-based cloud computing environment. *Journal of Ambient Intelligence and Humanized Computing* 1(1): 56–69

[2] Braga R, Mota E and Passito A 2010 Lightweight DDoS flooding attack detection using NOX/OpenFlow. *IEEE Local Computer Network Conference* 1(1): 22–34

[3] Yan Q, Yu F R, Gong Q and Li J 2016 Software-defined networking (SDN) and distributed denial of service (DDoS) attacks in cloud computing environments: a survey, some research issues, and challenges. *IEEE Communications Surveys & Tutorials* 18(1): 602–622

[4] Bannour F, Souihi S and Mellouk A 2018 Distributed SDN control: survey, taxonomy, and challenges. *IEEE Communications Surveys & Tutorials* 20(1): 333–354

[5] Amin R, Reisslein M and Shah N 2018 Hybrid SDN networks: a survey of existing approaches. *IEEE Communications Surveys & Tutorials* 1(1): 1–21

[6] Tamanna T, Fatema T and Saha R 2017 SDN, A research on SDN assets and tools to defense DDoS attack in cloud computing environment. *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)* 1(2): 78–89

[7] Aamir M and Zaidi M A 2013 A survey on DDoS attack and defense strategies: from traditional schemes to current techniques. *Interdisciplinary Information Sciences* 19(2): 173–200

[8] Jonker M and Sperotto A 2015 Mitigating DDoS attacks using OpenFlow-based software defined networking. *Intelligent Mechanisms for Network Configuration and Security* 1(1): 129–133

[9] Mariette J and Villa-Vialaneix N 2016 Aggregating self-organizing maps with topology preservation. *Advances in Self-Organizing Maps and Learning Vector Quantization* 1(2): 27–37

[10] Cottrell M, Olteanu M, Rossi F and Villa-Vialaneix N 2016 Theoretical and applied aspects of the self-organizing maps. *Advances in Self-Organizing Maps and Learning Vector Quantization* 1(1): 3–26

[11] Karnwal T, Thandapanii S and Gnanasekaran A 2013 A filter tree approach to protect cloud computing against XML DDoS and HTTP DDoS attack. *Advances in Intelligent Systems and Computing* 1(1): 459–469

[12] Cotton M 2017 DDoS attacks: defending cloud environments. *Advances in Intelligent Systems and Computing* 1(1): 907–909

[13] Wang B, Zheng Y, Lou W and Hou Y T 2015 DDoS attack protection in the era of cloud computing and software-defined networking. *Computer Networks* 81(1): 308–319

[14] Khimabhai Y A and Rohokale V 2016 SDN control plane security in cloud computing against DDoS attack. In: *Proceedings of the International Conference on Advances in Information Communication Technology & Computing - AICTC '16* 1(1): 45–56

[15] Mousavi S M and St-Hilaire M 2017 Early detection of DDoS attacks against software defined network controllers. *Journal of Network and Systems Management* 26(3): 573–591

[16] Kilari N and Sridaran R 2017 A novel approach to protect cloud environments against DDOS attacks. *Advances in Intelligent Systems and Computing* 1(1): 515–523

[17] Johnson Singh K and De T 2017 Mathematical modelling of DDoS attack and detection using correlation. *Journal of Cyber Security Technology* 1(3–4): 175–186

[18] Dang-Van T and Truong-Thu H 2017 A multi-criteria based software defined networking system architecture for DDoS-attack mitigation. *REV Journal on Electronics and Communications* 1(1): 45–55

[19] Pillutla H and Arjunan A 2019 Fuzzy self organizing maps-based DDoS mitigation mechanism for software defined networking in cloud computing. *Journal of Ambient Intelligence and Humanized Computing* 10(4): 1547–1559

[20] Zhao C and Liu F 2018 DDoS attack detection based on self-organizing mapping network in software defined networking. *MATEC Web of Conferences 176*(1): 01026

[21] Kohonen T 2001 Self-Organizing Maps. *Springer Series in Information Sciences* 1(1): 78–102

[22] Horio K, Aikawa A and Yamakawa T 2014 Pattern Recognition based on relative position of local features using self-organizing map. *First International Conference on Innovative Computing, Information and Control - Volume I (ICICIC'06)* 1(2): 12–25

[23] Mohebi E and Bagirov A 2014 A convolutional recursive modified Self Organizing Map for handwritten digits recognition. *Neural Networks* 60(1): 104–118

[24] Deepali and Bhushan K 2017 DDoS attack mitigation and resource provisioning in cloud using fog computing. *2017 International Conference On Smart Technologies For Smart Nation (SmartTechCon)* 1(1): 56–65

[25] Challagidad P S and Birje M N 2017 Trust management in cloud computing. *2017 International Conference on Smart Technologies for Smart Nation (SmartTechCon)* 1(1): 45–56

[26] Xu Y and Liu Y 2016 DDoS attack detection under SDN context. *IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications* 1(1): 45–56

[27] Wang W, Ke X and Wang L 2018 A HMM-R approach to detect L-DDoS attack adaptively on SDN controller. *Future Internet* 10(9): 83

[28] Aamir M and Zaidi S M 2019 Clustering based semi-supervised machine learning for DDoS attack classification. *Journal of King Saud University - Computer and Information Sciences* 1(1): 56–67