International Journal of Electrical and Computer Engineering (IJECE)
Vol. 10, No. 2, April 2020, pp. 1599~1611
ISSN: 2088-8708, DOI: 10.11591/ijece.v10i2.pp1599-1611

Preemptive modelling towards classifying vulnerability of DDoS attack in SDN environment

Narender M., Yuvaraju B. N.
Department of Computer Science and Engineering, The National Institute of Engineering Mysuru, India

Article Info

Article history:

Received May 30, 2019
Revised Oct 12, 2019
Accepted Oct 22, 2019

Keywords:

Attack
Distributed denial of service
Intrusion detection system
Machine learning
Software define network

ABSTRACT

Software-Defined Networking (SDN) has become an essential networking concept towards escalating the networking capabilities that are highly demanded future internet system, which is immensely distributed in nature. Owing to the novel concept in the field of network, it is still shrouded with security problems. It is also found that the Distributed Denial-of-Service (DDoS) attack is one of the prominent problems in the SDN environment. After reviewing existing research solutions towards resisting DDoS attack in SDN, it is found that still there are many open-end issues. Therefore, these issues are identified and are addressed in this paper in the form of a preemptive model of security. Different from existing approaches, this model is capable of identifying any malicious activity that leads to a DDoS attack by performing a correct classification of attack strategy using a machine learning approach. The paper also discusses the applicability of best classifiers using machine learning that is effective against DDoS attack.

Corresponding Author:

Narender M.,
Department of Computer Science and Engineering,
The National Institute of Engineering, Mysuru, India.
Email: narender@nie.ac.in

1. INTRODUCTION

Software-Defined Networking (SDN) has been evolved to cater up to the exponentially increasing demands of the clients [1, 2]. These rises of the demands are arising from the usage of future internet architecture, which is very different from existing networking schemes [3]. The capability of the centralized architecture used in the network is furthermore enhanced as it can now segregate the control plane to forwarding plane and can further offer more extensive programming capability to the network [4, 5]. Irrespective of immense advantage, there are various security issues associated with the SDN [6-10]. Apart from the problem associated with the failure of a centralized point in the legacy centralized architecture, the SDN offers a mechanism using its controller system in order to study the security. However, there is a long way to go as SDN is comparatively a new networking concept which will require more prototyping, more exhaustive investigation, and more validation in order to claim its success factor. There is various review work on SDN environment to claim that it is inflicted with potential security problems [11-13]. As SDN uses the software as the core part; therefore, the risk of intrusion is quite more over the network, and various changes of security breaches can be expected. In this regard, Distributed Denial-of-Service (DDoS) acts as a single point of attack which also invites various other security breaches too. This is possible as a DDoS attack paralyze the entire system on the target network by illegitimately capturing the complete availability of such a system. Such attacks are quite easier to be launched in SDN owing to the presence of a centralized controller system that controls the entire network and formulates a decision of a server to be victimized. Normally, when a switch receives a data packet from various traffic flow that is found to be quite unmatched with each other than the data packet is forwarded to the SDN controller. A traffic flow rule is then forwarded by the SDN controller that is constructed on the basis of the application running on the network.

In the case of DDoS attack, attackers use multiple compromised systems for generating massive sizes of data packets. These massive data packets are forwarded to the SDN controller. The number of incoming requests to the server will increase exponentially enough to consume the complete resources that are allocated for a defined communication channel. These resources are in terms of the channel capacity, storage, memory, etc. of the regular user. One of the vulnerable resources within the switch is the forwarding table capacity. This information can be compromised by the attacker in DDoS. All sorts of illegitimate resources initiate towards saturating the tablespace when a massive quantity of the spoofed packets is accumulated in the network switch. Finally, the DDoS attack completely cripples the communication link existing between the SDN controller and the network switch. This causes the final step of DDoS attack due to heavier congestion of such illegitimate traffic.

Although there is various existing research work, they all have its benefits as well as a constraint which let the DDoS problem still unsolved in present time. Therefore, this paper presents a solution to such a problem by developing a mathematical model towards resisting DDoS problem in SDN. The paper illustrates the model with a machine learning approach to offer more gain in preventive measures towards DDoS. The organization of the paper is as follows: Section1.1 discusses the related work of existing approaches toward resisting DDoS attack on SDN environment. Section 1.2 briefs the research problems addressed in current work while discussion of the adopted research methodology is carried out in Section 2. Algorithm discussion is carried out in Section 3 while result discussion is carried out in Section 4. Finally, the conclusive remarks are made in Section 5.

This section presents a briefing of research approaches towards resisting DDoS attack. Most recently, Phan and Park [14] have used *a self-organizing map* and *support vector machine* in order to introduce a node classification concept for enhancing the rate of DDoS attack detection over SDN cloud. Another recent work of Zhu *et al*., [15] has showcased the usage of *encryption* approach for safeguarding privacy while *the KNN* approach is used for enhancing the efficiency of detection over SDN. The work of Yu *et al*., [16] has presented detection as well as prevention mechanism by constructing a network table using *OpenFlow protocol*. The study has also used *a support vector machine* for training. A *correlation-based approach* has been presented by Zheng *et al*., [17] by using network switches designed using *a commercial off-the-shelf*. Existing system also reports of using entropy-based approaches towards DDoS detection scheme. The work of Kalkan*et al*., [18] has used a *statistical solution* that enables the controller to generate a dynamic rule. Simpson *et al*., [19] have used a *local filtering scheme* with a more precise specification of cookies as well as proofs. Different technologies associated with machine learning are also found in existing approaches. Assis *et al*. [20] have jointly used *game theory, digital signature*, and *fuzzy logic* for developing an autonomous DDoS attack detection. A unique implementation scheme has been presented by Yan *et al*., [21] where a *scheduling* approach has used as security approach. The approach extracts different time variant information of the schedules allocated on the basis of the attacks of DDoS intensity. Similar work is also carried out by Lim *et al*., [22] where the scheduling approach has been promoted by the author as a security solution. The approach is capable of confining the DDoS attack from infecting the SDN controller. Lawal and Nuray [23] have presented a technique where the packet flows are investigated, and it leads to the generation of the rule set to the SDN controller over a virtual machine. The work of Lima and Fernandez [24] has presented a *statistical* technique as well as the entropy factor of traffic. Usage of *network function virtualization* has also been proven to assists in boosting security system against DDoS attack. The work of Zhou and Guo [25] has introduced a framework for mitigating DDoS attack by monitoring the network traffic to filter out any *anomaly-based behavior*. A prevention-based mechanism was carried out by Dong *et al*., [26] where the traffic flow is classified for the purpose of *sequential probability decision* making. Another study carried out by Hong *et al*., [27] have addressed the problem of slow DDoS, which invites all form of lethal threats towards the web server. The solution presented by the author is mainly a defense system to identify and resist the slow mode attack pattern of DDoS. Usage of network function virtualization is also seen in the work of Hyun *et al*., [28] for resisting security breaches. The authors have developed a *threshold-based mechanism* which unblocks the regular data packets and block the suspicious data packets for being circulated in the network. The study towards the resistance measures against DDoS attacks on the SDN system is carried out by Kalkan*et al*., [29]. According to the study, it can be said that research work carried out on or before the year 2017 has been associated with diversified features where very less number of work is focussed combined on detection and mitigation. The study carried out by Kuerban*et al*., [30] has emphasized on *FlowSec*, which is an approach for resisting DoS threat on the controller. Study considering resource as the factor to identify DDoS threat is carried out by Xu and Liu [31] that is carried out while monitoring the neighboring nodes. Dao *et al*., [32] have presented a mechanism to surveil the data packets generated from the switches followed by conditional checking if the packets match with the security logic presented. Blocking and unblocking of packets are followed by this as a solution towards DDoS attack. Saravanan*et al*., [33] have demonstrated a novel

structure to reduce DDoS vulnerabities into the cloud environment.Sofia*et al*., [34] have presented constructive router helped congestion control for the Software defined networking (SDN).Ali *et al*., [35] have focused on Load balancing in to the data center for Software defined networking. Therefore, there are various studies carried out towards resisting DDoS attack over the SDN environment. Various approaches presented by the researchers do have their advantages as well as limitation. The next section highlights the research problems that are constructed after reviewing the existing system.

Following are the research gaps that have been evolved while studying existing methodologies as well as results obtained by them:

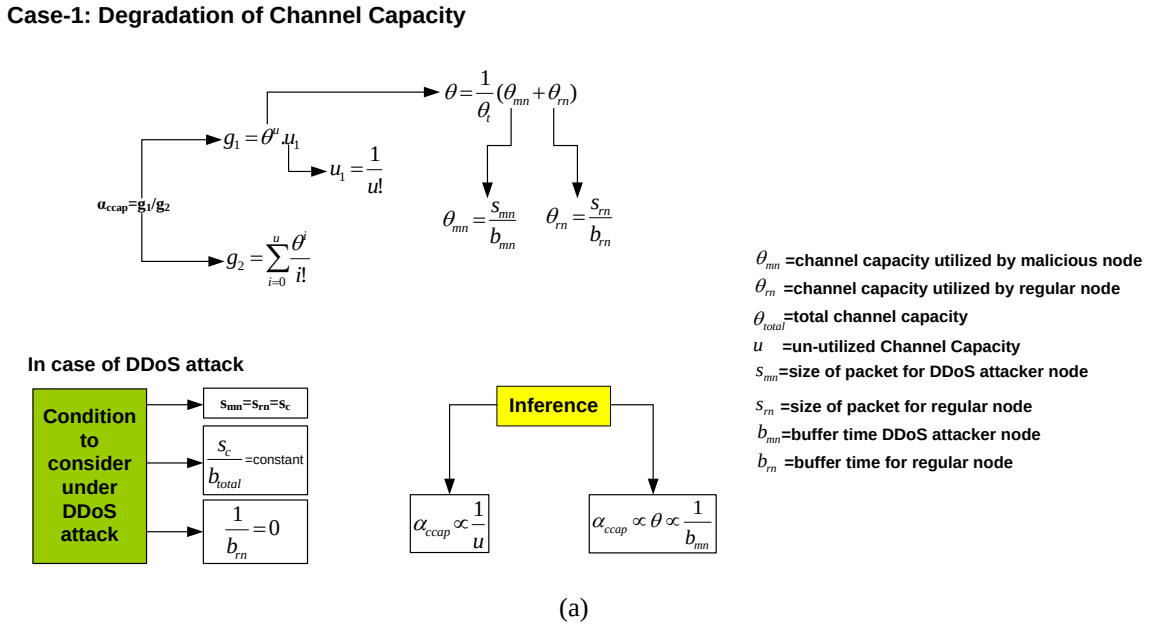
a. *Few Standard Research Work*: There is no doubt of the presence of research-based solution towards DDoS attack in the context of SDN [36, 37]. However, a closer look will show that there is no significant novelty in the approaches.
b. *Defined Execution*: Existing solutions towards resisting DDoS is only successful in the predefined information about the attack strategy is known to the controller. Such approaches are not applicable in order to stop the dynamic attack scenario, which is complex in a distributed system.
c. *Less scope of Machine Learning*: Machine learning is tremendously gaining success in solving a complex problem. This aspect cannot be seen in reality when it comes to developing resistance against DDoS attack. It is because the attack model is quite complex, which further complicates the training process.
d. *Lack of adversarial modeling*: 100% of all the research work is carried out in the existing system by assuming the adversarial model but not modeling it. The existing solution is successful only when eavesdrop and the one-by-one system is compromised until it leads to the victim server. However, an adversary can adopt any number of strategy to initiate or launch an attack, while this is the instant of time which requires to be identified. Unfortunately, there is no existing system which addressed this problem.

Hence, there is a need of *"a computational model which can offer preemptive measure for early identification following by resisting DDoS attack in SDN environment*." The next section briefs of the research methodology adopted as a solution to this problem.

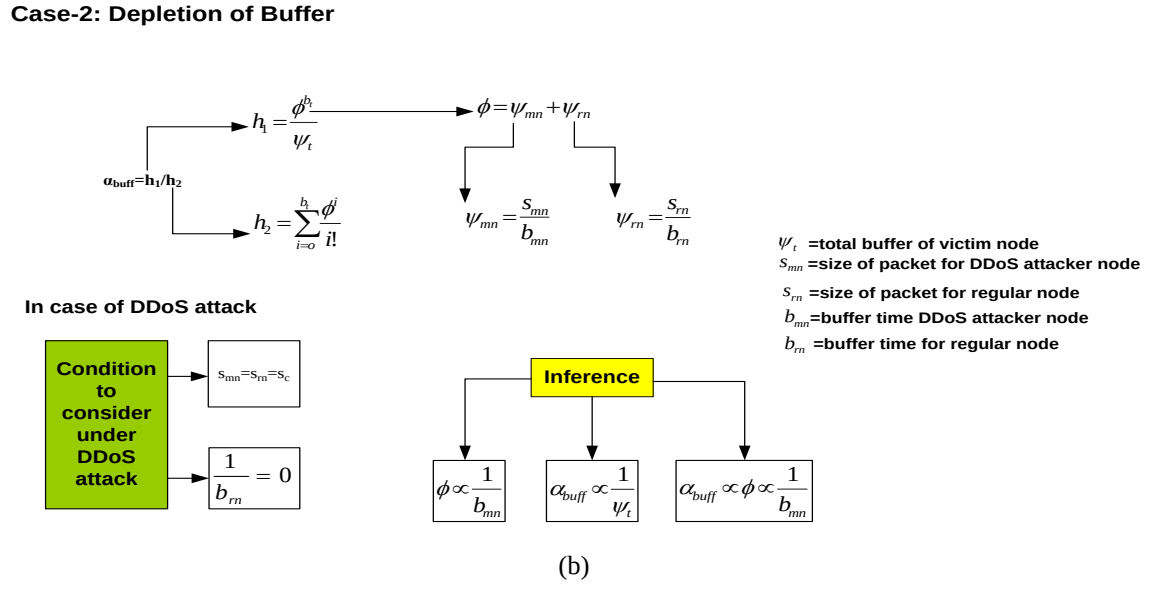
2. RESEARCH METHODOLOGY

The proposed study is designed and developed using analytical research methodology where both empirical and simplified mathematical approach has been used. The complete planning of implementation of the proposed study is based on the fact that it is quite challenging to differentiate a regular node or malicious node presence in SDN and differentiating them is a quite challenging task. Therefore, the proposed system formulates concrete characteristics of DDoS attack in SDN where channel capacities, as well as a buffer of the nodes in SDN, are targeted to be exhausted. Hence, the proposed study develops a simple form of mathematical modeling where two case studies are selected viz. a) case-1 represents the primary vulnerable scenario of DDoS attack that significantly depletes the channel capacity as shown in Figure 1(a) and (b) case-2 represents the vulnerable secondary scenario of DDoS which over-saturate the memory system of the nodes present in SDN as shown in Figure 1(b). The complete implementation of the proposed system is developed using three different stages. The *first* stage of implementation is about formulating a simplified expression for computing the cumulative intrusion that considers both the case studies. The outcome of this study assists in obtaining a qualitative value of the DDoS attack on the SDN node. The complete formulation of the cumulative intrusion is carried out on the basis of probability theory for comprehensive inference about the attack classification. The *second* stage of implementation is about constructing mathematical expression along with condition towards measuring the probability of attack considering both the case of DDoS attack scenario. In the case of degradation of channel capacity, the prominent parameters that are affected during the attacks are the size of data forwarded for malicious as well as a regular node as they tend to remain the same. It is because the adversary doesn't offer any chances to get them identified by different size of packet from a regular node during flooding the server with iterative flooding packets from a different direction. Another significant and dependable parameter considered is buffer time of both regular and malicious node, which is computed as the time of two packets forwarded by the same node. This parameter of buffer time is utilized in assessing the respective buffer rate of the malicious as well as a regular node on the perspective of the victim node. Apart from this, un-utilized channel capacity consumption is another dependable attribute that is used in mathematical modeling in order to compute the criticality score of DDoS attack. In the case of declination of the buffer of the victim node, the dependable parameter remains the same except they are chooses on the context of the buffer value of the respective nodes. The total size of the buffer of the victim node is another essential parameter that is used in formulation towards assessing the criticality score. The third stage of implementation is associated with applying the machine learning approach in order to optimize the classification performance of the proposed system with respect to the DDoS attack scenario in

SDN. The implication of machine learning is another mechanism of better Classification of the complex variables and establishes a robust relationship among the variables. Therefore, the proposed system performs a group-wise analysis of the outcomes of cumulative intrusion value, followed by assessing the similarity score of these parameters in the group. This significantly assists in better classification process towards studying intrusion mechanism in DDoS. Apart from this, the proposed system contributes towards a qualitative investigation towards exploring the impact of machine learning towards the classification performance of the DDoS attack in SDN. Different from any existing system, the proposed system offers a comprehensive insight to the potential capabilities of machine learning approaches. Figure 1 and Figure 2 discuss the operation flow of the proposed system with respect to case study and adoption of machine learning mechanism.



(a)



(b)

Figure 1. Extracting inference for DDoS attack in proposed system,
(a) Mathematical computational flow for case-1, (b) Mathematical computational flow for case-2

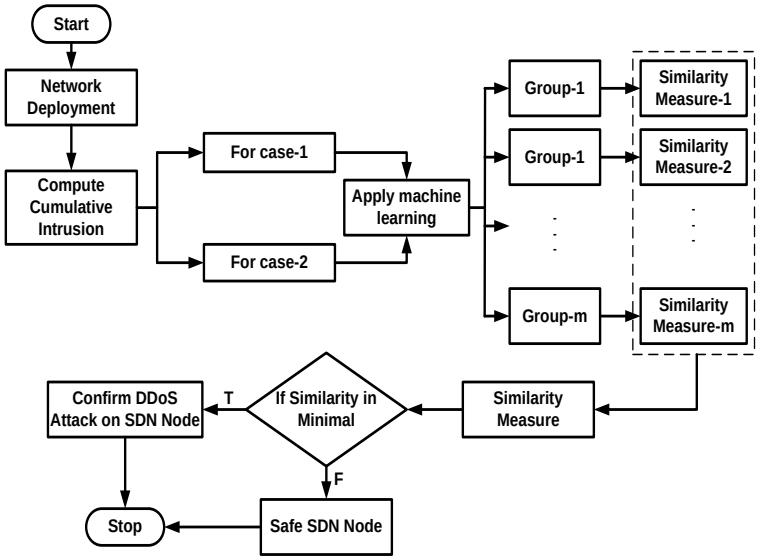


Figure 2. Process flow of DDoS attack confirmation using machine learning

3. ALGORITHM DESIGN

The prime motive of the proposed system is to carry out a discrete analysis of the DDoS attack and its adversarial impact over the nodes with SDN. The complete design strategy of the proposed system is carried out on the basis that a DDoS attack will lead to serious degradation of the services relayed from the server to be victimized. Developing this scenario of the attacker is essential as it will assist in carrying out a qualitative analysis of the impact of attack strategies on the server. This network performance will also directly contribute towards better identification process of DDoS attack. From theoretical understanding, it is well known that when the legitimate clients are denied access over the server, there is a state of tremendous resource depletion of the server as a part of attack strategy. Such forms of depletion of resources could be carried out over channel capacity or node buffer (victim). The process is further optimized by allowing a machine learning algorithm to participate in the process of classifying. Therefore, the proposed system constructs an algorithm that offers unique identification of the type of attacks on the basis of the adversarial behavior.

3.1. Algorithm for classifying DDoS attack

The core idea of the development of this algorithm is that the DDoS attack leads to declination of resources either from channel capacity viewpoint or from buffer viewpoint. The algorithm takes the input of n (number of nodes) and th_1/th_2 (threshold) which after processing yields and outcome of flag message of C_{int} (Cumulative Intrusion). The complete illustration of the proposed algorithm is based on two cases m=2 (Line-4) for computing utilized channel capacity and buffer. The first switch case (Line-5-7) is associated with depletion of channel capacity using various attributes of it. The proposed algorithm formulates an expression called a degree of channel capacity consumption $\alpha_{ccap,}$

$$\alpha_{ccap}=g_1/g_2 \tag{1}$$

In the above expression, the variable g_1 and g_2 represent favorable channel capacity depleted and total channel capacity. The computation of g_1 and g_2 is carried out as follows,

$$g_1 = \theta^u . u_1 \text{ and } g_2 = \sum_{i=0}^{u} \frac{\theta^i}{i!} \tag{2}$$

In the expression, the variable θ can be further broken down to following empirical expression,

$$\theta = \frac{1}{\theta_t}(\theta_{mn} + \theta_{rn}) \tag{3}$$

In the expression, the variable θ_{mn}, θ_{rn}, and θ_t represent channel capacity used by the malicious node, regular node, and total channel capacity, respectively. The computation of θ_{mn}can be carried out by dividing the size of data forwarded for the malicious node (s_{mn}) to buffer time of malicious node (b_{mn}). The computation of θ_{rn}can be carried out by dividing the size of data forwarded for the regular node (s_{rn}) to buffer time of malicious node (b_{rn}). The proposed system introduces another practical concept that no malicious node will initiate an attack in such a way that behavior looks suspicious and the size of the packet is one such factor which gives a rough idea about it. Therefore, the proposed system considers that size of the data packet forwarded during invoking DDoS attack in SDN node to be equivalent to that of regular communication environment. This strategy is adopted by an adversary so that they don't get themselves caught. Therefore, in such case, expression (3) will look like following,

$$\theta = \frac{1}{\theta_t}(\theta_{mn} + \theta_{rn}) = \frac{1}{\theta_t}\left(\frac{s_{mn}}{b_{mn}} + \frac{s_{rn}}{b_{rn}}\right)$$

In case of DDoS attack $s_{mn}=s_{rn}=s_c$,

$$\theta = \frac{1}{\theta_t}\left(\frac{s_c}{b_{mn}} + \frac{s_c}{b_{rn}}\right) = \frac{s_c}{\theta_t}\left(\frac{1}{b_{mn}} + \frac{1}{b_{rn}}\right) = \chi.\frac{1}{b_{mn}} \quad (4)$$

It is because, during the state of DDoS attack, s_c/θ_t is constant say χ and $1/b_n$ are nearly equivalent to zero. From expression (4), it can be stated that θ is inversely proportional to the buffer time of the malicious node, i.e., b_{mn}. The proposed study considers buffer time to the duration of time between two data packets relayed from a single SDN node. Therefore, with the help of expression (1) and (4), the proposed system can infer that,

$$\alpha_{ccap} \propto \frac{1}{u} \text{ and } \alpha_{ccap} \propto \theta \propto \frac{1}{b_{mn}} \quad (5)$$

The inference of the above expression (5) is that during DDoS attack on SDN node, degree of bandwidth consumption is directly proportional to the variable θ and inversely proportion to buffer time of the malicious node. The algorithmic steps are as follows:

Algorithm for classifying DDoS attack

Input: *n* (number of nodes), th_1/th_2 (threshold)
Output: Flag message of C_{int} (Cumulative Intrusion)
Start
1. init*n*, th_1, th_2
2. **For** i=1: *n*
3. $C_{int} \rightarrow 1-(P_1.P_2)$
4. **Switch** (m)
5. **case**-1:
6. $\alpha_{ccap}=g_1/g_2$
7. *break*;
8. **case**-2:
9. $\alpha_{buff}=h_1/h_2$
10. *break*
11. **End**
12. **If**Bt<th_1&&N_c>th_2
13. *Flag*C_{int} as LARGE
14. **If**Bt>th_1&&N_c<th_2
15. *Flag*C_{int} as LESS
16. **End**
End

The next part of the algorithm (Line-7-10) is about the second case of vulnerability considering the effect of buffer declination during a DDoS attack in SDN node. In this case, the proposed system computes the degree of buffer declination as,

$$\alpha_{buff}=h_1/h_2 \quad (6)$$

In the above expression, the variable h_1 and h_2 are formulated nearly in a similar way just like it was done for expression (1) as below,

$$h_1=\frac{\phi^{b_t}}{\psi_t} \text{ and } h_2=\sum_{i=o}^{b_t}\frac{\phi^i}{i!} \quad (7)$$

The similar scenario used during expression (4) is also used here, which can be re-written for buffer declination as,

$$\phi=\psi_{mn}+\psi_{rn}=\frac{s_{mn}}{b_{mn}}+\frac{s_{rn}}{b_{rn}} \quad (8)$$

As during DDoS attack on SDN node,

$$\phi=s_b(\frac{1}{b_{mn}}+\frac{1}{b_{rn}})=\frac{s_b}{b_{mn}} \quad (9)$$

With similar logic that $s_{mn}=s_{rn}=s_b$ and $1/b_{rn}$ is equivalent to zero during attack event, the proposed logic now derives down to the following inference,

$$\alpha_{buff}\propto\frac{1}{\psi_t} \text{ and } \alpha_{buff}\propto\phi\propto\frac{1}{b_{mn}} \quad (10)$$

Using all the above expression, the proposed system finally makes the following inference as shown in Table1 for performing analysis of DDoS attack.

Table 1. Analysis of attack scenario

Attack Scene	b_{mn}	Communication channel	The buffer of the victim node, *b*	Cumulative intrusion C_{int}
1	✡	✡	✡	✡✡✡
2	✡✡✡	✡✡✡	✡✡✡	✡
3	✡	✡✡✡	✡✡✡	Node successfully captured

(Legend: Less: ✡, High: ✡✡✡)

3.2. Optimizing classification using machine learning

This prior operational step assists in classifying the type of DDoS attack in SDN node with respect to the intensity of the attack. If the attack intensity is high, it flags as large alarm and vice versa. However, the best result of Classification can be only obtained if the machine learning approach is used. The basic logic of the precise classification process is that – if the degree of channel capacity consumption as well as buffer utilization is found prominent. However, such associated data are so diversified that if they are not grouped, then it will be challenging to improve the classification accuracy of it. Hence, the proposed system applies a machine learning approach for improving classification performance. The primary reason behind this adoption is that machine learning classifications don't have any dependency on experts or specialized skill set, and it is a highly automated mechanism for performing training/learning operation for a given set of data. Apart from this aspect, a machine learning approach is also cost-effective as well as highly flexible mechanism with more scope of applicability on any form of the task associated with learning operation. Hence, applying machine learning algorithm could further offer better classification performance for DDoS attack in SDN node. The steps of the algorithm are as:

Algorithm for classification optimization

Input: *n* (nodes), *thr* (threshold)
Output: Flag of DDoS attack / Safe
Start
1. check*n* (mn = rn)
2. retainB_t
3. **For** i=1:n
4. a=group n (id, B_t)
5. Apply f(x)→a;
6. sim→[sim_1, sim_2, ….sim_m]
7. **For** j=1:m
8. **If** sim<thr
9. Flag DDoS Attack
10. **Else**
11. Flag Safe
12. **End**
13. **End**
End

The initial process of the algorithm is to formulate a group of all the nodes *n* that has a nearly similar identity of regular *rn* and malicious node *mn* (Line-1). All the similar forms of the SDN nodes are kept in the same group, which is carried out on the basis of the buffer time B_t (Line-4). An explicit function *f*(x) is developed, which is applied to the extracted group of similar SDN nodes (Line-5). This operation results in a matrix *sim* which retains all the similarity score of obtained groups respectively, e.g., sim_1, sim_2,…sim_m, where *m* is the number of groups formulated (Line-6). For all the number of formulated groups (Line-8), the proposed system compares the cumulative similarity score *sim* with a threshold score *thr* (Line-8). If the similarity score is found to be less than the threshold than it represents a case where the presence of malicious event can be ascertained. In such a case, the algorithm flags, the outcome as confirmed DDoS attack (Line-9) or else it is safer (Line-11). One of the interesting points to be noted in this algorithm is that it just contributes to performing precise classification operation by performing training on the data obtained from the prior algorithm. As this algorithm is a continuation of the prior algorithm, therefore, it adds more precision in the process of classification operation. Irrespective of the dynamic behavior of the regular and malicious SDN nodes, the proposed machine learning approach can perform a precise classification of criticality score of the attacker influence as well as confirmation of the DDoS attack. The next section discusses the result obtained.

4. RESULT ANALYSIS

This section discusses the results being obtained after implementing the algorithm. Scripted in MATLAB, it should be noted that the proposed system aims for quantifying the Classification of DDoS attack performance. The classification outcomes of the first algorithm are about the intensity of attack while the classification outcome of the second algorithm is about further confirming the presence of a DDoS attack. Therefore, the performance parameter considered for result analysis is mainly accuracy. As the proposed system implements machine learning approach for improving and qualitative analysis of classification performance; therefore, prediction time and training time is also assessed. Further discussion of result analysis is as:

4.1. Analysis of strategy

The final objective of the proposed system is to ensure that a node running over SDN should be able to perform safer communication and could potentially resist DDoS attack, even before it takes place. Majority of the existing approaches towards resisting DDoS attack is only applicable when any even of DDoS takes place. It is because such methods fail to identify the malicious behavior that an adversary could adopt to initiate a DDoS attack. Therefore, the proposed security process is assumed to be executed in the control layer of SDN standard architecture, which resides in between the application layer and infrastructure layer. The input to this control layer could be network/communication-related information, whereas the output of the controller will be a set of rules that will be followed by all the nodes running within it. The proposed system constructs the rule on the basis of the information and inference discussed in Table 1. According to the proposed rule set, following are the distinct characteristics under DDoS attack e.g.

a) cumulative depletion of overall resources of the vulnerable node (victim node) is inversely proportional to the rate of buffer time of the adversary node, b) the probability of the resource consumption increases of the buffer time is smaller, c) if the size of buffer is maintained higher than resource depletion could be controlled, iv) the presence of many numbers of un-utilized channel capacity can also significantly control the channel capacity drainage for the vulnerable node. Referring to Table 1 and the first classification algorithm, there are three inference extracted. The *first inference* (1st attack scenario) obtained is - it can be stated that cumulative intrusion C_{int} value has the highest impact on nodes when a vulnerable node is found with a) smaller buffer time, b) less availability of un-utilized channel capacity, c) smaller size of the buffer (or memory). The *second inference* (2nd attack scenario) obtained is – the cumulative intrusion C_{int} value is less for the vulnerable node with high buffer time and higher size of the buffer (or memory) as well as a higher number of un-utilized channel capacity. The *third inference* (3rd attack scenario) obtained is – the probability value of the cumulative intrusion is said to be compromised if buffer time is found less along with a large number of availability of un-utilized channel capacity and higher size of the buffer. Therefore, un-utilized channel capacity and size of the buffer is the dominant predictor of DDoS attack and hence can also be used for resisting such intrusion.

4.2. About dataset

The analysis of DDoS attack will need a real form of DDoS attack to be involved in the experiment. However, it is challenging to test the proposed algorithm waiting for a real-time scenario of a DDoS attack. In order to carry out the implementation of the proposed algorithm for Classification of DDoS attack, the system demands networking data when subjected to a DDoS attack. At the same time, the proposed system will also need ground truth data (data without DDoS attack, i.e., safe network) for carrying out validation. Therefore, the proposed system uses the dataset associated with real-time traces of cyber risk [38]. This dataset is collected by using standard probing mechanism over the internet (IPv4 and IPv6) with a standard collection of more than 60 countries with more than 60 megabytes of data collected per day. Availability of cyber data as shown in Figure 3.

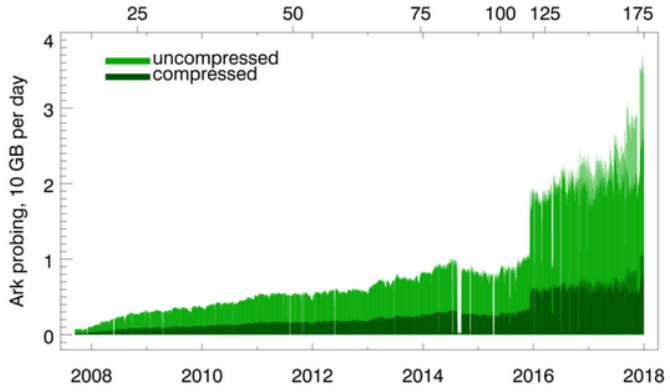


Figure 3. Availability of cyber data [35]

4.3. Core study findings

The analysis of the proposed system is carried out considering 200 unique nodes with a distinct identity. The dataset is curated with an equal distribution of 100 regular nodes and 100 malicious nodes. Only information related to buffer time, data arrival frequencies (sec), and throughput are considered for the complete dataset which is split into two parts, i.e., first part represents for normal communication scenario while second part represents for DDoS attack scenario. The analysis is carried out considering the intruding time to be 30 seconds for all the nodes. As the proposed system claims of using machine learning for optimizing the classification performance, it is essential to perform a comparative analysis to showcase the better version of machine learning. The comparative analysis is carried out using almost all types of existing classification methods using machine learning mechanism viz. i) logistic regression, ensemble learning, K-Nearest Neighbor (KNN), Support Vector Machine (SVM), and tree-based classification approach.

The numerical outcomes obtained from the proposed study are showcased in graphical form Figure 4 to Figure 5. The outcomes are quite interesting and exhibit multiple sets of inference. From Figure 4, it can be just said that logistic regression and tree-based offers the highest accuracy of more

than 97% in comparison to other existing classification technique. The reason behind this are:-logistic regression is best suited for solving binary classification problem (although, we have four classification factors – two factors (large and smaller intensity of DDoS attack) obtained from first algorithm and two (presence or absence of DDoS attack) from the second algorithm), it always does binary classification stage-wise. The classification outcome exhibited by the logistic regression model is potentially enriched with information as compared to other classifiers. This aspect can be seen as Figure 4 shows that logistic regression offers a maximum number of observation to be predicted in contrast to other classifiers however it is carried by consuming more training time to confirm faster convergence as shown in Figure 5. The next analysis is for decision tree-based approach, which has considered averaging three tree-based methods (i.e., coarse, medium, and fine tree). The accuracy of the decision tree (98.6%) is slightly higher than the logistic regression approach.

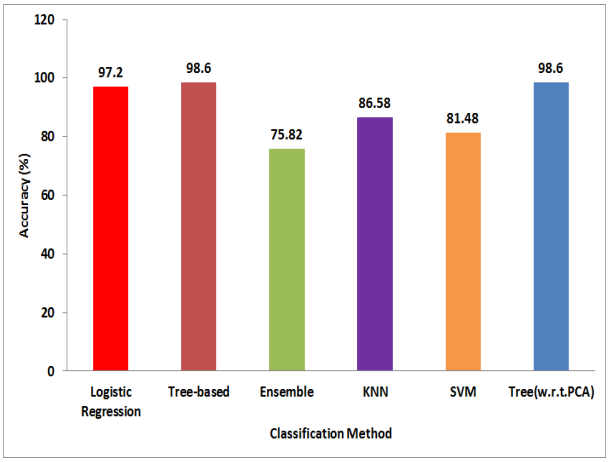


Figure 4. Comparative analysis of accuracy

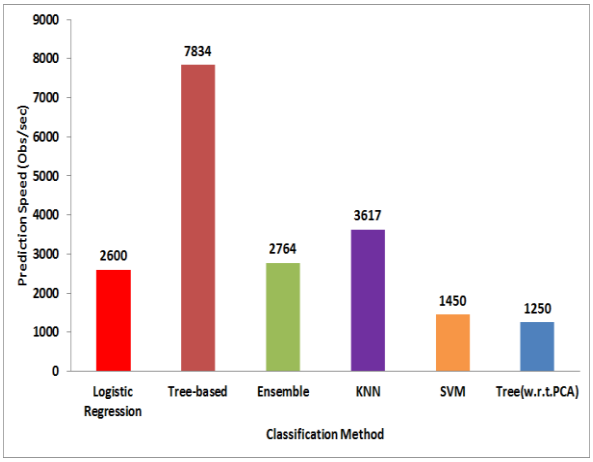


Figure 5. Comparative analysis of prediction speed

The prime reason for this is decision tree based approach assists in efficient feature extraction which is absolutely not impacted by any non-linear relationship (if they arrive owing to DDoS attack) because of this reason, the training time as shwn in Figure 6 is least for decision tree and its observation capability as shown in Figure 4 is also quite high compared to other classifiers. The performance of accuracy, as well as observation reading capability of the ensemble classifier, is found to be lower than other classifiers. The biggest challenge in the ensemble mechanism is that it cannot give unbiased outcome for one model, which is dominant and hence could result in interoperability problem. The fourth classifier is KNN approach whose accuracy is found to be higher than the ensemble approach and less than logistic regression and decision tree-based classifier. The proposed system has been analyzed with different variants of the KNN approach (e.g., fine, medium, coarse, cosine, cubic, and weighted). The fifth and sixth classified, i.e., SVM and Tree considering with/without principal component analysis, doesn't exhibit better outcomes. Therefore, the core study findings are as: a) although logistic regression approach offers higher accuracy as well as it cannot capture maximum observation as seen by Tree-based classifier. Moreover, its training time is significantly higher because of extensive predictive outcomes. Hence, *the logistic regression approach is well suited for offline analysis and not for online analysis*. b) A decision tree is the best suited classified for proposed research problem of Classification of DDoS attack. It is well suited for both offline as well as online analysis as it offers highly reduced training time exhibiting the least computational complexity involved with more capability to read maximum observation. It is also observed that a fine-structured decision tree is more recommended for slightly higher accuracy as compared to the course and medium decision tree. c) The nature of the ensemble approach is quite computationally expensive and hence not recommended for online analysis of DDoS attack. iv) KNN algorithm is another suitable algorithm for analyzing the attack in offline mode. It is because offline mode will generate massive traces of data, which is the highly suitable scenario for KNN classification. Reduced training time is another advantage of KNN. However, owing to maximum computation cost and dependency, only weighted KNN approach is recommended for adoption.

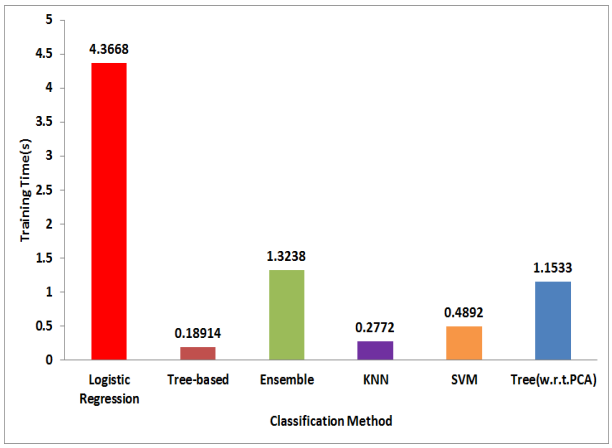


Figure 6. Comparative analysis of training speed

5. CONCLUSION

The role of SDN is essential in a distributed networking system in present time. With a distributed mechanism, the communication system is also becoming challenging to be protected. While the existing system is about protecting the network from DDoS attack once it has already been initiated, but proposed system offers a preemptive solution towards identifying the malicious behavior of the DDoS attack and stop any form of the malicious event even before the attack is launched. The complete modeling has been carried out considering a resource that area targeted to be depleted during DDoS attack, i.e., channel capacity and buffer. The complete implementation is carried out considering two implementation stage-where the first stage of implementation focuses on classifying the intensity of DDoS attack, and the second part is about classifying the presence or absence of DDoS attack. It will mean that machine learning has been used to precisely confirm the presence of attack so that if there is any form of outliers than it is compensated using a machine learning approach. The proposed system is also investigated using a different type of machine learning approaches where logistic regression, decision tree, and KNN algorithms are found to offer better performance in DDoS attack classification. Our future work will be further optimizing the preventive measure. The future research direction can considered other security parameters for performance analysis.

REFERENCES

[1] Dijiang Huang, AnkurChowdhary, Sandeep Pisharody, Software-Defined Networking, and Security: From Theory to Practice, CRC Press, 2018.
[2] Dumka, Ankur, Innovations in Software-Defined Networking and Network Functions Virtualization, IGI Global, 2018.
[3] BhawanaRudra, Flexible Network Architectures Security: Principles and Issues, CRC Press, 2018.
[4] SiamakAzodolmolky, Software Defined Networking with OpenFlow, Packt Publishing Ltd, 2013.
[5] RahamatullahKhondoker, SDN and NFV Security: Security Analysis of Software-Defined Networking and Network Function Virtualization, Springer, 2018.
[6] Dacier, Marc & Koenig, Hartmut&Cwalinski, Radoslaw&Kargl, Frank & Dietrich, Sven. (2017). Security Challenges and Opportunities of Software-Defined Networking. IEEE Security and Privacy Magazine. 15. 98-102. 10.1109/MSP.2017.46.
[7] M. Dabbagh, B. Hamdaoui, M. Guizani and A. Rayes, "Software-defined networking security: pros and cons," in *IEEE Communications Magazine*, vol. 53, no. 6, pp. 73-79, June 2015.
[8] A. Feghali, R. Kilany and M. Chamoun, "SDN security problems and solutions analysis," *2015 International Conference on Protocol Engineering (ICPE) and International Conference on New Technologies of Distributed Systems (NTDS)*, Paris, pp. 1-5, 2015.
[9] M. C. Dacier, H. König, R. Cwalinski, F. Kargl and S. Dietrich, "Security Challenges and Opportunities of Software-Defined Networking," in *IEEE Security & Privacy*, vol. 15, no. 2, pp. 96-100, March-April 2017.
[10] V. Patil, C. Patil and R. N. Awale, "Security challenges in software defined network and their solutions," *2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, Delhi, pp. 1-5, 2017.
[11] S. Scott-Hayward, S. Natarajan and S. Sezer, "A Survey of Security in Software Defined Networks," in *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 623-654, Firstquarter 2016.
[12] Heng Zhang, ZhipingCai, Qiang Liu, Qingjun Xiao, Yangyang Li, and ChakFoneCheang,"A Survey on Security-Aware Measurement in SDN," *HIndawi-Security and Communication Networks*, Volume 2018.

[13] Ahmad, Ijaz&Namal, Suneth&Ylianttila, Mika &Gurtov, Andrei. (2015). Security in Software Defined Networks: A Survey, IEEE Communications Surveys & Tutorials. 17, 4, 1-1, 10.1109/COMST.2015.2474118.
[14] T. V. Phan and M. Park, "Efficient Distributed Denial-of-Service Attack Defense in SDN-Based Cloud," in *IEEE Access*, vol. 7, pp. 18701-18714, 2019.
[15] L. Zhu, X. Tang, M. Shen, X. Du and M. Guizani, "Privacy-Preserving DDoS Attack Detection Using Cross-Domain Traffic in Software Defined Networks," in *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 3, pp. 628-643, March 2018.
[16] Y. Yu, L. Guo, Y. Liu, J. Zheng and Y. Zong, "An Efficient SDN-Based DDoS Attack Detection and Rapid Response Platform in Vehicular Networks," in *IEEE Access*, vol. 6, pp. 44570-44579, 2018.
[17] J. Zheng, Q. Li, G. Gu, J. Cao, D. K. Y. Yau and J. Wu, "Real-time DDoS Defense Using COTS SDN Switches via Adaptive Correlation Analysis," in *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 7, pp. 1838-1853, July 2018.
[18] K. Kalkan, L. Altay, G. Gür and F. Alagöz, "JESS: Joint Entropy-Based DDoS Defense Scheme in SDN," in *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 10, pp. 2358-2372, Oct. 2018.
[19] S. Simpson, S. N. Shirazi, A. Marnerides, S. Jouet, D. Pezaros and D. Hutchison, "An Inter-Domain Collaboration Scheme to Remedy DDoS Attacks in Computer Networks," in *IEEE Transactions on Network and Service Management*, vol. 15, no. 3, pp. 879-893, Sept. 2018.
[20] M. V. O. De Assis, A. H. Hamamoto, T. Abrão and M. L. Proença, "A Game Theoretical Based System Using Holt-Winters and Genetic Algorithm With Fuzzy Logic for DoS/DDoS Mitigation on SDN Networks," in *IEEE Access*, vol. 5, pp. 9485-9496, 2017.
[21] Q. Yan, Q. Gong and F. R. Yu, "Effective software-defined networking controller scheduling method to mitigate DDoS attacks," in *Electronics Letters*, vol. 53, no. 7, pp. 469-471, 30 3 2017.
[22] S. Lim, S. Yang, Y. Kim, S. Yang and H. Kim, "Controller scheduling for continued SDN operation under DDoS attacks," in *Electronics Letters*, vol. 51, no. 16, pp. 1259-1261, 6 8 2015.
[23] B. H. Lawal and A. T. Nuray, "Real-time detection and mitigation of distributed denial of service (DDoS) attacks in software defined networking (SDN)," *2018 26th Signal Processing and Communications Applications Conference (SIU)*, Izmir, pp. 1-4, 2018.
[24] N. A. S. Lima and M. P. Fernandez, "Towards an Efficient DDoS Detection Scheme for Software-Defined Networks," in *IEEE Latin America Transactions*, vol. 16, no. 8, pp. 2296-2301, Aug. 2018.
[25] L. Zhou and H. Guo, "Applying NFV/SDN in mitigating DDoS attacks," *TENCON 2017 - 2017 IEEE Region 10 Conference*, Penang, pp. 2061-2066, 2017.
[26] P. Dong, X. Du, H. Zhang and T. Xu, "A detection method for a novel DDoS attack against SDN controllers by vast new low-traffic flows," *2016 IEEE International Conference on Communications (ICC)*, Kuala Lumpur, pp. 1-6, 2016.
[27] K. Hong, Y. Kim, H. Choi and J. Park, "SDN-Assisted Slow HTTP DDoS Attack Defense Method," in *IEEE Communications Letters*, vol. 22, no. 4, pp. 688-691, April 2018.
[28] D. Hyun, J. Kim, D. Hong and J. P. Jeong, "SDN-based network security functions for effective DDoS attack mitigation," *2017 International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju, pp. 834-839, 2017.
[29] K. Kalkan, G. Gur and F. Alagoz, "Defense Mechanisms against DDoS Attacks in SDN Environment," in *IEEE Communications Magazine*, vol. 55, no. 9, pp. 175-179, Sept. 2017.
[30] M. Kuerban, Y. Tian, Q. Yang, Y. Jia, B. Huebert and D. Poss, "FlowSec: DOS Attack Mitigation Strategy on SDN Controller," *2016 IEEE International Conference on Networking, Architecture and Storage (NAS)*, Long Beach, CA, pp. 1-2, 2016.
[31] Y. Xu and Y. Liu, "DDoS attack detection under SDN context," *IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications*, San Francisco, CA, pp. 1-9, 2016.
[32] Nhu-Ngoc Dao, Junho Park, Minho Park and Sungrae Cho, "A feasible method to combat against DDoS attack in SDN network," *2015 International Conference on Information Networking (ICOIN)*, Cambodia, 2015, pp. 309-311
[33] A. Saravanan, S. SathyaBama, SeifedineKadry, Lakshmana Kumar Ramasamy, "A new framework to alleviate DDoS vulnerabilities in cloud computing", *International Journal of Electrical and Computer Engineering (IJECE)*, Vol. 9, No. 5, pp. 4163-4175, 2019.
[34] Sofia NaningHertiana, AditKurniawan, HendrawanHendrawan, UdjiannaSekteriaPasaribu, "Effective Router Assisted Congestion Control for SDN", *International Journal of Electrical and Computer Engineering (IJECE)*, Vol. 8, No. 6, pp. 4467-4476, 2018.
[35] Tariq Emad Ali, Ameer Hussein Morad, Mohammed A. Abdala, "Load Balance in Data Center SDN Networks", *International Journal of Electrical and Computer Engineering (IJECE)*, Vol. 8, No. 5, pp. 3084-3091, 2018.
[36] Q. Yan, F. R. Yu, Q. Gong and J. Li, "Software-Defined Networking (SDN) and Distributed Denial of Service (DDoS) Attacks in Cloud Computing Environments: A Survey, Some Research Issues, and Challenges," in *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 602-622, Firstquarter 2016.
[37] Q. Yan, F. R. Yu, Q. Gong and J. Li, "Software-Defined Networking (SDN) and Distributed Denial of Service (DDoS) Attacks in Cloud Computing Environments: A Survey, Some Research Issues, and Challenges," in *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 602-622, Firstquarter 2016.
[38] "Caida", https://www.caida.org/data/, Retrieved on 30-05-2019.

BIOGRAPHIES OF AUTHORS

Mr. Narender M is assistant professor in the department of computer science & engineering, National Institute of Engineering, Mysuru, Karnataka, India. He has received his M. Tech. and B.E. from Visvesvaraya Technological University (VTU), Belagavi, Karnataka, India. He is pursuing his PhD from Visvesvaraya Technological University (VTU), Belagavi, Karnataka, India. His teaching and research interests are in the field of Software-defined networking, Cognitive radio networks and Data mining. He has around total teaching experience of 5 years.

Dr. Yuvaraju B. N.is a professor in the department of computer science & engineering, National Institute of Engineering, Mysuru, Karnataka, India. He has received his B.E from Mysore University, Karnataka, India and M.Tech as well as PhD from Visvesvaraya Technological University (VTU), Belagavi, Karnataka, India. He has around total teaching experience of 23 years. His teaching and research interests are in the field of Computer Networks, Data Mining and Storage Area Networks.