A Framework for Mitigating DDoS Attacks in Named Data Networking Using Machine Learning Technique

[1]Hashmat Fida[2]Tayeed Ahmad Hajam,[3]Dr. K. Suresh Joseph
[1]MTech N&IE ,[2]Research Scholar ,[3]Assistant Professor
[1]Department of Computer Sciences,
[1]Pondicherry University, Pondicherry, India
[2] Department of Computer Applications
[2]SBBS University Punjab ,India
[3]Department of Computer Sciences,
[3]Pondicherry University, Pondicherry, India.

Abstract: As internet architecture has undergone a drastic change due to the emerging technologies. It is very difficult to implement the improved and updated security models and make the network more secure. The future technologies must offer additional and better security over current networks. Attacks such as alteration, spoofing and DDoS are unconditionally directed in NDN architecture. Moreover, some specific DDoS attacks are only meant for NDN. DDoS attacks in NDN are generally incorporated with the help of interest flooding. A distributed denial-of-service (DDoS) attack is an attempt, rather a destructive aim to distort the usual traffic of a selected server, network or service by overpowering the target or its corresponding systems or infrastructure with a spate of internet traffic. This attack has various disadvantages over the security of the network. This paper thrives a framework which uses a machine learning technique to mitigate DDoS. Firstly, we choose a training data set to describe the characteristics of the traffic. Then as per our survey, we found TF-IDF technique to transform the data set into matric form. To reduce the dimensions of the features and time complexity, we discovered that LDA algorithm is more feasible for our work. Also, RNN Algorithm was found useful and suitable for our work as a detection model. LDA reduces the prediction time significantly. For real data set, LDA can achieve a remarkable performance in terms of accuracy, sensitivity etc. compared to other detection models of DDoS.

IndexTerms **- DDoS, LDA, Interest Flooding, Alteration, Spoofing, RNN, PIT, FIB, DGEA.**

I. INTRODUCTION

As NDN is a step towards future internet architecture, it should offer better security than current IP network.NDN architecture is also surrounded by various attacks such as spoofing, alteration and various types of DDoS attacks. In NDN, DDoS attacks are incorporated with the help of interest flooding. In this attack, the attacker makes use of bunch of compromised hosts which we called as zombies. DDoS attacks can be implemented in various ways. Firstly, the attacker can try to overflow the Pending Interest Tables (PITs) in the routers to stop them from managing the genuine requests. The attacker uses a large number of compromised hosts to initiate a stream of interest packets. Secondly, the DDoS attacks can be dynamically initiated on the basis of requirement. In this attack, all requests are routed to content producers. This consumes bandwidth of the router's PIT resulting the denial of genuine and valid content requests [1].

The elemental requirements of Named Data Networking are security and privacy. In IP networks, DDoS utilize the resources of network and thus manipulating the services for genuine users. Such attacks can lead to most difficult security problems. Therefore, NDN architecture is susceptible to DDoS attacks and requires full observation. NDN offers improved security and privacy then current IP networks[2].

This paper studies an idea about the prevention of DDoS attacks using machine learning technique. The NDN data set containing URL's and domains is compressed and converted into text form using radix-tree master to reduce the size as shown in figure 1. The data set will be then transformed into matrix form using TF- IDF. After labelling, LDA algorithm will be applied as a dimension reduction method. Then RNN model will be used to detect the attack.

II. NDN BACKGROUND

Before we proceed further, we should have a better knowledge of NDN background. NDN has a vast and complex background that we won't be covering fully. In this paper, we will discuss about the security aspects only including DDoS in NDN. Named Data Networking is an advanced form of Content-Centric Networks (CCNs). Itis completely different from IP network paradigm because it mainly focuses on what is the content rather than where the content is located. NDN can be distinguished from IP network on the basis of assignment of names to the content, rather than IPs. The packets are routed and forwarded by their respective names[3].

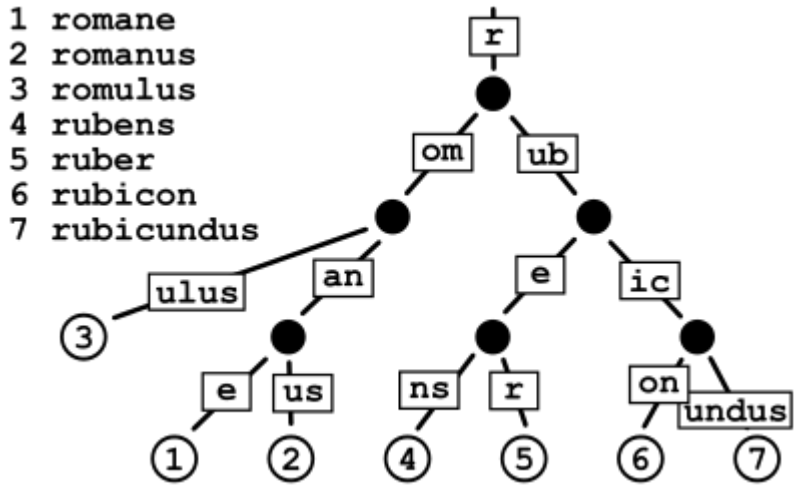


Fig. 1. Example of a radix tree

A. Names in NDN

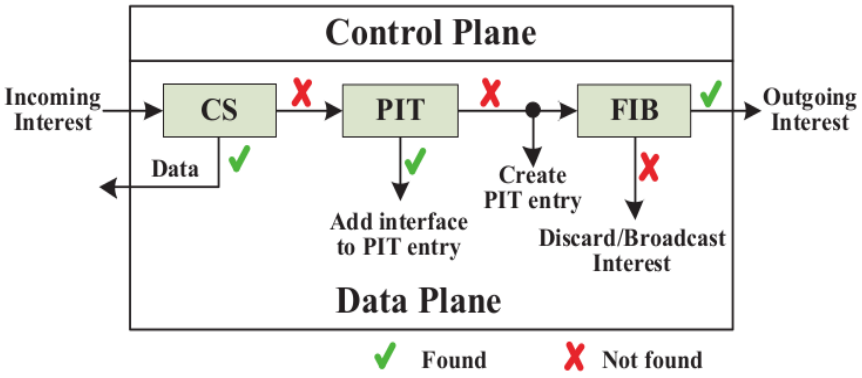
Names in NDN are transparent to the network and also application relying. NDN names share all ordinary characteristics that are structured hierarchically and contain fixed or limited components. The best example of NDN name is: *"edu/in/pondiuni/2019/cfp.html"* where *"edu/in/pondiuni/"* is the reversed domain name of *pondiuni.edu.in, 2019/cfp.html/* is path of the content's directory on the website server. *"/" is* not the part of name, it represents the boundary of the components: *edu, in, pondiuni, 2019* and *cfp.html* are five components of the name[2].

B. Communication Model

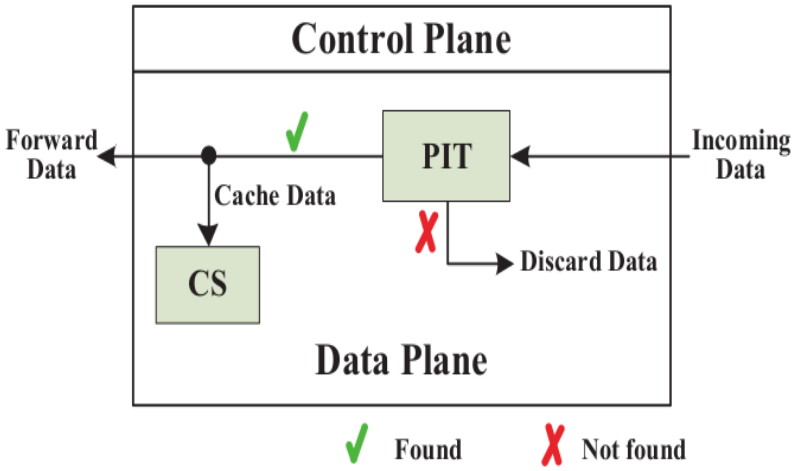
In Named Data Networking, there exists two types of packets: Interest packet (request packet) and Data packet (response packet). The interest packet contains a name that identifies the corresponding data. The name itself is banded in a data packet to represent the content. In NDN, an applicant sends an Interest packet for requested content. The server returns the Data packet with the attached content[2].

C. Packet Forwarding

As illustrated in Fig. 2(a) and Fig. 2(b), NDN router maintains three tables: FIB, PIT and CS. FIB abbreviated as Forward Information Base, is the routing table of NDN. PIT abbreviated as Pending Interest Table keeps a detailed log of restless Interest packets and their incoming interfaces. Content Store (CS) uses a strategy to cache Data packets to serve the ensuing Interest packets that are requesting the same content [2].



(a) Interest lookup and forwarding process.



(b) Data lookup and forwarding process.

Fig. 2. Packet lookup and forwarding process.

III. DISTRIBUTED DENIAL OF SERVICE ATTACKS IN NDN

There are variety of DDoS attacks which aim to reduce the efficiency of the network, software of the protocol layers. A typical DDoS attacks is composed of three segments: (1) a bunch of zombies controlled by master node/s, (2) master node/s, (3) set of victims e.g., hosts or routers [4]. We will now discuss some popular DDoS attacks addressed in TCP/IP based network and their negative effect on NDN architecture.

A. *Reflection Attack*

This attack is generally managed by three sources: the adversary, a victim host or router, few secondary victims. The main purpose of the attacker is to make use of secondary victims. A massive amount of traffic is sent towards the victim hosts to degrade the efficiency. To obtain the goal of adversary, an IP packet is used with copied or fake addresses. The attacker replaces its own address with the address of intended victim and thus sends these IP packets to the secondary ones. For such packets, responses are not acknowledged to the attacker. Instead of this, the victim host or router is downgraded by sending enormous amount of traffic. To occur is attack more effectively, attackers amplify the process by using less data for themselves and flooding the victim with a huge amount of data[4].

This type of attack is generally addressed in NDN as each Interest packet has a uniform path and the related content. The content packet should follow the path initiated by the preceding Interest packet. An NDN router is only permitted to broadcast an arriving Interest packet to some or all of its corresponding interfaces.

B. Bandwidth Depletion

In this attack, the zombies are controlled by the adversary. The zombies flood the victims with enormous amount of traffic to degrade the performance of the network and its resources. The main goal of the attacker is to make the victim hosts disappear from others. In other words, the attacker blocks the communication channel of the intended victim[4].

Bandwidth depletion can also be labeled in NDN architecture. In this attack, the adversary directs a large number of zombies towards the particular content of a specific victim. The efficacy of this attack can be easily determined [4].

C. *Prefix Hijack and Black-Hole*

In this type of attack, an autonomous system is made hostile and configured incorrectly which publicizes its unwanted or fake routers to influence the genuine autonomous system so that it can forward all the traffic to the fake one. This results in Black-Holing where an attacker disposes all the traffic sent to it. This attack is mainly incorporated in IP networks. Once the attack is injected, it is very tough for the routers to detect and retrieve from this muddle[5].

NDN architecture also holds Prefix Hijacking via Black- Holing. NDN routers are very advanced and contain necessary information than IP ones. This information is used to observe the malicious activities in the network during the communication process. Since the content packet follows the path of Interest packet, the amount of expired Interests can be analyzed to identify the hijacking attack[5].

IV. RELATED WORK

Authors from [6] proposed a defense model for Dynamically-Generated Existing Attacks (DGEA). DGEA is less harmful than DGNEA (Dynamically-Generated Non-Existing Attack). The distraction or harmness caused by DGEA can be ignored. Some metrices that are used to observe the DGNEA are not feasible for DGEA. In some cases, certain matrices are used to detect DGEA

using Pending Interest Table (PIT) utilization rate. The authors assume that the genuine users send request for a particular content to a single producer only. In some cases, users might request the content from two or more producers at an equal interval of time.

Authors from [7] has introduced a new technique to target an existing attack called Collusive IFA (CIFA). This method uses a centralized controller to gather information from the routers in the network. The authors have successfully detected the CIFA and they have assumed that this method can be also applied to disharm the DGEA.

Authors from [8] have mentioned a model to mitigate DGEA and DDoS. This model is called FROG. This method is very simple and more effective. FROG runs on those routers that are connected directly to the NDN consumers and thus monitors packet hop-counts. The authors have proved by their simulation results that FROG improves flexibility against DDoS attacks.

In article [9], the authors have mentioned another defense mechanism to mitigate the Distributed Denial-of-Service (DDoS) attacks. This mechanism is called POSEIDON that uses a push-back approach to mitigate interest flooding attacks. POSEIDON is a combination of specific algorithms that run on NDN routers. The goal is to monitor the traffic and identify the malicious activities and then disharm or nullify the attack effects. The authors have simulated IFA using a realistic AT&T network topology as shown in figure. The authors have used R*x*, C*x*, P*x* and A*x* to represent the *x-th* router, producer, consumer and the attacked node controlled by adversary. The continues lines in the figure represent the connections between NDN routers. The dashed lines denote connections between routers and consumers. Similarly, the dotted lines represent the connections between producers and routers respectively. Theauthorshaveused16consumersandonlytwoproducers.

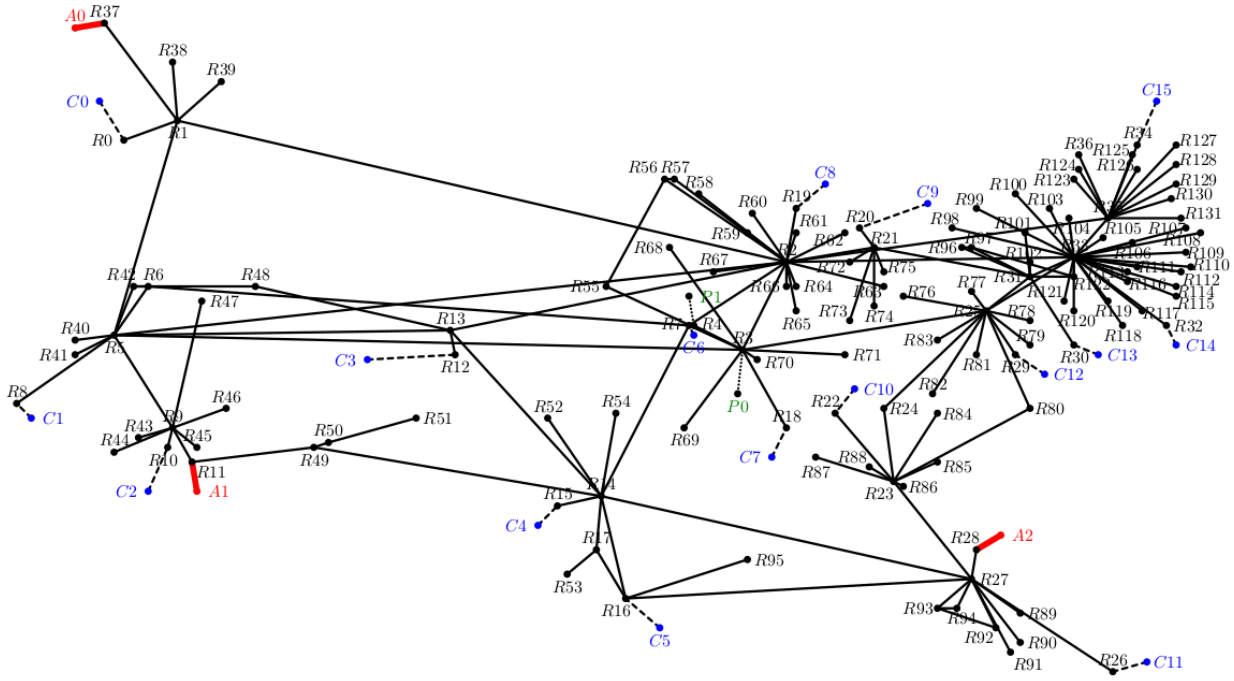


Fig. 3. AT&T Topology.

V. DDoS Discrimination by Linear Discriminant Analysis(LDA)

As per our survey, some mathematical models are featured to identify the prediction degree based of the data arrival rates. The data is thus categorized into two types: predictable and unpredictable data. The mathematical models should determine the data using the threshold value. In this paper, we will discuss a brief about LDA as a detection mechanism [10]. LDA is a mathematical approach that is used to classify objects like people, things, etc. The classification is done on the basis of set of features placed in more than two characteristic groups.

Authors from [10] have mentioned about the Bayes' rule that minimizes the total errors. The objects are assigned to the group *i*which expresses the higher conditional probability against group*j*.

$$P(i|x) > P(j|x), \text{ for } \forall\, j \neq i \quad (1)$$

The authors have obtained the measurement and computed the probability for every class by defining P(*x*|*i*).

$$P(i|x) = \frac{P(x|i) \cdot P(i)}{P(x)} = \frac{P(x|i) \cdot P(i)}{\sum_{\forall j} P(x|j) \cdot P(j)} \quad (2)$$

The equation (1) *i., e* Bayer's rule thus becomes:

$$\frac{P(x|i)\cdot P(i)}{\sum_{\forall k} P(x|k)\cdot P(k)} > \frac{P(x|j)\cdot P(j)}{\sum_{\forall k} P(x|k)\cdot P(k)}, \quad for\ \forall\, j \neq i \tag{3}$$

Our proposal uses LDA as a dimension reduction technique that mainly focuses on maximizes the separability among the known categories. As per our survey, we found LDA more feasible and accurate for NDN data set than PCA. Finally, we gathered the knowledge about RNN algorithm and found suitable for out data set [11]. The implementation would be carried in next phase of research as future work. Using RNN, DDoS can be performed from multiple perspectives. The architecture of the proposed work is shown below and the same will be implemented in our next phase of research. Our survey has observed that LDA is more effective then PCA in both IP architecture as well as NDN architecture, figure 5 clearly defines the measure of effectiveness between PCA and LDA in IP networks and Named Data Networks[11].

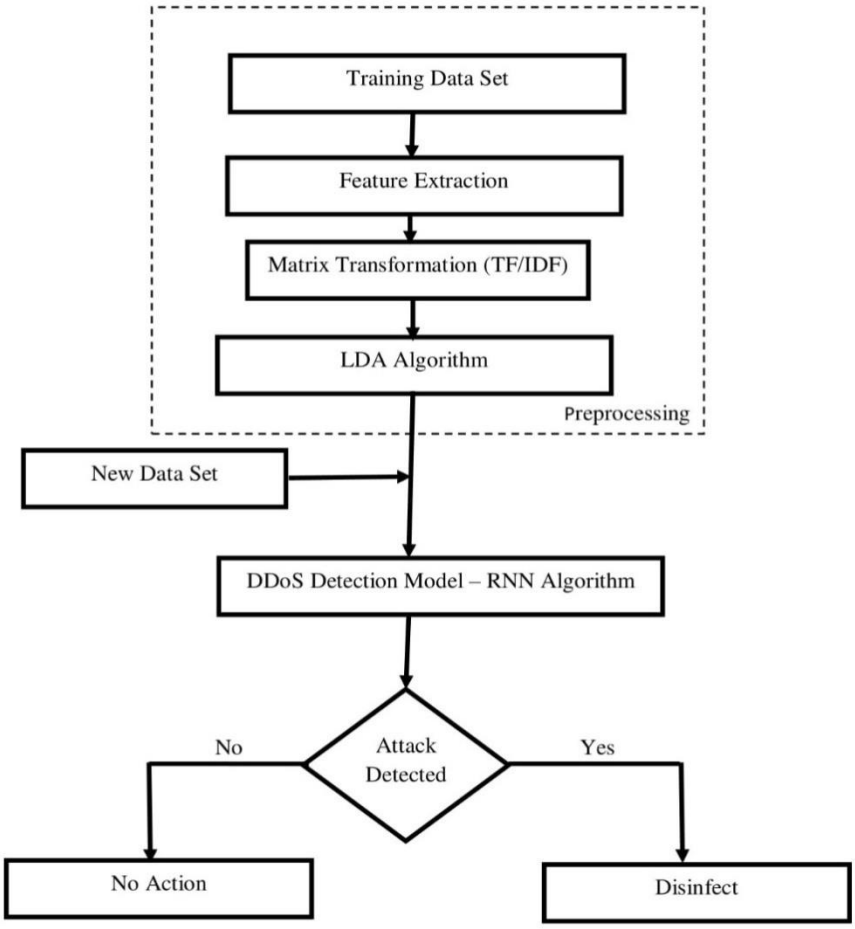


Fig. 4. Proposed LDA - RNN Model.

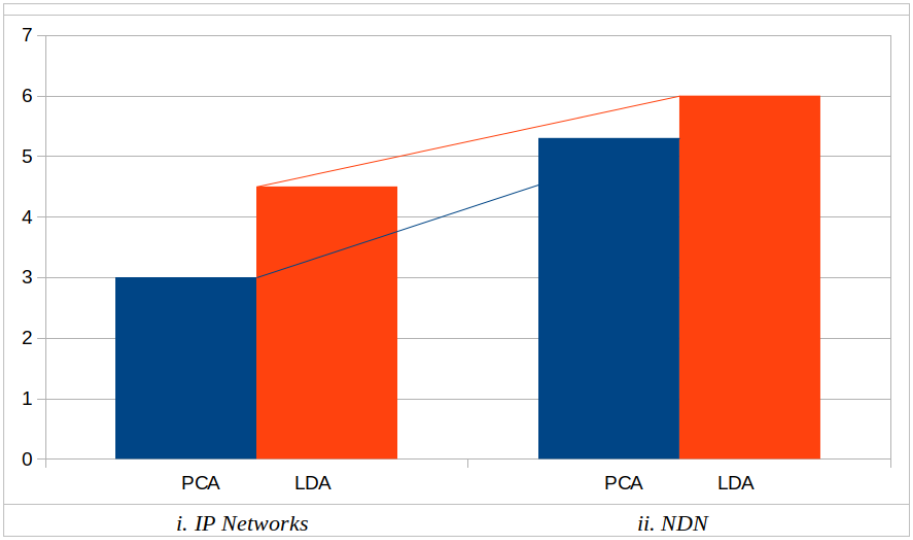


Fig. 5. Accuracy of PCA and LDA on two different network scenarios

VI. CONCLUSION &FUTUREWORK

This paper studies a framework about mitigating DDoS attacks using LDA mathematical model for dimension reduction. We have also described a brief about RNN model as a detection mechanism. Moreover, we have assumed that LDA- RNN model is highly effective in Named Data Networks. In future work, we will be implementing the same model in NDN scenario. The experimental evaluations for LDA & DDoS would be carried using NDN Sim for ns-3. Also, we will be using sci-kit learn & Tensor Flow as additional libraries for RNN. Figure 6 shows the performance metrics of LDA-RNN model compared existing ones[11].

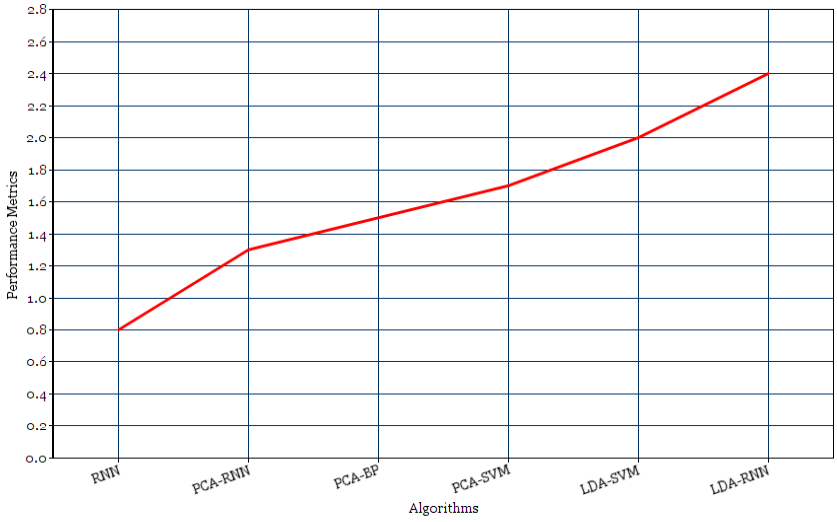


Fig. 6. Performance Metrics Compared with Existing Methods

REFERENCES

[1]. T. Chatterjee, S. Ruj, and S. D. Bit, "Security issues in named data networks," *THE IEEE COMPUTER SOCIETY*, vol. 0018-9162/18, pp. 70–71, JANUARY 2018.
[2]. H. Dai, Y. Wang, J. Fan, and B. Liu, "Mitigate ddos attacks in ndn by interest trace back," *IEEE INFOCOM 2013*, vol. 978-1-4799-0056-5/13, pp. 381–382, 2013.
[3]. L. Zhang, D. Estrin, V. Jacobson, and B. Zhang, "Named data network- ing (ndn) project"," *Technical Report*, vol. NDN-0001, 2010.
[4]. H. Ballani, P. Francis, and X. Zhang., "A study of prefix hijacking and interception in the internet," *SIGCOMM Comput. Commun. Rev.*, vol. 37(4), p. 265–276, AUGUST 2007.
[5]. P. Gasti, G. Tsudik, E. Uzun, and L. Zhang, "Dos-ddos in named-data networking," *arXiv*, vol. 1208.0952v2, pp. 3–4, AUGUST 2012.
[6]. Y. Nakatsuka, J. L. Wijekoon, and H. Nishi, "Frog: A packet hop count based ddos countermeasure in ndn," *IEEE Symposium on Computers and Communications*, vol. 978-1-5386-6950-1/18, pp. 00 492–00 493, 2018.
[7]. H. Salah and T. Strufe, "Evaluating and mitigating a collusive version of the interest flooding attack in ndn," *IEEE Symposium on Computers andCommunication(ISCC)*,p.938–945,June2016.
[8]. A. Afanasyev, P. Mahadevan, I. Moiseenko, E. Uzun, and L. Zhang, "In- terest flooding attack and countermeasures in named data networking," *IFIP Networking*, p. 26–31, 2013.
[9]. A. Compagno, M. Conti, P. Gasti, and G. Tsudik, "Poseidon: Mitigating interest flooding ddos attacks in named data networking," *IEEE Con- ference on Local Computer Networks*, vol. 978-1-4799-0537-9/13, pp. 632–633, 2013.
[10]. T. Thapngam, S. Yu, and W. Zhou, "Ddos discrimination by linear discriminant analysis (lda)," *IEEE*, vol. 978-1-4673-0009-4/12, pp. 533– 534, 2012.
[11]. Q. Liu, L. Meng, J. Yan, and Y. Zhang, "DDos attacks detection using machine learning algorithms," *APNet*, pp. 1–2, August 2-3 2018.