

Preventing DDoS attacks in Path Identifiers-Based Information Centric Networks

Basheer Al-Duwairi* and Öznur Özkasap†

*Department of Network Engineering & Security

Jordan University of Science & Technology

Irbid 22110, Jordan, Email: basheer@just.edu.jo

†Department of Computer Engineering

Koç University, Istanbul, Turkey, Email: oozkasap@ku.edu.tr

Abstract—Information Centric Networks (ICNs) have emerged in recent years in order to provide effective mechanisms for content distribution and retrieval. However, these future network architectures are vulnerable to Distributed Denial of Service Attacks (DDoS). In this paper, we propose an effective mechanism to counter these attacks in path identifier-based information centric networks. The proposed mechanism, called GET Message logging based filtering (GMLF), employs Bloom filter based logging to store incoming GET messages such that corresponding content messages are verified, while filtering packets originating from malicious hosts.

I. INTRODUCTION

Distributed denial of service attacks represent a major threat for traditional and future network architectures. Today's Internet has witnessed major DDoS attack incidents and attack rates have exceeded unprecedented limits, and these attacks continue to grow in number and sophistication. Among the major attack incidents is the one that targeted Github in March 2018 with aggregate traffic rate that exceeded 1.7 Tbps [1]. This comes after the major incident of the well known DDoS attack on the DNS provider Dyn in October 2016 with a traffic rate around 1.2 Tbps. The Dyn attack originated from millions of IoT devices controlled by the Mirai botnet [2]. DoS attacks are not limited to traditional computer networks. They represent a major issue in software defined networks, cloud computing, fog computing, vehicular networking, smart grid systems and named data networks. It is clear that preventing and mitigating these attacks is becoming a priority in network security field as evident by the substantial amount of research in this area. Addressing the issue of DDoS attacks in future Internet architectures is equally important.

Information Centric Networks (ICNs) [3] and Named Data Networks (NDNs) [4] have emerged in recent years as future Internet architectures to address the need of having effective ways to provide content distribution and retrieval. NDNs are based on Content Centric Networking (CCN) paradigm. Generally, content providers distribute their content to large number of nodes (i.e., content stores), and content consumers obtain certain content by sending interest messages. A node that has the requested content replies back with that content.

Routing in these networks is different than the traditional IP networks. Two main routing approaches have been adopted. NDNs rely on maintaining a Pending Interest Table (PIT) at each NDN router where outstanding interests and their arrival interface is stored such that content producers inject their traffic to be sent to the consumer traversing the same path followed by the interest messages. ICNs, which received more attention in recent years, rely on using Path Identifiers (PIDs) to identify paths between network domains where these PIDs are added to interest messages along the path to the content provider such that content follow back the same path. Details of two approaches are presented in Section II.

Both PIT routing and PID-based routing suffer from a major vulnerability to DDoS attacks [6], [18], [19]. In PIT based named data networks, PIT can be overloaded with flood of fake interest messages that would populate PITs with huge number of entries that correspond to fake interest messages resulting in a PIT overflow problem. On the other hand, in PID based information centric networks the feasibility of DDoS attacks depends on attackers ability to learn path identifiers. In some PID based architectures, PIDs are globally advertised and become public. Therefore, PIDs can be learned easily by attackers. However, in other PID based architectures, PIDs (e.g., [19]) are kept secret which imposes additional requirements on attacker to learn PIDs. In both cases, an attacker controlling a set of compromised machines can know path identifiers to send content messages to another node in the network leading to potential DDoS attack.

In this paper, we propose a hybrid approach named GMLF: GET Message logging based filtering. GMLF is similar to PID based ICN architectures in the sense that it uses PIDs to perform inter-domain routing. However, we add interest message digests at ICN routers along the path to the content provider in a flavor similar to that used in PIT based NDN architecture aiming at filtering content packets that do not correspond to previously requested content. Storing interest message digests allows for verifying content messages while maintaining minimal storage at intermediate routers.

The rest of this paper is organized as follows. In Section II, we present the background information and motivation. In Section III, we present the attack model and the previous work is discussed in Section IV. Our GMLF approach is

presented in Section V. Finally, conclusion and future work is presented in Section VI.

II. BACKGROUND AND MOTIVATION

Information centric networking (ICN) is emerging as a future Internet architecture to satisfy the current trend of using the Internet mainly for information dissemination as users are mostly interested in accessing information regardless of its location, and to address shortcomings of current host centric Internet and to provide better security and mobility support. ICN focuses on providing efficient ways to deliver content (information) to recipients. This includes the use of in-network storage and caching mechanisms, content naming approaches, and secure routing techniques. To this end, several information centric architectures have been proposed to meet these requirements in recent years. In this paper, we focus our discussion on two architectures that represent main approaches for content request and content routing in Information Centric Networks. These architectures are the Named data networks [4] and CoLoR [5].

A. Named Data Networking

Named data networks adopt packet logging approach at intermediate routers, where each NDN router maintains three data structures that include: (i) Pending Interest Table (PIT) to record incoming interest messages and their arrival interfaces (ii) The Forwarding Information Base (FIB) (iii) The Content Store (Cs). Two types of packets are used in NDN: interest messages and DATA messages. A content message includes a name, a payload and a digital signature computed by the content producer. Names are composed of one or more components, which have a hierarchical structure. In NDN notation, / separates name components, e.g., /cnn/politics/frontpage. Content is delivered to consumers only upon explicit request. Each request corresponds to an interest message. Unlike content, interests are not signed. An interest message includes a name of requested content. In case of multiple content under a given name, optional control information can be carried within the interest to restrict desired content. Content signatures provide data origin authentication. NDN routers forward interests towards the content producer responsible for the requested name, using name prexes for routing. When an NDN router receives an interest message it checks the content store to see if the requested content is available in which case a content message is sent back through the same interface where the interest is received. Otherwise, the router adds an entry to its PIT along with the router interface where the packet is received. Then it relies on its FIB to forward the interest message to upstream router along the path towards the content provider.

B. CoLoR Architecture

CoLoR was proposed in [5] as an information centric networking architecture that is based on Path Identifier based routing. In CoLoR, every content is assigned a location and

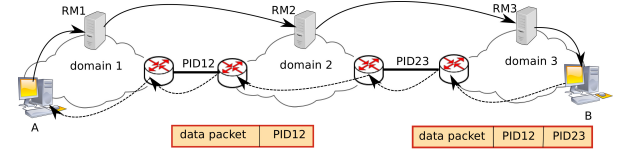


Fig. 1: Operation of CoLoR information centric network architecture. Figure adopted from [19]

application-independent and unique service identifier (SID). The SID is self-certified and remains valid as long as the content is available. Besides, each node is assigned a unique self-certified and persistent node identifier (NID) that does not depend on the node location. Self certification in both SID and NID serves the purpose of authentication without the need of an external authority. The reachability information of content is maintained by a logical resource manager (RM) in each domain. Resource managers in different domains maintain certain relationship with each other (e.g., provider, customer, or peer) and each node within a domain knows the location of the resource manager of that domain. Routing within a domain is done based on intra routing algorithm. Routing between domains is done based on virtual path name which consists of list of path identifiers (PIDs) assigned to links between border routers. In contrast to NDN routers that store PIT messages, in path-ID based ICNs, routers do not log any information about forwarded GET messages.

The operation of CoLoR is illustrated in Figure 1 which shows a network of three domains with a resource manager for each domain. A GET message is issued from node A to obtain certain piece of content. The GET message includes content SID and node's A NID (i.e., the consumer), and is forwarded to the local resource manager (RM1). Assuming that the request content is available at some provider B in domain 3, RM1 appends the path identifier of the path connection domain 1 to domain 2 (i.e., PID12) which is along the path to the requested content. Once the message reaches domain 2, it is forwarded to RM2 in order to decide on the next hop towards the requested content. RM2 appends PID23 as the GET message and forwards it to the router connecting domain 2 with domain 3. Finally, RM3, forwards this message to node B which is located in domain 3. Since the list of PIDs along the path from consumer to provider is already encapsulated in the message, the provider (i.e., node B in this example) can send the content back along the same path. At each router along the backward path, path identifier connecting one domain to another is removed from the list.

III. DDoS ATTACKS IN PATH IDENTIFIER-BASED INFORMATION CENTRIC NETWORKS

Data flooding attack is a type of DDoS attacks that can potentially target path identifier based information centric networks. In this attack, botnet machines are instructed to flood the target machine/network with high volume of data packets. Typically, in these networks, content is delivered from a provider to a consumer along the path followed by the

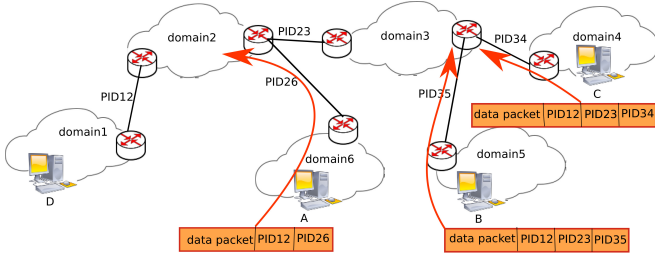


Fig. 2: Data flooding attacks in CoLoR. Figure adopted from [19]

GET message by including the list of received PIDs in the reply message as explained in Subsection II-B. However, in an attack scenario, data packets are flooded towards the target directly (i.e., not in a response to GET requests). This requires the attacker to know the path identifiers (PIDs) used in the network.

Figure 2 shows a typical attack scenario. Nodes A, B, and C represent attack nodes that flood the target (node D) with data packets along the paths, PID34-PID23-PID12, PID35-PID23-PID12, PID26-PID12, respectively. It is clear that making path identifiers public allows attackers to easily launch data flooding attacks. Changing path identifiers dynamically as proposed in [18] can reduce the effect of data flooding attacks and imposes additional constraints on attackers. Also, it was shown in [19] that keeping PIDs secret can reduce the impact of these attacks. However, both strategies do not prevent DDoS attacks completely as attackers can use different techniques to learn path IDs in real-time. These include interest luring technique and botnet cooperation technique.

IV. PREVIOUS WORK ON DDoS ATTACKS IN ICN

Distributed denial of service attacks are growing in number and sophistication. Defending against these attacks represents a top priority in today's Internet due to their devastating effect. Research efforts in this field has resulted in several techniques and approaches for attack detection, mitigation, prevention or traceback. For current Internet infrastructure, large number of DDoS countermeasures were proposed. For example, several techniques (e.g., Kill-bots [7], phalanx [8], JUST-Google [9], and Netfence [10]) have been proposed to mitigate application-layer DDoS attacks.

With the increased interest in Information Centric Networking in recent years and the potential threat of DDoS attacks in these networks, several mechanisms have been proposed to address this important issue. Some of these mechanisms were designed to address the Interest flooding attacks in named data networks (NDN) [11]. In these attacks, the content store or the Pending Interest Table of a router are targeted by high request rate with spoofed content names. In [12], cache pollution attack is a form of DDoS attacks where attackers populate routers' caches of an information centric network with unpopular content. The mechanism proposed

in [12] relies on threshold values for request rate, cache hit ratio, contents rating, and publishers ratings to detect cache pollution attacks. It also devises some defense mechanisms that depend on the attack scenario, where cache content is replaced based on the rating of the content and its usage pattern. Several research papers focused on interest flooding attacks in information centric networks. In [14], a detection mechanism based on the statistical hypothesis testing theory was proposed where the detector employs threshold values that do not depend on users' behavior on each interface of a router within the protected network. However, the paper did not address advanced attack types such as those described in [16] and did not provide mitigation of the attack once detected. In [15], the importance of implementing Interest NACK into content centric networks was analyzed and evaluated as a potential solution to interest flooding attacks in these networks because it helps in evacuating interest tables of ICN routers without necessarily waiting for each interest until it expires. Poseidon was proposed in [17] as a framework for interest flooding attack detection and mitigation by relying on both local metrics and collaborative techniques for early detection of interest flooding.

DDoS attacks in information centric networks that use path-identifiers (PIDs) for inter-domain routing were investigated in [18] and [19]. In [18], it was shown that dynamically changing path IDs between domains can potentially prevent attackers who have the ability to learn PIDs from flooding targeted system by irrelevant content packets. Moreover, the work presented in [19] emphasized that keeping PIDs secret is much better than making them public from security point of view and reduces attackers ability to perform DDoS attacks unless they use some advanced techniques by which they learn PIDs. The works presented in [18] and [19] represent a milestone for evaluating PIDs based ICNs resilience for DDoS attacks by using dynamic PIDs and by keeping PIDs secret. However, the two approaches suffer from major problems. On one hand, changing PIDs dynamically would result in an unnecessary overhead and does not prevent DDoS attacks in case that attacker learn PIDs dynamically. On the other hand, there are several mechanisms that attackers can use to learn PIDs even if they are kept secret.

V. GMLF MECHANISM: LOGGING BASED FILTERING

We propose a simple, yet an effective scheme, named GET Message logging based filtering (GMLF), to prevent data flooding attacks in path identifier based information centric networks. The main idea of the proposed GMLF mechanism is to log some information contained in the GET messages passing through routers such that these routers can verify that response data packets correspond to previously forwarded GET messages. For example, when a router receives a GET message containing SID as the requested content and P1,P2,P3 as the list of path IDs already appended to this message as shown in Fig. 3, it logs this information such that when the response message is received by the same router (as this

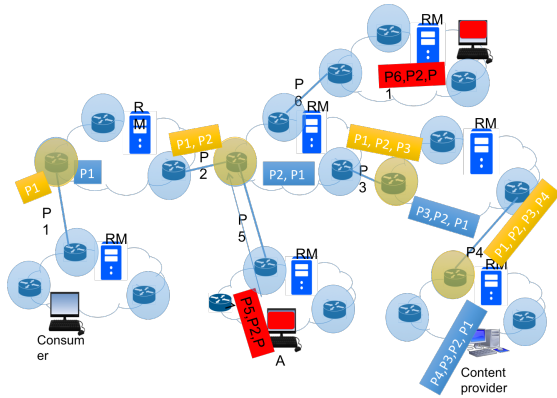


Fig. 3: Illustrative example of the proposed GMLF approach

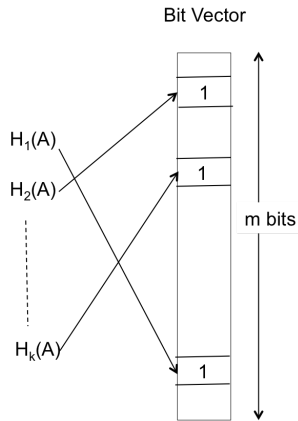


Fig. 4: A Bloom filter with k hash functions

message will traverse the same path according to the list of path IDs), the packet is passed if it contains the same information.

In order to reduce the storage overhead associated with GET messages’ logging, we propose to use Bloom filters [20] to log GET message digests instead of logging the message itself or part of it. A Bloom filter is a data structure for representing a set of n elements (also called keys) to support membership queries. The idea, as illustrated in Fig. 4, is to allocate a vector R of m bits, initially all set to 0, and then choose k independent hash functions, each with range $\{1, \dots, m\}$. For each element, A , the bits at positions $H_1(A)$, $H_2(A)$, ..., $H_k(A)$ in R are set to 1. (A particular bit might be set to 1 multiple times.) Given a query for B , we check the bits at positions $H_1(B)$, $H_2(B)$, ..., $H_k(B)$. If any of them is 0, then certainly B is not inserted in the filter. Otherwise, we conjecture that B is inserted in the filter although there is a certain probability that we are wrong. This is called a “false positive”. The parameters k and m should be chosen such that the probability of a false positive is acceptable.

The false positive rate of a Bloom filters is given by the following equation:

$$p = (1 - (1 - \frac{1}{m})^{kj})^k \approx (1 - e^{-\frac{kj}{m}})^k \quad (1)$$

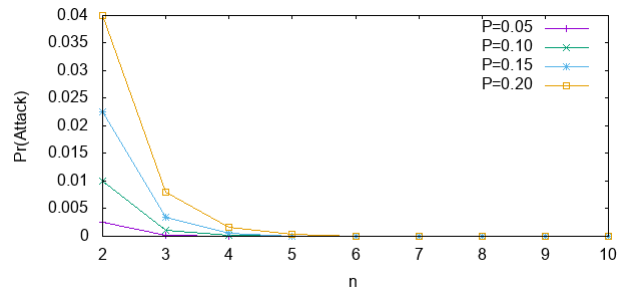


Fig. 5: Probability that an attack packet reaches the victim.

In GMLF, we use the SID and the list of PIDs in the GET message as input for the hash functions of the bloom filter as they uniquely identify each GET message. In this paper, we refer to logging of all GET messages at each router as *Comprehensive Logging*. In theory, if there are n ICN-enabled routers along the path from an attacker to a victim, then the probability that an attack packet reaches the victim is given by Equation 2.

$$Pr(attack) = p^n \quad (2)$$

where p is the bloom filter's false positive probability given by equation 1. Figure 5 depicts the probability that a packet originating from an attacker reaches a victim machine for various false positive rates.

As a future work, we plan to introduce other versions of the proposed mechanism where logging is done in different way in order to reduce the logging overhead. It is to be mentioned that logging is done at ICN enabled router where the GET message leaves a domain, while verification is done at ICN routers where the response message enters a domain. This way, we ensure that logging and verification are done at the same router, and that a GET message is logged only one time in each domain.

VI. CONCLUSIONS AND FUTURE WORK

Addressing the significance of dealing with DDoS attacks in future Internet architectures, we propose GMLF (GET Message logging based filtering) mechanism to counter these attacks in path identifier-based information centric networks. GMLF utilizes Bloom filter based logging to store incoming GET messages and verify them by storing a message only once per domain. As part of work-in-progress, we aim at developing versions of GMLF that employ different logging methods to reduce the overhead. Also, we will develop experimental models and conduct performance evaluations of the GMLF comprehensive logging and its versions.

ACKNOWLEDGMENT

The authors thank Ahmet Uysal, Ceren Kocaogullar, and Kaan Yildirim for their contributions to the development of the simulation platform.

REFERENCES

- [1] Steve Ranger. 2018. GitHub hit with the largest DDoS attack ever seen. ZDNet, <https://goo.gl/BmqekG>
- [2] Antonakakis, Manos, Tim April, Michael Bailey, Matt Bernhard, Elie Bursztein, Jaime Cochran, Zakir Durumeric et al. "Understanding the mirai botnet." In 26th USENIX Security Symposium (USENIX Security 17), pp. 1093-1110. 2017.
- [3] Vasilakos, Athanasios V., Zhe Li, Gwendal Simon, and Wei You. "Information centric network: Research challenges and opportunities." *Journal of Network and Computer Applications* 52 (2015): 1-10.
- [4] V. Jacobson et al., Networking named content, in CoNEXT, 2009.
- [5] Luo, Hongbin, Zhe Chen, Jianbo Cui, Hongke Zhang, Moshe Zukerman, and Chunming Qiao. "CoLoR: an information-centric internet architecture for innovations." *IEEE Network* 28, no. 3 (2014): 4-10.
- [6] Nguyen, Tan N., Rmi Cogranne, Guillaume Doyen, and Florent Retrait. "Detection of interest flooding attacks in named data networking using hypothesis testing." In 2015 IEEE International Workshop on Information Forensics and Security (WIFS), pp. 1-6. IEEE, 2015.
- [7] S. Kandula, D. Katabi, M. Jacob, and A. Burger, "Botz-4-Sale: Surviving DDos Attacks that Mimic Flash Crowds," in *Proc. USENIX NSDI 2005*. Boston, MA, May 2005.
- [8] C. Dixon, T. Anderson and A. Krishnamurthy, "Phalanx: Withstanding Multimillion-Node Botnets," in *Proc. of NDSI 2008*.
- [9] B. Al-Duwairi and G. Manimaran, "JUST-Google: A search engine based defense against botnet-based DDoS attacks," in *Proc. IEEE ICC 2009*, Dersin, Germany, June 2009.
- [10] Xin Liu, Xiaowei Yang, and Yong Xia. 2010, "NetFence: preventing internet denial of service from inside out" In *Proc. of the ACM SIGCOMM 2010 conference on SIGCOMM (SIGCOMM '10)*. ACM, New York, NY, USA
- [11] L. Zhang et al., Named data networking, *ACM Comput. Commun. Rev.*, vol. 44, no. 3, pp. 6673, Jul. 2014.
- [12] AbdAllah EG, Zulkernine M, Hassanein HS. DADI: Defending against distributed denial of service in informationcentric networking routing and caching. *Security and Privacy*. 2018 Mar;1(2):e16.
- [13] Luo H, Chen Z, Li J, Vasilakos AV. Preventing distributed denial-of-service flooding attacks with dynamic path identifiers. *IEEE Transactions on Information Forensics and Security*. 2017 Aug;12(8):1801-15.
- [14] Nguyen T, Cogranne R, Doyen G. An optimal statistical test for robust detection against interest flooding attacks in ccn. In 2015 IFIP/IEEE International Symposium on Integrated Network Management (IM) 2015 May 11 (pp. 252-260). IEEE
- [15] Wang K, Zhao Y, Tong X. On the urgency of implementing Interest NACK into CCN: from the perspective of countering advanced interest flooding attacks. *IET Networks*. 2017 Dec 21;7(3):136-40.
- [16] Signorello, Salvatore, Samuel Marchal, Jrme Franois, Olivier Festor, and Radu State. "Advanced interest flooding attacks in named-data networking." In 2017 IEEE 16th International Symposium on Network Computing and Applications (NCA), pp. 1-10. IEEE, 2017.
- [17] Compagno, Alberto, Mauro Conti, Paolo Gasti, and Gene Tsudik. "Poseidon: Mitigating interest flooding DDoS attacks in named data networking." In 38th annual IEEE conference on local computer networks, pp. 630-638. IEEE, 2013.
- [18] Luo, Hongbin, Zhe Chen, Jiawei Li, and Athanasios V. Vasilakos. "Preventing distributed denial-of-service flooding attacks with dynamic path identifiers." *IEEE Transactions on Information Forensics and Security* 12, no. 8 (2017): 1801-1815.
- [19] Luo, Hongbin, Zhe Chen, Jiawei Li, and Thanos Vasilakos. "On the Benefits of Keeping Path Identifiers Secret in Future Internet: A DDoS Perspective." *IEEE Transactions on Network and Service Management* 15, no. 2 (2018): 650-664.
- [20] Burton Bloom. Space/time trade-offs in hash coding with allowable errors. *Communications of the ACM*, 13, July 1970.