PhD Student paper: Tan NGUYEN Detection and mitigation of emerging attacks in virtualized Named Data Networking (NDN) environment

Ngoc Tan Nguyen, Guillaume Doyen and Rémi Cogranne
ICD - STMR - UMR 6281 CNRS
Troyes University of Technology
Troyes, France
{ngoc_tan.nguyen ; guillaume.doyen ; remi.cogranne }@utt.fr

I. Context

The IP network, which was originally designed to merely connect two computers from a distance, is exposing its limits in front of emerging Internet users' and business' requirements which transform the Internet in a planet-scale framework for content delivery. In this context, an evolution of current network architectures is occurring, inspired by the idea of naming content objects rather than naming nodes with IP addresses. Such an idea has been implemented into many Information Centric Network (ICN) proposals. Among them, **Named Data Networking (NDN)** [1] is the one receiving the most attentions from the research community. NDN shifts the semantics of network service, from delivering the packet to a destination, to retrieving data identified by a given name. This concept brings up important changes to the way the network works: (1) communication is driven by user requests; (2) seamless connection between content providers and users is no longer necessary; (3) mobility support is ensured by removing end-host identification and bringing in-network caches; (4) the latter in-network cache eventually increases the global content delivery performance; and finally (5) multi-path forwarding brings multicast delivery. NDN also brings in security primitives, by implementing signatures for all named data packets, and self-regulation of network traffic, through flow balance between Interest and Data packets.

While NDN appears as a promising solution for the Future Internet, its deployment is still limited to dedicated research testbeds. Furthermore, its adoption by Internet Service Providers (ISP) remains a challenge due to required time and prohibitive cost of large scale deployment. However, the emerging *Network Function Virtualization* (NFV) [2], a concept where network functions can be virtualized and installed over shared pools of standardized commodity hardware resources, emerges as an advantageous means to accelerate and facilitate the deployment of NDN by stakeholders. From a security perspective, NFV also brings challenges and opportunities by (1) clearly separating the infrastructure level from the virtualized one, thus bringing an intrinsic solution to malicious network function isolation, and (2) through the use of *Software-Defined Networking* (SDN) [3], enabling an easy configuration and orchestration of network functions.

In this context[1], the main questions which are raised by subsequent research topic are: *"By leveraging NFV for the deployment of NDN, what are the security threats these novel network functions expose? What are the most appropriate detection solutions and what are the related counter-measures?"*

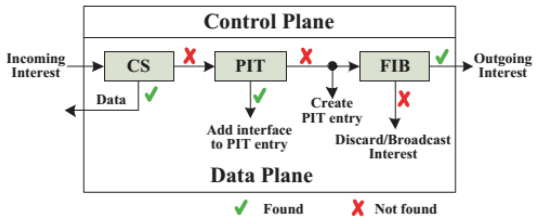
II. Named Data Networking Background

In this section, we briefly introduce key concepts and operations of NDN.

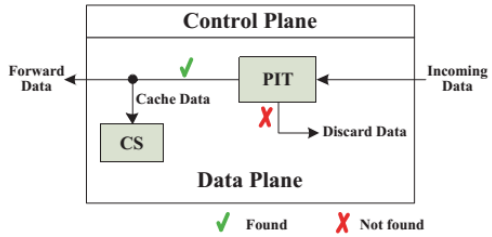
A. Key concepts

NDN is one of ICN proposals - a networking paradigm which is based on data objects. The key concept in ICN is that it names each data object in the network, instead of using IP addresses for naming

[1]The PhD is part of ANR project DOCTOR [4] (DeplOyment and seCurizaTion of new functiOnalities in virtualized networking enviRonments).



(a) Interest lookup and forwarding process.



(b) Data lookup and forwarding process.

Fig. 1: NDN router's operation [5]

hosts and nodes. Secondly, a node in ICN does not have to connect to one specific server to get data. Alternately, this node will send a request with the name of the required data object. Then, the network will return the corresponding object to this node. The third key concept is that ICN deploys in-network caching. Every time a packet passes a network elements, it will be cached. Based on these concepts, many ICN architectures have been introduced.

Among ICN proposals, NDN is the most popular one in research community. It currently allows researchers to evaluate their results with both simulator and implementation. In NDN, communications are based on requests for hierarchical content names and are performed by two type of packets: *Interest* and *Data*. A user sends an Interest packet when he wants to retrieve content and will receive a Data packet in return.

B. Operations

A router in NDN includes three main data structures: (1) *Forwarding Information Base* (FIB); (2) *Content Store* (CS) and (3) *Pending Interest Table* (PIT). The FIB works like a routing table, while the CS acts like a local cache inside, storing every Data packet passing through. The PIT maintains a routing state for each forwarded Interest packet and uses these states to forward the corresponding Data back to the requester. A PIT entry contains a NDN name and multiple incoming interfaces. Figure 1(a) presents the Interest lookup and forwarding process in NDN. Whenever receiving an Interest for a content name, the router will check the CS first. If a cached copy exists, router will send this copy back to the incoming interface. If a cached copy doesn't exist but a PIT entry for this content name is already created, the incoming interface of the Interest will be added to this entry and Interest will be dropped. If a matching PIT entry doesn't exist, a new entry will be created and then the Interest will be forwarded using routing information in FIB. If no matching route is found, the Interest can be discarded or broadcast, depending on the routing policy of the router.

Figure 1(b) presents the Data lookup and forwarding process in NDN. When receiving a Data packet, it checks the PIT. If a matching PIT entry is found, it will cache the Data packet before forwarding it to all the corresponding interfaces in the PIT entry and then this entry will be removed. If the router did not request for this Data, there is no matching PIT entry and the Data packet is dropped. The whole process ensures that one Interest only results in one Data packet.

III. Summary of Achieved and Ongoing Work

In order to identify main security threats which could prevent the emerging NDN technology from being deployed, we have first performed a careful state of the art, based on current literature in ICN security [6]. Given our initial objectives, in order to highlight the most relevant research direction, we have evaluated all the revealed attacks not only on the basis of their impacts on privacy, content delivery and damage scale, but also on their feasibility with the current NDN implementation and the amount of previous works on the corresponding issue. As a result, we have focused our next studies on Denial of Service (DoS) related attacks since they are among the easiest to carry out, while requiring the least efforts from an attacker (i.e. no need to corrupt a router or server, no need to cheat with an certification authority).

Especially, the *Interest Flooding Attack* (IFA) [7] appears as an easily-created attack while providing the most serious damages. The attack consists, for the attacker, in flooding the network with *Interest* packets for non-existing content. In NDN, the router has a stateful nature: it maintains a *Pending Interest Table (PIT)* in order to send Data packet back to the requester. Exploiting this nature, the goal of IFA is to overload the PIT, thus preventing the well-operating of the architecture for legitimate content delivery.

While the state of the art already provides early solution against IFA, especially ones for attack detection, none of them can provide a well-grounded result. Specifically, they cannot provide a clearly-defined threshold for the detector, hence raising question for the network managers when they want to implement the proposed detectors. In addition, previous works cannot provide an expected theoretical performance, hence weakening authors' claim for an optimal result.

As such, we address the attack by proposing a detector based on the statistical hypothesis testing theory. The proposal [8], [9] is evaluated based on data simulated in ndnSIM - a NDN network simulator largely adopted in the community. The hypothesis testing theory allows the proposed detector to have indisputable advantages over previous solutions for IFA. Figure 2 briefly visualizes these advantages our proposal.

First, the theoretical performance of the proposed detector (dash line) can be analytically established and the sharpness of those results have been confirmed with numerical experiments (solid line). This can help a lot in assessing the confidence given in the results, as it is possible to set the false-alarm (false-positive) and missed detection (false-negative) probabilities.

Secondly, the detector's decision threshold is clearly defined, with simple enough computation, hence making the detector easy to be set up. Besides, the threshold does not depend on the attack's behavior but only on the desired false-alarm probability, which can be chosen to satisfy a trade-off between early detection of attacks and decision reliability. In our measurement for detector's input, we expected that there will be a number of samples corrupted by the attack (denoted as M). Figure 2 shows a comparison between the theoretical and the empirical power of the proposed GLRT for three number of corrupted sample, denoted M, 1, 3 and 7. As one would expect, the power increases with the number of corrupted samples. This result emphasizes that the proposed method can be adapted to focus on the quickest detection, hence aiming at detecting only if the last sample is corrupted at a cost of lower detection accuracy. On the other hand, it is also possible to increase the detection delay, hence focusing on the detection of several lasts samples corrupted by the IFA, to ensure a higher detection accuracy.

Finally, while most of proposed solutions apply countermeasure universally, all the time, thus consuming resources and decreasing the content delivery

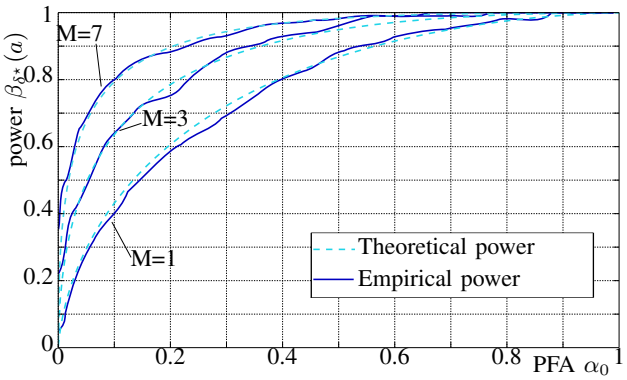


Fig. 2: Receiver Operational Characteristic (ROC) curves for the proposed GLRT with different number of samples corrupted.

performance, applying a low-computation-cost detector helps saving a lot of resources for the network as the counter measure is applied only when an attack is suspected.

IV. Future work

On the basis of this current work, our perspective of future research works are the followings:

- **Evaluate our detection solution with real data:** Although the results of the proposed detector for IFA are very promising, the current empirical performance was estimated with simulated data. Even though those data were obtained using very realistic traffic models, there are still many aspects that can hardly be considered within simulation environments, such as the detection time, impact of real data against simulated ones, computation cost, etc. Hence, we plan to move from a pure simulation framework to real traffic data produced by the reproduction of the attack within a real environment. As such, in the context of DOCTOR project, a testbed for NDN will be deployed in a near future bringing the expected deployment environment. At first, the collected data will be processed off-line. However, dedicated monitoring probes will be deployed and the detector will be implemented in each of them to detect attack on-line.
- **Moving to a distributed detection scheme:** It is important to note that the current detector can be applied on any interface of NDN. However, in cases of distributed attack, the detection can hardly be carried out on single interface. Hence, a distributed detection scheme has to be proposed based on the detector deployed over

each NDN device interface. Since the proposed detector has analytically known statistical performance, this could greatly help in the design of a reliable distributed detection method. On a more practical point of view, this would also greatly help pushing back distributed attacks and also the application of counter-measures.

- **Design a mitigation solution for IFA:**
  As a straight next step of this first perspective, we plan to design and implement a mitigation strategy, built as a part of an autonomous solution, which will especially leverage the SDN technique to realize counter-measure actions and the NFV to implement it.
- **Explore a second major threat in NDN:**
  As a second case of attack detection and mitigation, we have selected another form of DoS attack which can easily occur in NDN: *cache poisoning* [10], an attack aiming at injecting forged content into caches and profiting from caching system to spread such content among users. Although NDN routers are allowed to verify data packets received and to remove altered packets, signature verification of each packet is computationally infeasible and not realistic. Hence, this is a dangerous attack which however receives less attention from the research community. With a methodology similar to the work achieved for IFA, we plan to assess the attack process of such an attack, reproduce it in a simulated environment at first, design and dedicated detection and mitigation solutions and then move to a real implementation in the Doctor testbed [11].

ACKNOWLEDGMENT

This work is partially funded by the French National Research Agency (ANR), DOCTOR project <ANR-14- CE28-000> and supported by the French Systematic ICT cluster.

SPEAKER BIOGRAPHY

Tan NGUYEN is a first year PhD student in Troyes University of Technology (UTT), France. His PhD is co-supervised by Dr. Rémi COGRANNE and Dr. Guillaume DOYEN. His research area focuses on security issues in Information Centric Networks and especially the NDN proposal. His work takes part of the DOCTOR project, started in December 2014 and funded by the French National Agency of Research (ANR). This project, led by Orange, and gathering Thales, the Montimage, the LORIA-CNRS lab and UTT, aims at proposing solutions for the deployment, monitoring and security of ICN architectures deployed in a NFV context.

REFERENCES

[1] L. Zhang, D. Estrin, J. Burke, V. Jacobson, J. D. Thornton, D. K. Smetters, B. Zhang, G. Tsudik, D. Massey, C. Papadopoulos *et al.*, "Named data networking (ndn) project," *Relatório Técnico NDN-0001, Xerox Palo Alto Research Center-PARC*, 2010.

[2] M. Chiosi, D. Clarke, P. Willis, A. Reid, J. Feger, M. Bugenhagen, W. Khan, M. Fargano, C. Cui, H. Denf *et al.*, "Network functions virtualisation: An introduction, benefits, enablers, challenges and call for action," in *SDN and OpenFlow World Congress*, 2012, pp. 22–24.

[3] N. McKeown, "Software-defined networking," *INFOCOM keynote talk*, vol. 17, no. 2, pp. 30–32, 2009.

[4] Doctor project website. [Online]. Available: http://doctor-project.org/

[5] H. Dai, Y. Wang, J. Fan, and B. Liu, "Mitigate ddos attacks in ndn by interest traceback," in *Proc. of IEEE INFOCOM NOMEN Workshop, IEEE Press, Piscataway, NJ, USA*, 2013.

[6] E. AbdAllah, H. S. Hassanein, and M. Zulkernine, "A survey of security attacks in information-centric networking."

[7] A. Afanasyev, P. Mahadevan, I. Moiseenko, E. Uzun, and L. Zhang, "Interest flooding attack and countermeasures in named data networking," in *IFIP Networking Conference, 2013*. IEEE, 2013, pp. 1–9.

[8] T. Nguyen, R. Cogranne, and G. Doyen, "An optimal statistical test for robust detection against interest flooding attacks in ccn," in *Integrated Network Management (IM), 2015 IFIP/IEEE International Symposium on*. IEEE, 2015, pp. 252–260.

[9] T. Nguyen, R. Cogranne, G. Doyen, and F. Retraint, "Detection of interest flooding attacks in named data networking using hypothesis testing," in *Information Forensics and Security (WIFS), 2015 IEEE 7th International Workshop on*, November 2015, pp. 1–6.

[10] C. Ghali, G. Tsudik, and E. Uzun, "Needle in a haystack: Mitigating content poisoning in named-data networking," in *Proceedings of NDSS Workshop on Security of Emerging Networking Technologies (SENT)*, 2014.

[11] Named data networking codebase. [Online]. Available: https://github.com/named-data