An in-network collaborative verification mechanism for defending content poisoning in Named Data Networking

Haohao Kang
School of Computer Science and Communication Engineering
Jiangsu University
Zhenjiang, China
kanghaohaohn@163.com

Yi Zhu*
School of Computer Science and Communication Engineering
Jiangsu University
Zhenjiang, China
zhuyi@ujs.edu.cn

Yu Tao
School of Computer Science and Communication Engineering
Jiangsu University
Zhenjiang, China
771545495@qq.com

Jianlong Yang
JingJiang College of Jiangsu University
Zhenjiang, China
ArcYang@protonmail.ch

***Abstract*—The verification mechanism is the key to ensuring the content security in Name Data Networking (NDN). However, due to the limited computational capacity of NDN router, it is difficult to complete the verification task of all received data packets under heavy traffic. As a consequence, content poisoning has become one of the important security risks of current NDN. To solve this problem, we introduce the concept of data packet credibility and propose an in-network cooperative verification mechanism. In our design, the router calculates the credibility of received data packet from two aspects, one is the internal-evaluation estimated by itself, another is the external-evaluation from its upstream routers. After completing the combined evaluation, router further performs a probabilistic verification according to the credibility. For the data packet with high credibility, it will be verified with low probability. Then router forwards the combined evaluation result to its downstream routers on the reverse path by modifying the structure of data packet. Through building a collaborative verification relationship, this mechanism tries to avoid repeatedly verifying the data packets verified by upstream routers. Simulation results show that it can effectively defend content poisoning while significantly reducing content verification overhead.**

***Index Terms*—Named data networking, Content poisoning, Collaborative verification, Probability verification, Data packet credibility**

I. Introduction

Named Data Networking (NDN) [1] is a typical representative of the next-generation Internet architecture. Using name-based routing and distributed caching mechanism, NDN can achieve location-independent data transmission and effectively solve some problems troubling current Internet, such as bandwidth competition, network congestion, redundant transmission and etc. For security aspect, NDN also make a completely new design. Its security system is based on content itself rather than the traditional link protection. In NDN, each data packet must be signed by its publisher and the router or user will verify the signature to confirm its validity and integrity. Theoretically, this built-in signature verification mechanism can effectively ensure the safety of content. But, when the received data traffic is heavy, the router cannot complete the verification task due to its limited computational source. In addition, for lacking of trust, router will frequently execute repeated verification, which further decreases the in-network verification efficiency. As a consequence, content poisoning has become one of the important security risks of current NDN [2].

Motivated by this problem, we propose an in-network collaborative verification mechanism (named as ICoV in short). In ICoV, the router first evaluates the credibility of received data packet, and then verifies it with a probability according to the evaluation result. When router replies the data packet to its downstream nodes, it must insert the credibility value to data packet as external-reference.

The credibility is the key concept of ICoV, it is determined by two factors. One is the internal-evaluation which is depended on the verification successful probability of the data packet arrival interface, another is the external-evaluation from the upstream routers which is carried in the received data packet. After combining the above two results, the router can make a more accurate judgment on whether the received data packet is legitimate or not. Because the calculation of credibility is depended on the evaluation results from the upstream routers, this mechanism builds a collaborative relationship between routers of transmission path. For the data packet has already checked by upstream router, it will be avoided repeatedly checking in downstream router with high probability. So, our design can significantly reduce content verification overhead while defending content poisoning attack.

The rest of this paper is organized as follows. Section 2 introduces the content poisoning and existing solutions. Section 3 describes the ICoV mechanism in detail. Section 4 evaluates the benefits and overheads of ICoV. We conclude the paper and future works in section 5.

II. Content poisoning and its related works

Content poisoning attack [2], [3] is implements by hijacking a router or content source. Through injecting fake content

* Corresponding author:Yi Zhu(zhuyi@ujs.edu.cn)

978-1-5386-4870-4/18/$31.00 ©2018 IEEE

into network, the fake content will be served for the interest packets with same name, and then extended to the entire network. When the fake content is detected by users, users will repeatedly request the same content. If the repeated request traffic is very heavy, it will exhaust the network resources (such as link bandwidth and computational resource at routers) and then seriously degrade the network performance.

One countermeasure is to filter interest packets to forbid forwarding fake data packets, such as Ghali et al. [4] proposed a light-weight ranking algorithm to distinguish valid and invalid content. For a router, if the received interest packet matches the content with lower ranking, this interest will be discarded. However, some valid interests may also be excluded in this algorithm, as mentioned in [5].

Another important solution is to improve verification mechanism. In recent years, researchers have carried out some meaningful explorations in this direction.

Bianchi et al. [6] suggested a scheme called check before storing (CBS), In CBS, content is verified and cached with a certain probability p. Although the method is advisable, a fixed probability is not a good choice. On one hand, if network is relatively safe, high verification probability will produce more computational overhead. On the other hand, if network environment is dangerous, low verification probability cannot curb poisoned content effectively.

A similar strategy is "verification on hit" which is proposed by Kim et al. [7], [8]. This strategy divides content store (CS) into two parts, one is protected area, another is unprotected. When a data packet arrives, it will be stored into unprotected area first. After hit event occurs, the target data packet will be verified. If it is legitimate, it will be moved to protected area. Otherwise, it will be discarded. This design can reduce the verification pressure dramatically. But because the data packet is not verified when it is stored in CS, the poisoned content doesnt be controlled.

Wang at al. [9] propose a verification scheme named Router-Cooperation, in which the edge routers verify the contents and the core routers no longer verify them. This scheme effectively promotes the in-network verification capability. However, when the traffic is heavy, the edge routers cannot complete all verification tasks.

Gasti et al. [10] points out that a large overhead is expected if every content is verified before being inserted into CS. They proposed a probabilistic verification mechanism according to warning information from adjacent routers. In this mechanism, the router verifies the data packet with probability. If verification fails, the router must send a warning message to its one-hop neighbors. If verification successes, the router will ignore the subsequent warnings from the same interface. This mechanism can partly detect poisoned content within one-hop range, but the transmission of warning messages will lead to an extra traffic and consume the bandwidth resource.

From these research works, we can find the verification mechanism is the major method against content poisoning. But, how to balance the verification overhead and the content poisoning defense capability still afflict current verification mechanism design. It is also the research target of this paper.

III. In-network Collaborative Verification Mechanism Design

For effectively reducing the verification overhead while maximizing the in-network verification capabilities, we design a new mechanism named ICoV in short. In this section, we first introduce the calculation method of credibility from the aspects of internal-evaluation and external-evaluation, then present the probabilistic verification mechanism of ICoV in detail.

A. The internal-evaluation of credibility

In our design, the internal-evaluation of credibility of received data packet is based on the credibility of its arrival interface, the concept of interface credibility is defined as follow.

Definition 1 (Interface Credibility) The interface credibility is defined as the ratio of valid data packets to the total data packets received from this interface during statistical time . Now let $N_{\text{total}}(i,j)$ and $N_{\text{valid}}(i,j)$ denote the total data packets and valid data packets received from interface j in router i respectively, then $Credit_{\text{self}}(i,j)$, which is the credibility of interface j in router i, can be expressed as follow.

$$Credit_{\text{self}}(i,j) = \frac{N_{\text{valid}}(i,j)}{\min\left\{ N_{\text{total}}(i,j),\ \frac{N_{\text{total}}(i,j)}{\sum_{j=1}^{n} N_{\text{total}}(i,j)} \times \text{Limit} \times T \right\}} \quad (1)$$

Where, Limit is the maximum verification capability of a router (packets/s), n is the number of interfaces, $N_{\text{total}}(i,j) \times \text{Limit} \times T / \sum_{j=1}^{n} N_{\text{total}}(i,j)$ is the maximum verification amount on interface j for received data packets during statistical time T.

Since the interface credibility represents the credibility of contents retrieved from the upstream content sources, it can be also used to evaluate the credibility of received data packets as internal-evaluation by router.

B. The external-evaluation of credibility and combined result

Simply relying on the internal-evaluation result, the router only forms one-side cognition about the validity of received data packet. To make a more accurate judgment, ICoV further introduces the external-evaluation of credibility using the information from the upstream routers.

Now we add a new field named *Credibility-Info* to NDN data packet, which is used to record the credibility evaluation results from the routers on the reverse path of data packet, as shown in Fig.1. When data packet passes by a router, the router ID and calculated credibility result for this data packet will be recorded in *Credibility-Info* filed as a new entry. If the data packet has passed by M routers, there should be M entries in this field. Assuming $Credit(M+1)$ is the credibility result of current router (router $Credit(M+1)$) on reverse

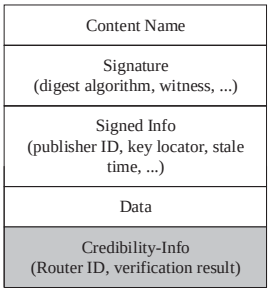


Fig. 1. Modified NDN data packet.

path, if router $M+1$ verifies this data packet and confirm its validity, $Credit(M+1)$=1; if router $M+1$ doesnt verify this data packet, it should estimate $Credit(M+1)$ according to its internal-evaluation and the history M entries recorded in *Credibility-Info* filed, as shown in equation (2).

$$Credit(M+1) = \min\left\{ \frac{\sum_{i=1}^{M} Credit(i)}{M}, Credit_{\text{self}}(M+1, j) \right\} \quad (2)$$

Where $\sum_{i=1}^{M} Credit(i)/M$ represents the average external-evaluation of M upstream routers, $Credit_{\text{self}}(i,j)$ represents the internal-evaluation of current router, so $Credit(M+1)$ represents the combined evaluation result for this data packet in essence. Obviously, as more routers passed by on reverse path, more external-evaluation results the current router obtains. Benefited from the extra information from multi-upstream routers, $Credit(M+1)$ shows more accurate estimation of data packet credibility.

C. The probabilistic verification mechanism of ICoV

According to the combined evaluation result, ICoV implements a probabilistic verification mechanism. When the received data traffic does not exceed the maximum verification capability of router, the router will verify all received data packets. When the received traffic exceeds the maximum verification capability of router, the router will select data packets to verify with a probability $\Pr ob$, its calculation is shown as equation (3).

$$Prob = \min\left\{ 1 - Credit(i), \frac{\text{Limit}}{\sum_{j=1}^{n} N_{\text{total}}(i,j)} \right\} \quad (3)$$

Where n is the number of interfaces, Limit(packets/s) is the limitation verification capacity of router, $\text{Limit}/\sum_{j=1}^{n} N_{\text{total}}(i,j)$ denotes the maximum verification probability of interface j of router i. Equation (3) shows that, the higher the credibility of data packet is, the smaller the verification probability is. With this probabilistic verification, a collaborative verification relationship is built between routers on reverse path. It tries to let router avoid verifying the data packet which has already been verified in the upstream routers.

D. The workflow of ICoV

The workflow of ICoV is briefly summarized in Fig.2. When router i receives a data packet, it will determine how to verify the signature according to its workload.

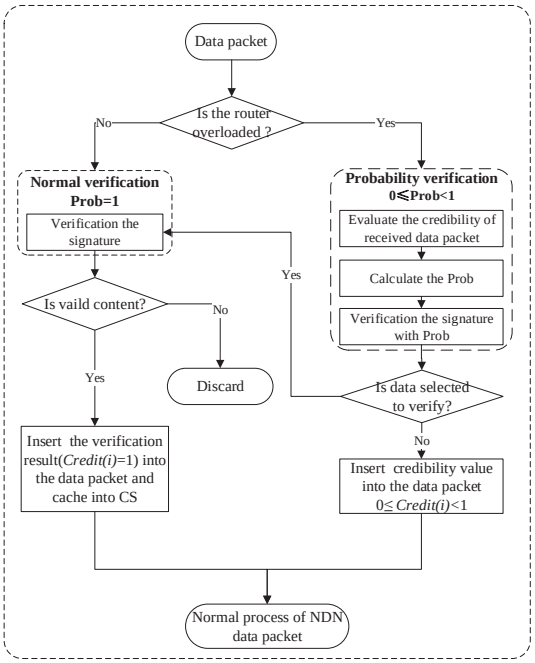


Fig. 2. The workflow of ICoV.

(i) If router i is not overloaded, the router will perform the ***normal verification mechanism*** ($\Pr ob$=1) for data packet. If the verification result is valid, the data packet will be cached, and then router i inserts the verification result into *Credibility-Info* field ($Credit(i)$=1). If the verification result is invalid, the data packet will be discarded directly;

(ii) If router i is overloaded, it will implement the ***probabilistic verification***. First, router i evaluates the credibility of received data packets using equation (2). Based on the combined evaluation result, router i calculates the verification probability $\Pr ob$ ($0 \leq \Pr ob < 1$) and then selects the data packets to verify with $\Pr ob$. For these selected data packets, if the verification result is valid, router i inserts $Credit(i)$=1 into *Credibility-Info* field; if the verification result is invalid, the data packet will be discarded. For these unselected data packets, router i inserts the $Credit(i)$ to the *Credibility-Info* field as credibility reference for the downstream routers.

IV. Performance Analysis

In this section we evaluate the performance of proposed mechanism from the following two aspects, and adopts CBS [7] as comparison mechanism. The simulation tool is ndnSim [11], [12].

(i) *Data Packet Drop Ratio* is measured as the ratio of dropped data packets to total received data packets. This indicator is used to evaluate the defense effect of against content poisoning. In our simulation, link bandwidth is set large enough to exclude the congestion effect, hence, packet drop phenomenon only occurs when the requested content is detected as invalid content.

(ii) *Verification Workload* is defined as the total amount of verified data packets in network. Verification workload represents the overhead of verification mechanism.

A. Simulation conditions

The ndnSIM runs on a Linux machine (Ubuntu v16.04 LTS) with an Intel(R) Core(TM) i7 (6700HQ CPU@2.60 GHz) and 16GB RAM, and enable the new NDN data packet format as shown in Fig.1. by modifying the source code of ndn-cxx. In addition, we use the Origin data processing tools to analyze and draw the experimental results.

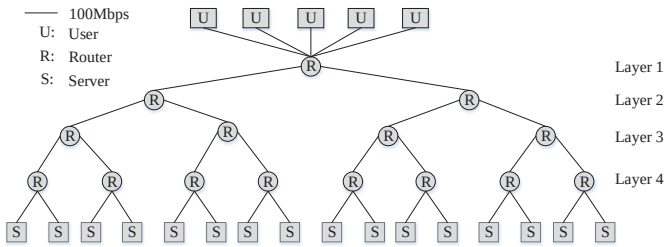
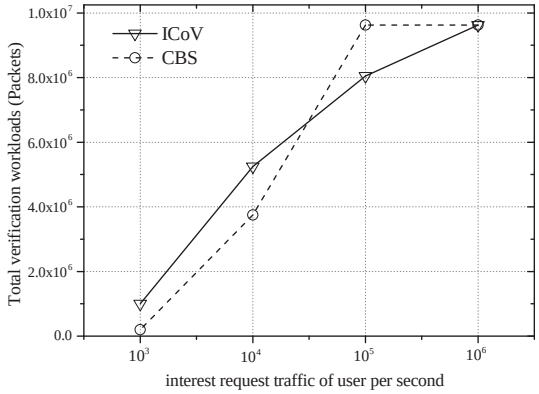


Fig. 3. Simulation topology.

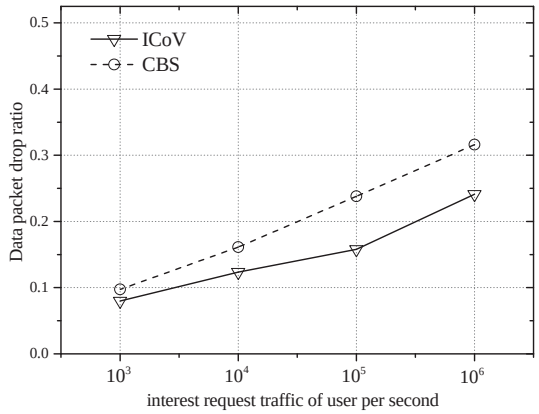
The simulation topology is shown in Fig.3. The NDN server provides 5×10^4 content files. These content files are divided into K=50 classes by content popularity of Zipf distribution [13] ($\alpha = 0.8$), each class has 10^3 content files. CS size of each router is same and set as files. The maximum verification capability of router is set as Limit $= 1.25 \times 10^4$ data $\mathrm{packets/s}$[11]. Requests from user for content in class are generated according to Poisson process with parameter λ. Moreover, the caching replacement adopts the default Least Recently Used (LRU) policy [14], and the simulation time is set as 50s. In this paper, we focus on how NDN router detect poisoned contents effectively. So, we simply generate poisoned contents for each class at the server with a specific probability.

B. Simulation Results

1) Impact of Interest request rate (λ): Now we set the average poisoning rate of servers as 0.2. Caching probability in CBS, p, is set to 0.2. Fig.4 gives the comparison results of ICoV with CBS by changing the interest request traffic of users (λ) from 10^3 to 10^6 (interests/s). In Fig.4.(a), it can be observed that the verification workload of ICoV is higher than CBS when $\lambda \leq 10^4$. Since under the situation of $\lambda \leq 10^4$, for the routers of Layer 3 and Layer 4, the workload of two mechanisms is below their maximum verification capability. With CBS, the lower verification workload is determined by the small verification probability $p = 0.2$. With ICoV, the routers will verify all received data packets, therefore, the



(a)Verification workload



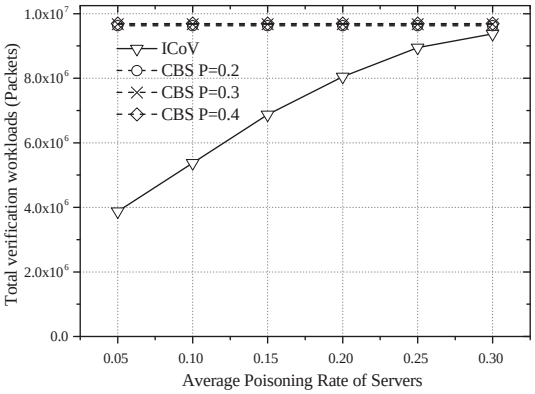
(b)Data packet drop ratio

Fig. 4. Impact of Interest request rate (λ).

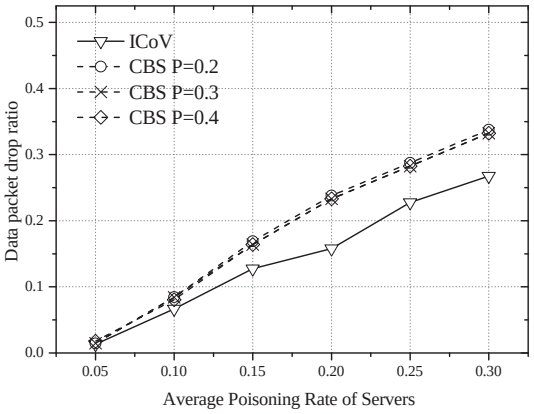
total verification workload of ICoV is higher than CBS. When $\lambda = 10^5$, the verification workload of ICoV is obvious lower than CBS. The reason is that the arrival traffic of $\lambda = 10^5$ exceeds the maximum verification capability of router. With CBS, each router still chooses the data packets to verify with $p = 0.2$. With ICoV, it implements the in-network collaborative verification, avoids repeated verification between routers, so the verification workload is significantly reduced. However, under the heavy traffic of $\lambda = 10^6$, router must verify the received data packet with maximum verification capability, the workload of two mechanisms is approximately same. Fig.4.(b) presents the data packet drop ratio with increasing of λ. When $\lambda = 10^3$, there is no significant difference between two mechanisms. But when $\lambda \geq 10^4$, the data packet drop ratio of ICoV is markedly lower than CBS, and as λ increases, the advantage of ICoV is more obvious. Since a fixed verification probability ($p = 0.2$) in CBS is no longer suitable with the increasing of λ, it is insufficient to defend content poisoning. But with ICoV, the verification probability is dynamically adjusted according to the evaluated credibility, so the defense capability of against content poisoning is better.

2) Impact of average poisoning rate: We further compare the ICoV with CBS under varying the average poisoning rate

of servers. Here, the request rate is set as $\lambda = 10^5$, it is a heavy traffic. With CBS, a group verification probability of p=0.2/p=0.3/p=0.4 is selected in simulation. From Fig.5.(a),



(a)Verification workload



(b)Data packet drop ratio

Fig. 5. Impacts of average poisoning rate ($\lambda = 10^5$).

we can see that the verification workload does not change in CBS with the increasing of average poisoning rate, since the verification workload of CBS is only determined by fixed probability p. With ICoV, although verification workload increases with the increasing of the average poisoning rate, it is overall less than CBS. When the average poisoning rate of servers is at the low level, the advantage is more obvious. The major reason is that ICoV can effectively avoid repeated verification between routers, especially for light poisoned situation. Fig.5.(b) gives the curves of data packet drop ratio, this indicator degrades for both two mechanisms with the increasing of the average poisoning rate of servers. But the performance of ICoV is always superior than CBS. And this advantage enlarges when the poisoning level of servers worsen. The simulation results show that ICoV can effectively defend the content poisoning attack than CBS mechanism.

V. Conclusion

In this paper, we propose an in-network collaborative verification mechanism to improve the verification performance of NDN routers while defending the content poisoning. Based on the combined credibility evaluation of received data packet, this mechanism builds a collaborative verification relationship between routers on reverse path and effectively reduces a large amount of unnecessary repetitive verification. The simulation results prove that ICoV has good performance for both verification workload and defense capacity of content poisoning under heavy traffic. In the future, we will explore how to apply block chain technology in NDN to construct the consensus trust mechanism and then further realize the defense system against content poisoning.

References

[1] Zhang L, Afanasyev A, Burke J, et al., "Named data networking," Acm Sigcomm Computer Communication Review,vol.44, no.3, pp:66-73, 2014.

[2] Lauinger T. "Security & scalability of content-centric networking," Darmstadt: TU Darmstadt, 2010.

[3] Ribeiro I, Rocha A, Albuquerque C, et al., "On the possibility of mitigating content pollution in content-centric networking," 2014 IEEE 39th Conference on Local Computer Networks, Edmonton: IEEE, pp.498-501, 2014.

[4] Ghali C, Tsudik G, Uzun E., "Needle in a Haystack: Mitigating Content Poisoning in Named-Data Networking," the Workshop on Security of Emerging NETWORKING Technologies, pp.68-73, 2014.

[5] Ghali C, Tsudik G, Uzun E., "Network-Layer Trust in Named-Data Networking," Acm Sigcomm Computer Communication Review, vol.44, no.5, pp.12-19, 2014.

[6] Bianchi G, Detti A, Caponi A, et al., "Check before storing: what is the performance price of content integrity verification in LRU caching. Acm Sigcomm Computer Communication Review, vol. 43, no.3, pp.59-67, 2013.

[7] D. Kim, S. Nam, J. Bi, and I. Yeom., "Efficient content verification in named data networking," Proceedings of the 2nd nternational Conference on Information-Centric Networking, New York: ACM, pp.109-116, 2015.

[8] Kim D, Bi J, Vasilakos A V, et al., "Security of Cached Content in NDN," IEEE Transactions on Information Forensics & Security, vol.12, no.12, pp.2933-2944, 2017.

[9] Wang Y, Qi Z, Lei K, et al., "Preventing "bad" content dispersal in named data networking," ACM Turing, Celebration Conference - China. pp.12-14, 2017.

[10] Gasti P, Tsudik G, Uzun E, et al., "DoS and DDoS in Named Data Networking," IEEE International Conference on Computer Communications and Networks, pp.1-7, 2012.

[11] Mastorakis S, Afanasyev A, Zhang L., "On the Evolution of ndnSIM: an Open-Source Simulator for NDN Experimentation," Acm Sigcomm Computer Communication Review, vol.47, no.3, pp.19-33, 2017.

[12] Mastorakis, Spyridon, et al., "ndnSIM 2.0: A new version of the NDN simulator for NS-3," NDN, Technical Report NDN-0028, 2015.

[13] Chiocchetti R, Rossi D, Rossini G, et al., "Exploit the known or explore the unknown?: hamlet-like doubts in ICN," Edition of the Icn Workshop on Information-Centric Networking, pp.7-12, ACM, 2012.

[14] Laoutaris, Nikolaos, Hao Che, and Ioannis Stavrakakis., "The LCD interconnection of LRU caches and its analysis," Performance Evaluation, vol.63, no.7, pp.609-634, 2006.