Content Poisoning Attack in Named Data Networking : A Solution Review

Rohazlin Binti Mohamed Iqbal
Department of Information and Communication Technology
Seberang Perai Polytechnic
Penang, Malaysia
rohazlin@psp.edu.my

Amran Ahmad
Internetworks Research Laboratory
School of Computing
Universiti Utara Malaysia
amran@uum.edu.my

Mohd. Hasbullah Bin Omar
Internetworks Research Laboratory
School of Computing, Universiti Utara Malaysia
mhomar@uum.edu.my

***Abstract*—Named Data Networking (NDN) is one of the potential future Internet architectures that attracts groups of active researchers to study and focusing on the aspects of domains in this architecture. One of the critical issues study by the researchers is the security concern for NDN to deploy in the future. Content Poisoning Attack (CPA) is one of the threats that is crucial to be solved before the deployment of NDN. CPA needs to be studied in terms of its attack mechanism and the results of affecting the network system. This paper provides a survey of the proposed solutions to overcome the attack, which are classified into four categories; hash function and algorithm, forwarding strategy, content verification and hardening network device. Furthermore, this paper shows the comparison between IP based and NDN architecture and the CPA sequence. Finally, there are challenges that need to be look into in order to mitigate CPA.**

Keywords— Content Poisoning Attack (CPA), Named Data Networking (NDN) security.

I. INTRODUCTION

Content Poisoning Attack (CPA) is a content-based attack in Named Data Networking (NDN) which distributing a fake content under a legitimate routable name. The data oriented in NDN, the in-network cache facility implemented in NDN architecture which aims to increase the ability to disseminate data in minimum time has put NDN in a vulnerable state of data integrity attacks, and one of the most concern issues is CPA.

The usage of Internet nowadays is as important as the need of electricity for the citizen. It is not too much to mention that the Internet is capable to make it possible for communication to happen regardless the global distance and in limitless time frame. The establishment of the Internet over more than 30 years is now directed to a different angle as massive content has been produced and distributed. Due to the necessity of the Internet in the community, the scarcity of host-to-host addressing scheme is alarming. Presently, the architecture of IP addressing in focusing on the source and destination between communicators. Since the trend of emphasizing the content despite the original source has promoting the new future Internet architecture; NDN. NDN architecture eliminating the usage of Internet Protocol (IP) address but assigning the routable name in forwarding data and rely in local cache to enhance the capability of disseminating data in lowest latency. The comparison between IP architecture and NDN architecture can be represented by the hourglass [22] in Fig. 1.

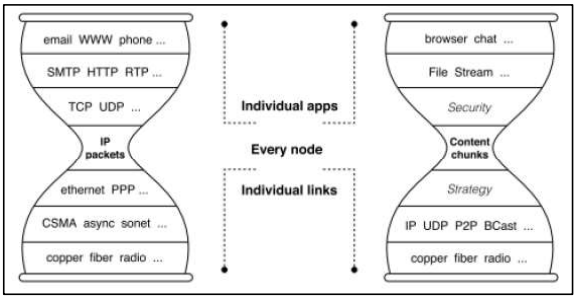

Fig. 1: Hourglass comparison of IP and NDN architecture [22]

Fig. 1 indicate in terms of security, the NDN architecture provides content-based security at immediate layer from content compare to IP architecture which implement security mechanism at the upper layers such as session or application layer. The hourglass compares the architecture build in two main domains of the Internet. IP based is presently in used by the host and server participating in the network and connection oriented between source and destination. In contrast, NDN architecture has put priority for the data instead of the content provider connection with the consumer. The in-network cache ability in the NDN router has increased the capability to disseminate data efficiently in the network compart to IP based architecture which emphasize on the connection between consumer and original content provider. This paper is going to discuss about CPA and approaches from researchers to overcome this issue.

II. CONTENT POISONING ATTACK

A. CPA definition

CPA is content oriented attack by distributing a fake content bind with a legitimate name which resulting in preventing the request consumer to receive a valid content [48]. The attacker respond the request from consumer before the compromised router able to respond with the valid content. To be able respond immediately upon request, attacker need to be located in the network nearby to the legitimate content provider. To overcome this attack is vital because CPA is easy

to launch since it does not need the information of content popularity ranking in-network cache. Hence, due to this nature of the attack, the poisoned content can affect the whole NDN network in a short period of time and degrade the network security performance [12].

B. CPA attack sequence

Research done by [19] describes the sequence of CPA to be launched by the attacker. When a consumer issues an Interest packet, *I_pack* requesting to obtain an information, the attacker will sense the request and producing a fake content. The attacker will bind the fake content with a valid name and signed digitally. The fake content and its digital signature are packaged, and stamped with a valid name as Data packet, *D_pack*. The *D_pack* will be forwarded to consumer who issued the *I_pack*, with retrieval public key. However, the consumer request is not fulfilled because the received *D_pack*, does not consist of valid data although named as a valid content. The aim of the attacker to distribute poisoned content in the network is accomplish in further success with low latency when other consumers requesting the same data due to the capability of in-network cache by NDN router. Poisoned content can be disseminated in the network and the capability on of in-network cache has worsen the situation. Due to this scenario of attack sequence has attract researchers in NDN domain to overcome this issue and proposed solutions in order to minimize the existent of CPA in NDN future Internet architecture. The attack summary sequence between consumer, *C* and attacker, *A* is illustrated in Fig. 2.

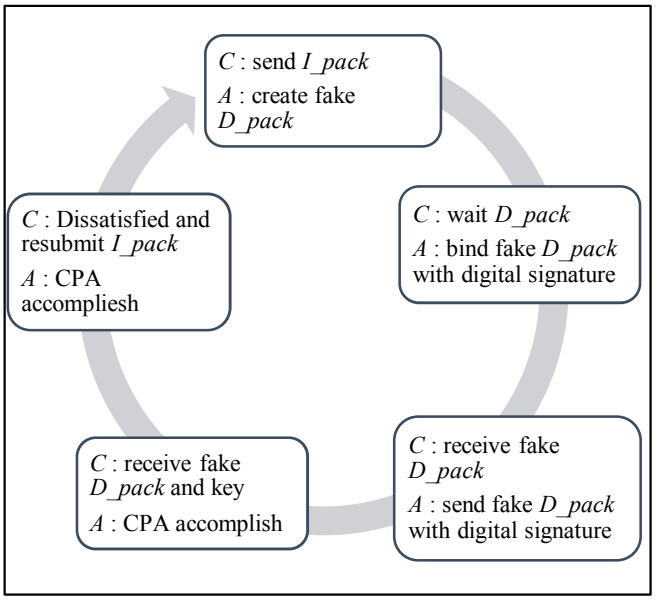

Fig. 2: CPA sequence

Fig. 2 illustrate the sequence of CPA from the beginning until the attack is accomplish. It is clearly display that the attack is in a complete and a never ended cycle. When the attack is accomplished, it does not fulfill the initial intention request by the consumer. Hence, the consumer needs to request again. As long as the attackers are exist in the network, the CPA is easily launched in a short time since there is no need to get cache popularity rank like in a cache poisoned attack, or any human intervention like in Denial of Service (DoS) attack. Due to this attack environment in the NDN architecture, CPA is a vital issue that need to be solved. This issue has attracts many researchers for doing active researches to proposed the solutions in overcoming the issue. This paper will proceed on the proposed solutions approaches by the related CPA field researchers.

III. SOLUTION APPROACHES

CPA occurs when poisoned or invalid content is disseminating by the attacker, matching to a valid name requested by the consumer. Poisoned content is either invalid data or fake content received by the consumer. The content is signed with valid signature or wrong key, or even invalid signature to the key. Focusing on the need to overcome the security issue of CPA, groups of researchers are doing active research on finding and proposing the solutions to solve the CPA. In this section, it will elaborate more on the approaches practiced by the researchers in proposing the solution.

A. Hash function and algorithm

Research done by [4] proposing a homomorphism of hash function which able to enhance the transmission of data and improving the security of network data. However, the proposed approach detecting CPA in theory and verification steps and process involved using quantitative research thus the limiting the actual effectiveness in real environment. A research done using a technique hash-based naming by [14] performing data validity without signature verification, which has an advantage in reducing overhead cost by the router however this approach is only suitable in static environment.

By using proposed ranking algorithm, the researchers [18] develop an access control policy mechanism to identify the existent of unauthorized attack in the simulated environment. This mechanism works only in ICN architecture and addresses the specific type of DDoS attack. Researchers [3] developed a detection algorithm with delay approach, which aims to conceal the data privacy for security purpose, but this mechanism is only can be applied only when the attack exist in the network, thus prevention to avoid the attack is not applicable by using this approach. Besides technique to detect and identify CPA, an approach by giving access control to the genuine owner of the data is proposed by [11] using Merkle Hash Tree which is light and capable to identify the data integrity, hence minimize the existent of CPA in the network.

However, the CPA in actual network environment may not as predictable in simulation environment. Hence, these approaches need to be implement in simulation environment which closely represent the actual network. Due to this issue and awareness, there are groups of researchers emphasizing the need to develop a simulated attack scenario which closely in representing the actual CPA environment in NDN. By simulating the scenarios, the proposed solution is developed and minimizing the gap between actual and experimented scenario.

To be able in simulating attack scenarios which closely represent the actual scenario is a great contribution because to implement a solution, the approach must be based on the situational issue. Knowing to the importance of this attack scenarios to overcome the CPA issue, there are researchers

are moving forward in doing research to develop the attack scenario as similar as the actual attack environment.

Research done by [6] has focusing on realistic attack scenario by developing it with suggested topology. The topology is based on attack behaviors and assumptions for design purpose. It also includes experimental setup with attack evaluation, emphasizing critical weaknesses in the protocol layout and its implementation. Due to the focus of the researchers in developing the attack scenarios, the experiment does focus on mitigating strategies and does not take into consideration of large scale topology.

B. Forwarding strategy

There are other groups of researchers contributing in proposing solution by forwarding strategy technique to solve the CPA issue in NDN architecture. Research by [7] proposing two evasion strategies in forwarding packet. The strategies able to identify valid content and dynamically adapt in topology where there is the existent of attacker hence, avoiding the attack. Due to the evasion strategy by defeating prefix attack, the solution may consume time as it takes many rounds to find the legitimate path thus increase the network latency to disseminate data and increase the overhead cost of network devices.

To overcome the single path issue, researchers [15] proposed a mechanism in in propagating multiple paths per destination, instead of one best path as it will prevents the prediction path by the attacker. This mechanism aims to prevent the attacker from disseminate invalid content to the consumers. Due to the mechanism in providing multiple path to forward the *D_pack*, it will put emphasize on end-to-end connectivity. In other words, regardless the forwarding path, the main concern is to get the *D_pack* safely delivered to the consumer. Emphasizing end-to-end connection does not reflect the research domain of NDN since NDN architecture is content-centric, data-oriented dissemination in the network. Same situation with proposed solution by [9], which specify the use of source and destination prefixes in stateless mode to overcome the problems to related stateful forwarding to aim tightening the security in NDN. However, the use of source and destination approach is more likely suitable in IP architecture compare to NDN architecture. It is not fully accordance with NDN implementation.

Similar approach in forwarding strategy to overcome the CPA, researchers [16] proposed on-path detection by per-hop content integrity check. The advantage of this proposed scheme is it's adaptable with content-centric format in general and the NDN routers are capable to detect any on-path attackers when forwarding the *D_pack*. However, there is no standard protocol for neighboring router to establish key exchange among them as there is no existence of global namespace ownership in the packet.

Although the forwarding strategy is another side of contribution to solve the CPA, there are strengths and weaknesses to be improved before it can be assured as the main solution to overcome CPA. Ongoing research on different approach is taking places before NDN architecture actual implementation such as self-certifying for *I_pack* and *D_pack*.

C. Content verification

Key binding rule involved self-certifying on routable names to verify content signature [8], with aims to reduce the network overhead and the burden of router. It is also targeting to reduce the overhead for publishing new content since the names assigned can be self-certified to verify the data integrity. In order to implement the self-certifying names using the key binding rule, the current *I_pack* format need to be adjusted to include the public key. In order the process of name verifying to be done before forward to router, it will slow down the whole process of forwarding hence, does not fulfill the objective of NDN architecture which aims to disseminate data in low latency. Furthermore, CPA can still be launched by the attacker because there is a possibility of its existent from compromised or malicious NDN router.

Research done [10] proposed a solution for CPA by per-packet verification. Eventually, the solution is suitable of ICN architectures in general but need to be further evaluated. Per-packet verification is good but in large network scalability, there will be too much verification overhead. Due to the need to overcome the verification overhead, researchers [20] proposed mechanism to minimized verification overhead by limiting it to popular content only. However, this mechanism requires router to identify popular content which will increase the router overhead and issues on finding the optimal data. Similarly, in the effort to contribute a solution for minimal verification overhead, [21] suggest for data groups verification but high computational overhead is expected as large size of group requires verification time consuming. Hence, it is not adaptable to different policy compare to initial one.

Verifying process involved the network device such as router. In realization of this, groups of researchers are taking consideration on proposing solution by tightening or hardening the security of network devices.

D. Hardening network device

Network devices such as routers and switches are important in forwarding packets and frames. Dissemination of data in a network depends on the capacity of the router to forward request from consumer in Interest packet, *I_pack* and return with *D_pack*, Data packet with valid data as requested. Due to the importance role of data, groups of researches are doing active researches in order to secure the router from being manipulate or compromised by the attacker. Furthermore, with the functions of in-network cache of router in NDN architecture, data are easily target by attacking the router and poisoned the content. This makes the CPA is vital to overcome as it is easily launched in a short time [12], does not need popularity rank in-network cache and human intervention and spread by compromised router.

Research done by [17] focusing on data transmission by uncompromised router to ensure the data integrity is consistent. The proposed solution does not rely the in-network verification to lessen the burden of router overhead. However, the suggested solution is not suitable to be applied in a large network because compromised router may still exist along the path with compromised host. It may alleviate the problem, but not solving the problem thoroughly. Group of researchers, [1] proposing a scheme to block bad content producer and block bad packet from entering the network.

This scheme is effective with in-network verification. The proposed solution consists of two major parts to overcome CPA. Router verification and digital signature verification by the edge router. Due to this scheme to be effective, although the edge router plays part in verifying the packet, the malicious or compromised intermediate router can still exist and launched the CPA. The network could not exclude malicious intermediate routers from the transmission path and the traffic will turn to heavy, and edge router cannot sustain to complete the verification task.

Another proposed technique by [13] emphasize on the detection mechanism and monitoring any abnormal behavior in the network. This technique rely on client feedback. The technique will monitor client exclusion when sending *I_pack*. The exclusion may represent the abnormal behavior which signed as the potential attacker collaborated with compromised network router or other host participating in the network. This technique solely reply on the client exclusion and may not represent the actually valid client in the network. Based on the inband probe, researchers [5] proposed a lightweight mechanism aiming on mitigating poisoned content inside a network. The proposed mechanism is well adapted to diverse network setting and dynamically suitable for various topology. However, the CPA attacker may still exist intermittently inside the network. Hence it is time consuming for the router to judge the host based on the change in character.

As discussed previously, the CPA proposed solution approaches can be classified into categories as shown in Table 1 that displays four main categories proposed solutions to CPA; hash function and algorithm, forwarding strategy, content verification and hardening network device. These categories represents groups of researchers proposing the similar classifications on proposing the approach of solutions to CPA.

TABLE 1 : CPA PROPOSED SOLUTIONS CLASSIFICATION

No.	CPA Proposed Solutions	
	Categories	***Technique / Mechanism / Scheme***
1	Hash function and algorithm	Homomorphism of hash function [4]
		Hash-based naming [14]
		Ranking algorithm [18]
		LIVE : Lightweight Integrity Verification [11]
		Realistic attack scenarios [6]
2	Forwarding strategy	Evasion strategies [7] • Immediate failover • Probe first
		Availability Centric Routing (ACR) [15]
		Per-hop content integrity check [16]
3	Content verification	Interest-Key Binding Rule [8]
		Cacheshield [10]
		CCNCheck [20]
		Lossy Caching [21]
4		FCPM scheme [12]
	Hardening network device	Router Oriented Mechanism (ROM) [17]
		Router-Cooperation scheme [1]
		Bayesian Network Technique [13]
		Lightweight Content Poisoned Mitigation Mechanism [5]

IV. CONCLUSIONS

NDN is a future Internet architecture that is potential to deploy as contrast to the present IP based architecture. The increasing numbers of participants of hosts in the current network connections has been main triggers for the researchers to study on the future Internet architecture. Scarcity of connection-oriented based by using the IP address is the main driven objective to develop the future deployment of NDN. One of the aspects in NDN is the security concern by listing the potential threats in the network based on its characters of routable names. Due to its priority in disseminating data with low latency has attracts the attackers to launce CPA. The aim is to collapse the main objective of sending the valid data to the consumer. Altered content or invalid key signature is the way CPA works.

As a conclusion, the major challenges to overcome the CPA issue can be categorized as follows:

- Maintaining at the lowest latency on disseminating data.
- Reducing the router overhead and network burden.
- Totally block the existent of the CPA attackers in the network.
- Identifying and eliminating compromised routers and hosts.
- Differentiate between valid consumer and posing attacker.
- Realistic attack scenarios that represents closely to actual environment of NDN in dynamic topologies.

As NDN is the most potential future Internet architecture preferable to be implemented, the NDN security aspect is vital due to the nature of NDN in emphasizing data integrity. The validity of content with high dissemination is the main features of NDN. Data received by the consumer must be valid and trustable regardless the source of the data.

REFERENCES

[1] Y. Wang, K. Lei, B. Liu, and C. Tian, "Preventing " Bad " Content Dispersal in Named Data Networking," China Communications, vol. 15, no. 6, pp. 109 – 119, 2017.

[2] D. Saxena and I. I. T. Roorkee, "Named Data Networking: A Survey," Computer Science Review, Elsevier, vol. 19, pp. 15—-55, 2016.

[3] E. Dogruluk, A. Costa, and J. Macedo, "Evaluating Privacy Attacks in Named Data Network," in IEEE Symposium on Computers and Communication (ISCC), 2016.

[4] T. Feng, X. Ma, X. Guo, and J. Wang, "Secure Network Coding against Content Pollution Attacks in Named Data Network," vol. 3, no. 4, pp. 303–307, 2015.

[5] X. Hu, J. Gong, G. Cheng, G. Zhang, and C. Fan, "Mitigating Content Poisoning with Name-Key Based Forwarding and Multipath

Forwarding Based Inband Probe for Energy Management in Smart Cities," IEEE Access, vol. 6, pp. 39 692– 39 704, 2018.

[6] T. Nguyen, X. Marchal, G. Doyen, and T. Cholez, "Content Poisoning in Named Data Networking : Comprehensive Characterization of real Deployment," no. 1, pp. 72–80, 2017.

[7] S. Dibenedetto and C. Papadopoulos, "Mitigating poisoned content with forwarding strategy," Proceedings - IEEE INFOCOM, vol. 2016-Septe, pp. 164– 169, 2016.

[8] C. Ghali, G. Tsudik, and E. Uzun, "Elements of Trust in Named- Data Networking," 2014. [Online]. Available: http://arxiv.org/abs/1402.3332{%}0Ahttp://dx.doi.org/10.1145/2677046.2677049

[9] C. Ghali and C. A. Wood, "Living in a PIT- LESS World : A Case Against Stateful Forwarding in Content-Centric Networking," 2015.

[10] M. Xie, I. Widjaja, and H. Wang, "Enhancing cache robustness for content-centric networking," Proceedings - IEEE INFOCOM, pp. 2426–2434, 2012.

[11] Q. Li, X. Zhang, Q. Zheng, R. Sandhu, and X. Fu, "LIVE: Lightweight Integrity Verification and Content Access Control for Named Data Networking," Ieee Transactions on Information Forensics and Security, vol. 10, no. 2, pp. 308– 320, 2015.

[12] W. Cui, Y. Li, Y. Xin, and C. Liu, "Feedback-Based Content Poisoning Mitigation in Named Data Networking," 2018 IEEE Symposium on Computers and Communications (ISCC), pp. 1–7, 2018.

[13] H. L. Mai, T. Nguyen, G. Doyen, R. Cogranne, W. Mallouli, E. M. De Oca, and O. Festor, "Towards a security monitoring plane for named data networking and its application against content poisoning attack," IEEE/IFIP Network Operations and Management Symposium: Cognitive Management in a Cyber World, NOMS 2018, pp. 1–9, 2018.

[14] M. Baugher, B. Davie, A. Narayanan, and D. Oran, "Self-verifying names for read-only named data," Proceedings - IEEE INFOCOM, pp. 274–279, 2012.

[15] D. Wendlandt, I. Avramopoulos, D. G. Andersen, and J. Rexford, "Don't secure routing protocols, secure data delivery," In Proc. 5th ACM Workshop on Hot Topics in Networks (Hotnets-V, 2006.

[16] C. Ghali, G. Tsudik, and C. A. Wood, "Mitigating On-Path Adversaries in Content-Centric Networks," Proceedings - Conference on Local Computer Networks, LCN, vol. 2017-Octob, pp. 27–34, 2017.

[17] D.Wu, Z. Xu, B. Chen, and Y. Zhang, "What if routers are malicious? Mitigating content poisoning attack in NDN," Proceedings - 15th IEEE International Conference on Trust, Security and Privacy in Computing and Communications, 10th IEEE International Conference on Big Data Science and Engineering and 14th IEEE International Symposium on Parallel and Distributed Proce, pp. 481–488, 2016.

[18] N. Fotiou, G. F. Marias, and G. C. Polyzos, "Fighting spam in publish/subscribe networks using information ranking," 6th Euro NF Conference on Next Generation Internet, NGI 2010 - Proceedings, 2010.

[19] B. Hamdane, A. Serhrouchni, A. Fadlallah, S. Guemara, and E. Fatmi, "Named- Data Security Scheme for Named Data Networking," 2010.

[20] I. Ribeiro and F. Guirmaraes, "Content Pollution Mitigation for Content-Centric Networking," in Proc. 7th NOF, Buzios, Brazil, 2016, pp. 1–5.

[21] G. Bianchi, A. Detti, A. Caponi, N. Melazzi, G. Bianchi, A. Detti, A. Caponi, and N. Blefari-melazzi, "Check before Storing : Which Performance Price of Content Integrity Verification in LRU Caching?" vol. 43, no. 3, pp. 60–67, 2013.

[22] L. Zhang, D. Estrin, J. Burke, V. Jacobson, J.D. Thornton, D.K. Smetters, B. Zhang, G. Tsudik, K.C. Claffy, D. Krioukov, D. Massey, C. Papadopoulos, T. Abdelzaher, L. Wang, P. Crowley, and E. Yeh. (2010). Named Data Networking (NDN) Project. [Online]. Available: http://named- data.net/project/annual-progress-summaries/