Network Analyzer in NDN infrastructure

Alessio Civica, *1916744, Sapienza University of Rome,*
Federico Detomaso, *1903906, Sapienza University of Rome,*

Abstract—ICN is an innovative network architecture called Information-centric Networking, which aims to replace the current IP-based architecture through content and interest packets. Each content, that is uniquely identified with a name, can be cached into routers in order to immediately serve subsequent requests. One of the most ICN protocol is called Named Data Networking (NDN) that presents some limitations in protecting user privacy. The idea is to introduce an advanced network analyzer designed specifically for the NDN environment, aiming to enhance network security by proactive monitoring and immediate intervention. We propose a Network Analyzer based on 2 approach:

- Real-time reputation assessment mechanism: it's engines on AI that updates the reputation levels of network consumers based on their activity patterns, such as frequency and nature of requests.
- Signatures approach: it blocks any request that matches a blacklisted word contained inside the local database.

By placing the network analyzer in a border router, it preemptively addresses potential security threats before they propagate inside the network.

Index Terms—Information-centric Networking, Named Data Networking, Network Analyzer, Reputation-based, Artificial Intelligence

1 Introduction

Considering the exponential increase of IoT devices connected to the network, it is necessary to find a better model that preserves energy consumption and improves network efficiency. This change is critical to managing the network of interconnected devices. By focusing on energy-efficient models and smarter network management techniques, the goal is to maintain high performance without sacrificing network reliability or functionality.

Named Data Networking (NDN) has emerged as a promising model under the concept of Information-centric Networking (ICN). NDN changes traditional network paradigms by relying on content, identifying it through unique names. This fundamental change facilitates direct caching within the network, enabling efficient data retrieval and optimized use of network resources.

However, NDN introduces complex challenges, particularly with privacy and security. Recognizing these vulnerabilities, this paper introduces a Network Analyzer tailored for the NDN infrastructure. This solution leverages deep network insights to bolster security frameworks, employing an innovative real-time reputation assessment system. This system dynamically evaluates the reputation of network consumers, considering variables such as the frequency and specificity of data requests.By integrating this analyzer strategically into border routers, it acts preemptively to neutralize threats and mitigate potential breaches before they penetrate the network.

This analyzer aims to prevent and identify threats coming from malicious consumer, improving network control through various features. The analyzer processes for each consumer its reputation through which it is possible to identify possible compromised consumers in real time, paying attention to them directly at a control center or excluding them from the network in an autonomous way.

2 Background

NDN operates on two primary types of packets: Interest packets and Content packets. Interest packets are sent by a consumer to request data, and Content packets are the responses that carry the actual data back to the requester. This communication model leverages several key components within NDN routers to manage data requests and delivery efficiently:

Pending Interest Table (PIT): The PIT plays a critical role in NDN routers. It keeps track of all the Interest packets that have passed through the router and are awaiting corresponding Data packets. Each entry in the PIT contains the name of the requested data and a list of interfaces from which the requests were received. When a Data packet arrives that matches an Interest in the PIT, the router forwards the Data packet back through the interfaces listed in the PIT entry for that data name, thus completing the data request cycle. The PIT ensures that routers do not resend Interest packets redundantly and helps in tracing the path back to the original requester(s).

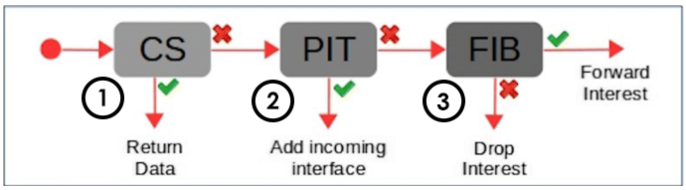
Forwarding Information Base (FIB): Similar to the routing table in IP networks, the FIB in NDN is used to forward Interest packets towards potential data sources. The FIB is populated with information about which interfaces lead to data sources for particular data names. It uses prefix matching to route Interest packets based on their names towards the data producers. This component is vital for efficiently guiding requests through the network.

Content Store (CS): The Content Store is an in-router cache where Data packets are temporarily stored. When an Interest packet arrives, the router first checks the Content Store to see if the requested data is already available locally. If found, the router immediately returns the Data packet from the Content Store, significantly reducing data access time and network traffic. This caching mechanism exploits

the temporal and spatial locality of data requests, making NDN inherently efficient in content distribution, especially popular content.

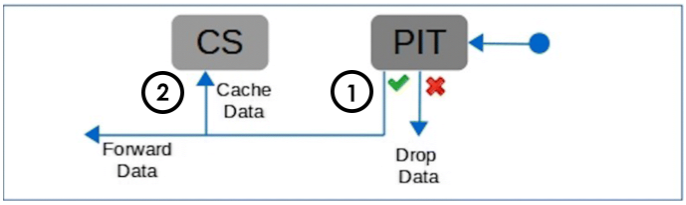
2.1 Operational Flow in NDN

Each NDN router in the network performs the following operations.
When an NDN router receives an interest:



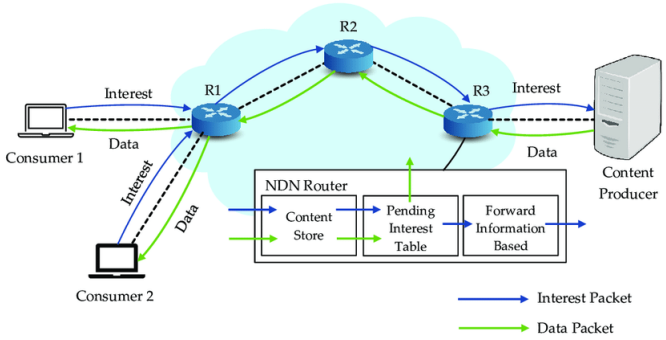
1) First check if the content is present in the CS (Content Store). If the outcome is positive, the router returns the content to the requesting consumer. If the result is negative, the check is carried out in the PIT.
2) In the second point the router checks whether the content with the relevant interface of the requester is already present in the table. If the interest was sent from another interface it is added to the requested content tuple. If the content is not present in the PIT, it is added and the interest is forwarded in the FIB. In the event that an interest has the same content and interface, the router discards that interest as it is already present in the PIT.
3) In the FIB, the router checks whether a route exists to route the interest to another interface. If it is not present, it drops the interest. If so, it forwards to the next node.

When an NDN router receives content:



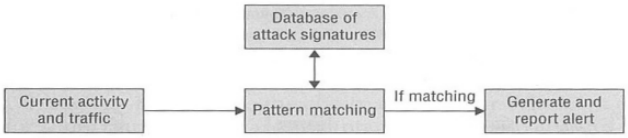
1) First the router checks the PIT table and examines if any pending interest has requested that particular content. If the result is negative, it drops the content.
2) If the result is positive, it stores the content in the CS and forwards the content to the requesting node.

This operational flow allows NDN to be highly efficient in data distribution, reduces redundancy, enhances security by decoupling data from location, and fundamentally supports a more resilient and scalable network.



2.2 Signature-Based Detection

A signature is a pattern that corresponds to a known threat. Signature-based detection is the process of comparing signatures against observed events to identify possible incidents. Signature-based detection is the simplest detection method because it just compares the current unit of activity, such as a packet or a log entry, to a list of signatures using string comparison operations.



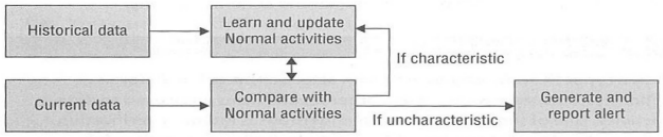
This approach has different advantages and disadvantages:

- **Advantages**:
  - High computation speed;
  - Low latency;
  - Powerful hardware not required.
- **Disadvantages**:
  - Extensive database;
  - Not adaptive to new types of malicious requests;
  - High frequency of database updates.

2.3 IPS with Artificial Intelligence approach

In the rapidly evolving landscape of cybersecurity, the deployment of Artificial Intelligence (AI) has made a significant impact. One area where AI has proven particularly promising is Intrusion Prevention Systems (IPS). AI-powered IPS systems leverage machine learning algorithms and advanced analytics to detect and prevent various forms of new cyber threats. These systems combine the power of AI with traditional IPS technologies, enhancing their capabilities to identify and mitigate attacks in real time. The integration of AI in IPS systems brings forth a new level of threat detection and prevention. Traditional IPS systems, based on behavior approach, often suffer from a high number of false positives, generating a flood of alerts that can overwhelm security teams. AI-powered IPS systems leverage machine learning algorithms to continuously refine their threat detection capabilities. By learning from historical data, they can better discern between genuine threats

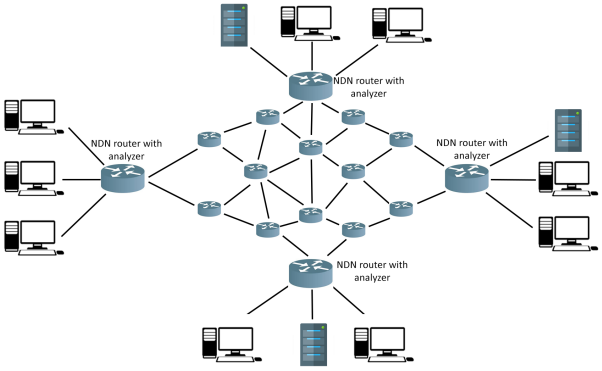
and benign network activities, leading to a significant reduction in false positives. This enables cybersecurity professionals to focus their attention on actual threats, improving operational efficiency.



3 Overview of our proposed approach

3.1 Network Design

Our proposal is based on a particular network where it requires that each border ruoter has an integrated analyzer and that each consumer is directly connected with a single NDN router, having multiple interfaces that uniquely identify a consumer, this model was chosen not only to uniquely identify a consumer but also to block threats as early as possible in order not to allow other routers or consumers to be infected by it, so as to create a kind of protected perimeter of the NDN network, this implementation is developed based on a network in which consumers frequently request contents through requests destined for the target server. It is assumed that at the beginning of the lifecycle the consumer is not malicious. The last assumption we considered is that each NDN border router with analyzer only checks incoming requests, since it is not possible to uniquely identify one consumer over another if the data packet has performed more than one hop inside the NDN network



3.2 Analyzer Design

In our proposal, we considered a hybrid IPS-like approach used to prevent different types of attacks and capable of mitigating them in different solutions.
Our analyzer consists of:

- **Signature-based approach**: used as a first frontier for detecting possible malicious attacks present within the local database of each border router;
- **AI-based approach**: based on AI-enhanced behavioral approach, focuses on analyzing consumer history and creating a personalized model based on consumer behavior.

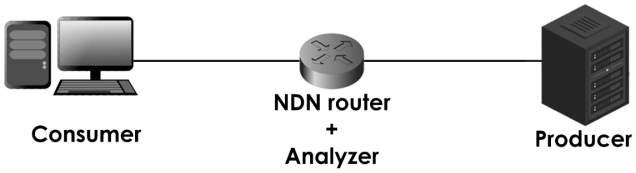
3.3 Signature-based approach

The idea of our approach is to integrate a database in each border router with the analyzer so as to preemptively perform like a signature-based IPS. This is implemented by defining a set of prohibited content names.

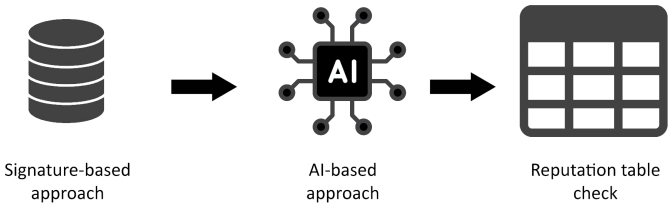
3.4 AI-based approach

To handle all those requests that are not prohibited, we thought we would rely on an AI-based system that defines and studies the behavior of a consumer. Starting with a predefined dataset for each consumer connected to the router, the router trains each model according to the consumers' requests. Arriving at a good security level, the model creates a profile of the consumer and therefore recognizes normal and abnormal behavior.
Through the model, the analyzer can calculate consumer reputation in the manner most inherent to its behavior.

3.5 Operational flow with the analyzer

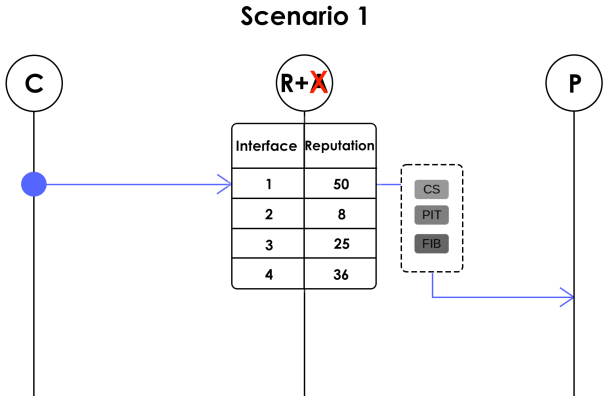


In our network with an NDN router with an analyzer inside, we will analyze the various possible scenarios to understand the behaviors and differences from a traditional NDN network.
The consumer sends an interest to the border router. The first check is performed by comparing the request with the database containing prohibited content names. If a match is found, the request is immediately blocked; otherwise, it is passed to the AI model that checks the request. If the request is inherent to the profile modeled for the consumer, the reputation is increased. If the request is out of normal behavior it is considered suspicious and reputation is decremented. Finally, the reputation table will be checked to verify the reputation of the node.



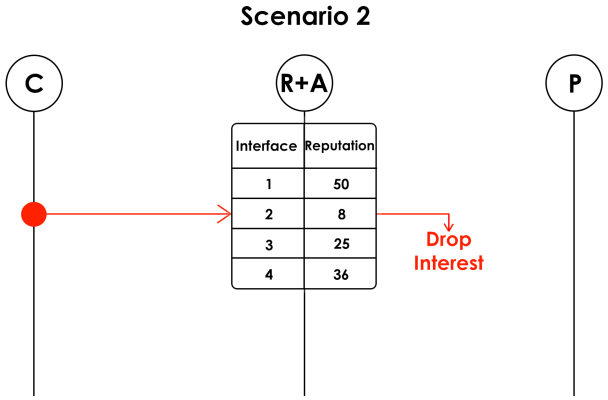
From here we can have different scenarios:

3.5.1 Scenario 1



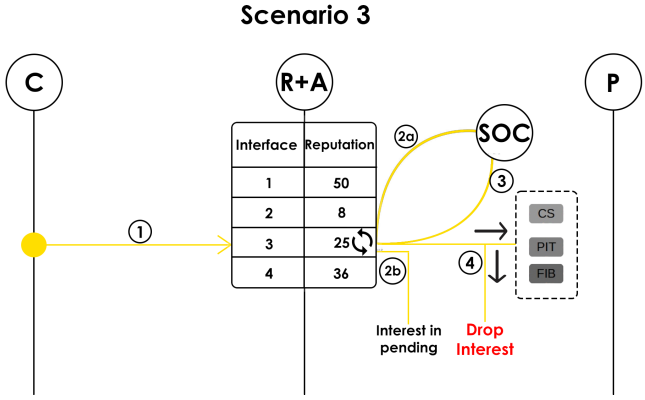
In scenario 1 the router receives the interest and checks the corresponding interface in the reputation table. The value will be greater than 30 therefore the router will behave like a normal NDN router carrying out the operations explained previously.

3.5.2 Scenario 2



In scenario 2 the router receives the interest and checks the corresponding interface in the reputation table. The value will be less than 10, therefore the router will drop the interest. This behavior prevents problems in the network since the router is located in the border and immediately blocks the malicious consumer.

3.5.3 Scenario 3



In scenario 3, in the corresponding entry of the interface, the reputation value will be less than 30 and greater than 0. The router will put the request in Pending and at the same time a report will be sent to the relevant SOC so as to carry out the checks in more detail by a specialized operator. Once the consumer has been analyzed, the operator will be able to respond with a Positive or Negative outcome.
If it is positive, the router will cut the connection to the malicious consumer and it will update the reputation of the interface to 0 for the future interests.
If it is negative, the router will execute the normal flow of the NDN router and will reset the reputation of the consumer to 50.

3.6 Features

3.6.1 Identification and Control

Given the NDN architecture approach in which devices connected to the network are not uniquely identified by IP address, the only method to uniquely identify a consumer is to connect it to the first border router, this approach goes a long way in increasing the timeliness in mitigating the malicious packet which will be cut off the communication promptly and reduce the impact in the whole network.

3.6.2 Trust Tracking

An important feature of this analyzer is its ability to track and manage the Reputation score associated with each consumer in the NDN network. Any border router maintains a Reputation table where consumer can be listed with reputation score.

3.6.3 Data Management

The analyzer stores detailed data on each consumer, which includes the interface ID, the time of the last update, the total number of requests made by the consumer. This data is crucial for maintaining the accuracy of AI model for ensure high control.

3.6.4 Dynamic Reputation Adjustment

Reputation scores are dynamically adjusted based on specific behaviors such as abnormal request patterns based on AI. Regular updates to the reputation scores are essential to reflect the current status of consumer based on their recent activities.

3.7 Algorithm for Consumer Reputation Update

3.7.1 Variables and Definitions

- **AI_result**: output is 1 if consumer behavior is consistent with the modeled profile otherwise 0;
- **reputation**: variable representing the reputation of a consumer. The initial value is 50;
  - **min_reputation**: 0;
  - **max_reputation**: 100;

- **incremental_factor**: factor representing the weight of each request in order to change the reputation.
- **current_weight**: the current weight of the request increases according to the number of requests made previously;
  - **initial_weight**: 3;
  - **max_weight**: makes sure that a request does not overly impact reputation;
- **num_consumer_requests**: the total number of requests from the consumer.

3.7.2 Algorithm Steps

- **Calculation of current weight**:
  $current\_weight = current\_weight + incremental\_factor * num\_consumer\_requests$
- **Calculation of Reputation**:
  If $AI\_result = 1$ then

$$reputation = reputation + current\_weight$$

Else

$$reputation = reputation - current\_weight * 2$$

4 Evaluation

4.1 Overview

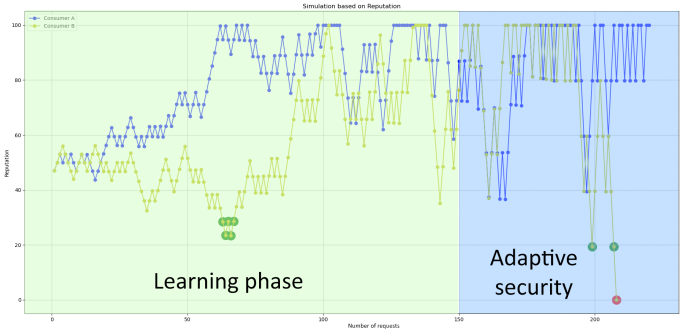
The scope of our experiment is to show how the approach with AI works in border routers with the analyzer.
The experiment considers two consumers (Consumer A and Consumer B) making a maximum of 220 requests that are not part of the blacklist in the signature approach database. Each consumer-related model is modeled according to the behavior of the consumer, and as we move forward with the requests, their weight becomes increasingly relevant.

4.2 Initial setup

- **Initial reputation**: 50
- **Incremental factor**: 0,001
- **Initial weight**: 3
- **Max weight**: 20
- **Max iterations**: 220

4.3 Simulation

To simulate the operation of the analyzer, we performed various experiments on Python until we arrived at the following graph that we think summarizes all the most important consideration to look at. We divided the graph into two macrophases in which all the different scenarios seen previously are present.



We can notice that among the two consumers, the one more inclined to anomalous behavior is Consumer B since in the learning phase it several times falls into scenario 3 (send the request to the SOC for additional checks), also in the adaptive security phase it falls directly into scenario 2 (drop packet). Only the SOC can later rehabilitate the consumer or leave it excluded from the network.

4.4 Phases

1) **Learning phase**: phase used to learn the characteristics of the consumer within the network.
2) **Adaptive security**: phase used to intercept and preemptively attentionalise requests that do not conform to the model created by the AI for the specific consumer

4.5 Consideration

- **Slow Start**: in the learning phase, we can see that the weight of requests that are not consistent with the model does not go to the consumer's reputation significantly because the model is still a suitable set for detecting suspicious behavior or otherwise. In fact, the weight changes according to the number of requests made;
  - **Advantages**: few false positives;
  - **Disadvantages**: requires learning time to confidently detect ambiguous requests.
- **AI features**: after several simulations, we summarized the advantages and disadvantages of integrating an AI model into a border router.
  - **Advantages**:
    - **Adaptive learning**: adaptive ability to scan requests;
    - **Simplification of the analyzing process**;
    - **Minimum maintenance**: only requires updating the AI but you don't have to go and develop the features of the model;
    - **Minimum maintenance**: only requires updating the AI but you don't have to go and develop the features of the model;
    - **Versatility of AI model**: does not require a single AI model, but each border router can use a different AI (local approach);
    - **Zero Trust approach**: always question the veracity of the consumer even after authenticating correctly in his device.

  - **Disadvantages**:
    - **Limitation in content control**: cannot, for example, learn the contents of media files;
    - **Forgery content name**: possibility of assigning a content a name not inherent to the content;
    - **Requires high-performance hardware**;
    - **Compliance and regulations**: consent to data treatment and security data.

- **Incremental weight**: in the adaptive security phase, it can be seen that after training the model, it is possible to gradually increase the weight of demands burdening reputation. Malicious requests in the adaptive security phase impact more than requests considered malicious in the learning phase.

5 Related work

Considering the approaches on reputation in NDNs, we noticed that control was being applied to contents but not to consumers. The only similar approach we found is in the FANET ad-hoc networks.
In contrast to a reputation-approach on a FANET network (such as the one in the third paper) which is based on direct or indirect interaction between UAVs (ad-hoc network), our approach is based on router-consumer interaction; the routers are in charge of "updating" the reputation of single consumer in the network based on its behavior. Unlike the FANET, in which each node in the network can contribute its feed to other nodes, thus being subject to possible attacks (such as spreading negative feedback towards non-malicious nodes and causing disruptions in the network), our system is more controlled, so that only routers deal with the reputation of consumers and there is no need for an exchange of information between them, thus reducing the load of data transmitted in the network.

6 Conclusions

Through our implementation, it is possible to have two simultaneous controls on the consumer. One is the **Active** control directly from an SOC and the second is the **Passive** control in which if the node reduces its reputation below the value 10 it is directly excluded from the network.
Summarizing we can considered several features that are added to the traditional NDN architecture:

- **Decentralized approach**: each border router has a local database that makes the right decisions independently;
  - **Advantages**:
    - Increased speed;
    - Less traffic on the network;
    - Independent failures;
    - Customized behavioral security.
  - **Disadvantages**:
    - Higher cost of hardware;
    - Increased maintenance;
    - Increased use of memory;
    - High computational load.

From a security perspective, we can see the following considerations:

- **Router-consumer interaction**: a consumer cannot alter the reputation of another consumer.
- **Hiding intentions**: a consumer can bypass the model by sending a malicious request for every so many consistent requests.
- **Training model**: the initial model must be trained with mostly non-malicious initial requests (we assume that at the beginning a user is not malicious).

In the future, artificial intelligence will be faster and more accurate so that our analyzer will be improved accordingly. In addition, the costs of hardware, such as memories, processors, will get lower and lower so that we will have better performance at a lower cost.

References

[1] Giovanna Carofiglio, Alberto Compagno, Mauro Conti, Fabio De Gaspari and Luca Muscariello, IaaS-Aided Access Control for Information-Centric IoT, 2018.
[2] Alberto Compagno, Mauro Conti, Paolo Gasti, Luigi Vincenzo Mancini and Gene Tsudik, Violating Consumer Anonymity: Geolocating Nodes in Named Data Networking, 2019.
[3] Ioanna Angeliki Kapetanidou, Paulo Mendes and Vassilis Tsaoussidis, Enhancing Security in Information-Centric Ad Hoc Networks, 2023.
[4] Amar Abane, Mehammed Daoui, Samia Bouzefrane, Soumya Banerjee and Paul Muhlethaler, A Realistic Deployment of Named Data Networking in the Internet of Things, 2019.
[5] Htet Htet Hlaing, Masahiro Mambo and Yuki Funamoto, Secure Content Distribution with Access Control Enforcement in Named Data Networking, 2021.
[6] Karen Scarfone and Peter Mell, Guide to Intrusion Detection and Prevention Systems (IDPS), 2007.
[7] The Rise of AI-Powered IPS Systems in 2023: Implications for Cybersecurity Professionals, 2023.
[8] Angelo Spognardi, Practical Network Defense Course - Slide IPS, Cybersecurity 2023,2024