Analyzing encrypted traffic using AI models

Louis Poidevin, Francoise Sailhan, Johanne Vincent

Analyzing encrypted traffic using AI models

A distributed and in-network approach

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

RÉPUBLIQUE FRANÇAISE
Liberté Égalité Fraternité

Affiliation

Lab-STICC

Authors

Louis Poidevin
Francoise Sailhan
Johanne Vincent

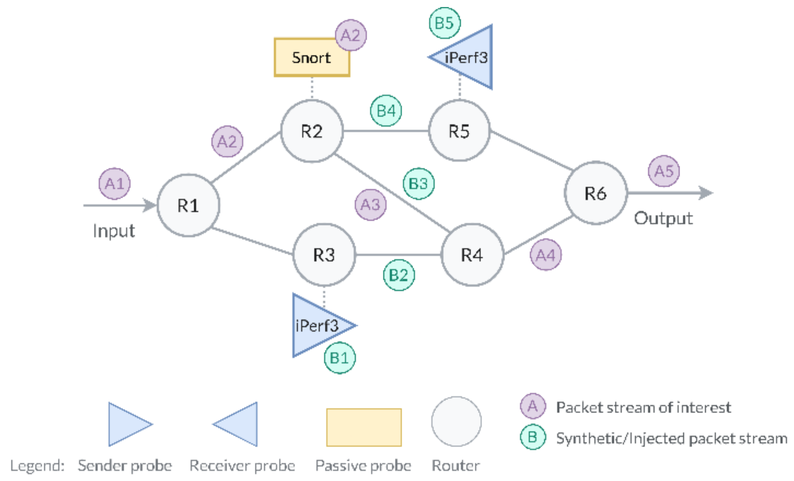
Funding

ME/CT PART

FRANCE 2030

Traffic Analysis

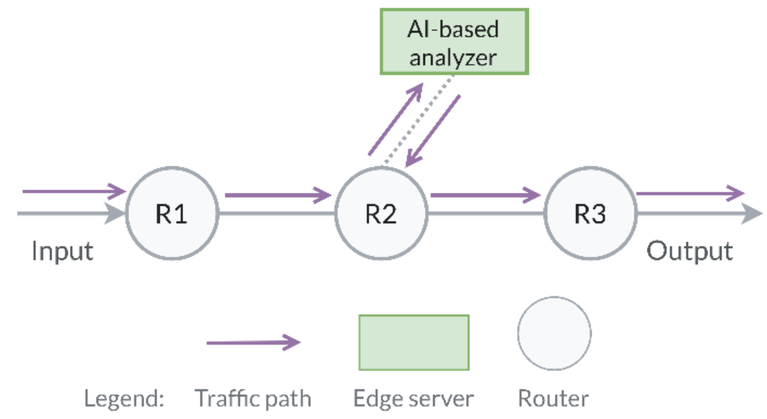
- Traffic analysis is essential for network operators to enhance **performance**, address **security** issues, manage **bandwidth**, and improve **Quality of Service (QoS)**.
- Traffic analysis methods include:
  - **Active monitoring**: generating and observing synthetic packet streams.
  - **Passive monitoring**: observing natural traffic streams without alteration.
  - **Hybrid approaches**: combining active and passive techniques for comprehensive insights.



Active vs. Passive Network Monitoring

Challenges of Encrypted Traffic Analysis

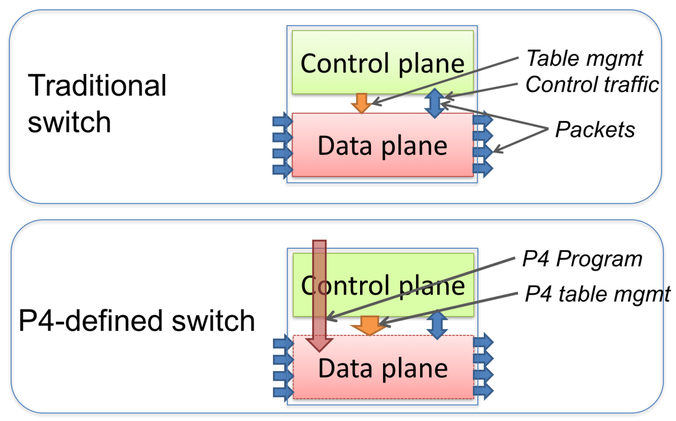
- The widespread encryption of internet traffic necessitates adaptation of network analysis tools.
- Traditional packet inspection methods are ineffective on encrypted traffic, requiring alternative approaches.
- **AI-based analyzers** offer solutions by classifying traffic based on packet statistics and behavioral patterns.
- However, AI-based analyzers have limitations in **throughput capacity** and **latency** due to their reliance on inference servers.



Traditional AI inference architecture

In-Network Computing: Pros and Cons

- To overcome AI model limitations, in-network computing deploys AI models directly on programmable network devices.
- This approach enables **high-performance**, **real-time processing** by offloading tasks from servers.
- Despite these advantages, programmable devices face significant constraints:
  - **Memory limitations** and restricted **data type support**.
  - Limited **computational capacity** optimized primarily for high-rate packet processing.
- Large AI models also present **coexistence challenges** with network operating systems (e.g., RARE).

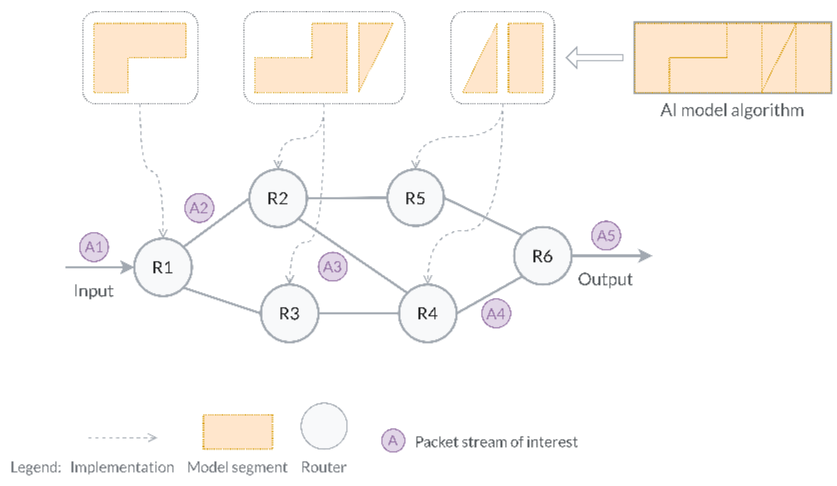


Traditional vs. programmable switches
Source: $P4_{16}$ Language Specification

Distributed and In-Network AI Inference for Encrypted Traffic Analysis

- To address resource constraints, [Zheng+23] propose to **decompose target programs** into segments and **distribute** them across network devices.
- We intend to apply this approach to our AI-based analyzer, achieving **line-rate traffic analysis** with minimal accuracy degradation.
- Our study will focus on the **QUIC protocol**, considered as a complex case in encrypted traffic classification.
- Additional investigations will address unique challenges in **data center** environments, **SDN/multi-tenant** networks, and assess the **energy efficiency** of the analyzer.

[Zheng+23] C. Zheng et al. *DINC: Toward Distributed In-Network Computing.* ACM CoNEXT, 2023.



Distributed in-network computing paradigm

November 2024 Winter training school Lille