## linea orizzontale



Riconoscimento di protocolli applicativi ignoti a ndpi tramite analisi del traffico e tecniche di fingerprinting avanzate

Titolo provvisorio della tesi 25/07/2025

**─**

Nicolò Fontanarosa

Università di Pisa

mat. 655436

# Panoramica

Strumenti come **nDPI** sono molto utili per identificare protocolli noti, ma si fermano generalmente a una classificazione generica (es. HTTPS, TLS, QUIC, ecc.) se non vengono integrate manualmente firme specifiche ( twitter, whatsapp etc … ). Tuttavia, molte applicazioni ( come VPN, proxy o app mobili ) utilizzano protocolli cifrati o personalizzati difficili da riconoscere senza analisi aggiuntiva

Attraverso l’uso di tecniche di fingerprinting ( come **SNI**, **JA3/JA3S**, **User-Agent** ) e regole personalizzate ( eventualmente anche tramite un LLM ), si vuole costruire un sistema che permetta di isolare e caratterizzare questi protocolli, anche se non noti a priori

# Obiettivi

Il progetto ha come obiettivo lo sviluppo di uno script che analizza file di traffico di rete in formato .pcap, con l’obiettivo di identificare protocolli applicativi utilizzati da app specifiche ( VPN, tunneling, browser particolari, etc … ) anche quando non sono riconosciuti direttamente da strumenti come **nDPI**

# Descrizione

Il progetto si divide in **tre fasi principali**:

### **1. Parsing e Pre-elaborazione dei file .pcap**

* Lo script prende in input un ( estendibile in futuro a più ) file .pcap
* Utilizza **nDPI** per una prima analisi del traffico e l’eventuale estrazione in formato **JSON**
* Viene effettuata una fase di **filtraggio**, per isolare flussi di traffico potenzialmente appartenenti a un singolo protocollo
* Questo filtraggio può basarsi su:
  + indirizzi IP o porte sospette
  + protocolli generici identificati da nDPI
  + pattern nel flusso ( SNI sospetti, quantità anomala di pacchetti )

### **2. Analisi e Identificazione del protocollo**

* Su ogni flusso filtrato vengono applicate tecniche di **fingerprinting**:  
  + **SNI (Server Name Indication)** del TLS
  + **JA3 / JA3S** fingerprint (impronta del Client Hello o Server Hello)
  + **User-Agent** nel caso di traffico HTTP
* I dati ottenuti vengono confrontati con una base di conoscenza (inizialmente anche manuale) di app note (esempi: debian.run, cdnexpress.org, fingerprint come f4febc55ea12b31ae17cfb7e614afda8, ecc.)
* In futuro si potrebbe anche applicare una logica tramite **script procedurali** o modelli **LLM** per riconoscere pattern o inferire il tipo di applicazione/protocollo.

### **3. Generazione di Regole e Report**

* Per ogni protocollo identificato, lo script:  
  + genera una **firma identificativa** (es. SNI fisso, fingerprint JA3, dimensioni dei pacchetti, tempo tra pacchetti, ecc.)
  + produce una **regola descrittiva**, che può essere usata in futuro per identificare traffico simile
* Nel caso di **protocolli ignoti**, lo script fornisce un report con:  
  + il volume di traffico osservato
  + eventuali pattern parziali (es. “questo traffico TLS ha JA3S X ed SNI Y, ma non è riconducibile con certezza a un'app nota”)
  + possibilità di aggiornare il database delle firme

**─**

hoxx\_vpn\_navigazione\_google\_chrome.pcap | grep SNI=[debian.run](https://debian.run/) io ho trovato SNI=account2.hoxx.com [JA3S: f33734dfbbff29f68bcde052e523c287]

cattura\_http\_injector.pcap | non c'ho troato nulla di particolare catture\_ha\_tunnel\_plus.pcap | JA3S: f4febc55ea12b31ae17cfb7e614afda8 cattura\_[hide.me](https://hide.me/)\_vpn.pcap | JAS3S: f4febc55ea12b31ae17cfb7e614afda8

io ho trovato

browsec\_vpn\_navigazione\_google\_chrome\_switch\_from\_netherlands\_to\_usa.pcap | SNI=[cdnexpress.org](https://cdnexpress.org/)