Metadata extraction from Network Traffic

@ NFSU, 26 APRIL 2025

-PRESENTED BY JALAJ BHASKAR

Who am I :

- Cybersecurity Enthusiast
- Various roles –
  - Security Operation Ceter Analyst (SOC)
  - Detection and Response Team
  - Security Engineer
  - Researching on Defensive AI
- Reach out to me at ..
  - Linkedin.com/in/teamblue

Today's Agenda

- Introduction to network traffic and metadata
- Why extract metadata?
- Basic networking concepts
- Tools for extraction
- Step-by-step demo using Wireshark
- Best Practices
- Demo
- QnA

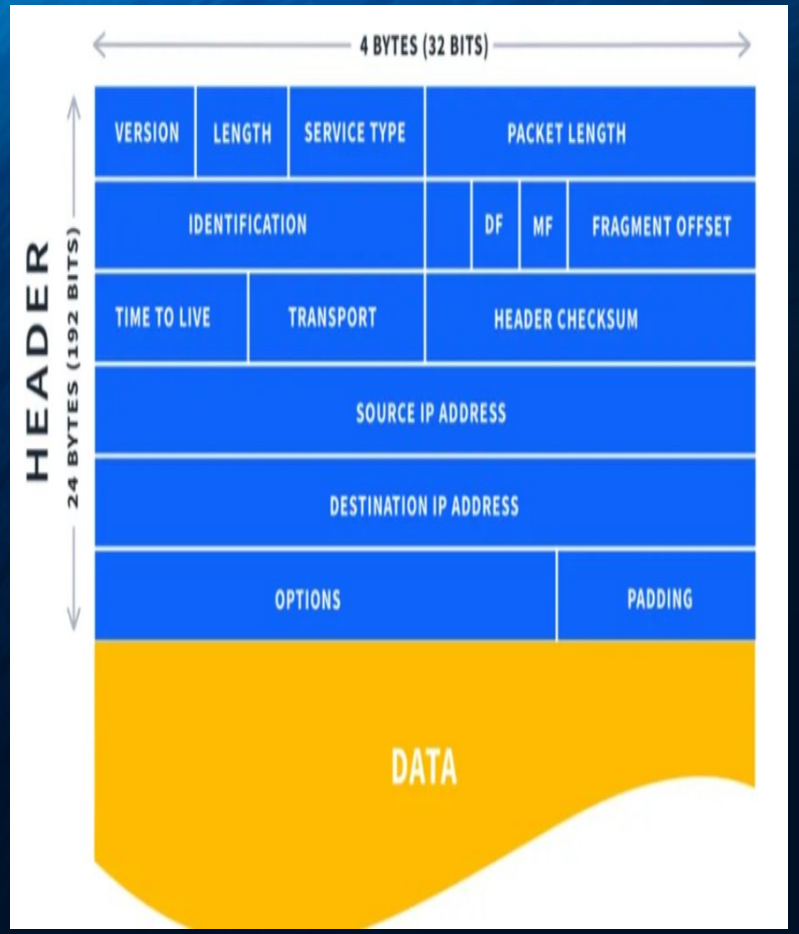
What is Network Traffic

Network traffic: Data moving between devices over the internet or local networks (like cars on a highway)

Examples: Loading a webpage, sending an email, streaming video

Traffic is made of "packets" – small chunks of data

Packets are like envelopes: Address on outside, letter inside



What is Metadata?

Metadata: "Data about data" – describes the packet without revealing the content

- Not the message itself (payload), but details like:
  - Who sent it (source IP address)
  - Who received it (destination IP address)
  - When it was sent (timestamp)
  - How it's sent (protocol, like HTTP for web)

Like envelope details (sender, receiver, postmark) vs. the letter inside

Metadata vs. Content

- Metadata: Headers in the packet (**quick to extract**, privacy-friendly in some cases)
- Content (Payload): Actual data (e.g., email text, photo) – **requires deeper inspection**
- Why focus on metadata? It's **efficient**, **legal** in many scenarios, and reveals patterns **without invading privacy**

Aspect	Metadata	Payload (Content)
Definition	Data about the packet- used for routing, handling, and analysis	Actual data being transmitted- the message or file itself
Examples	Source/Destination IP, Protocol, Port Number, Packet Length, Timestamp	Webpage content, Email body, Chat messages, Uploaded files
Extraction	Easy- available in packet headers	Harder- may require decryption or parsing
Use Cases	Traffic analysis, flow monitoring, intrusion detection, routing decisions	Deep packet inspection, malware analysis, data reconstruction
Visibility	Visible to network devices and monitoring tools	May be encrypted or obfuscated
Security Role	Helps identify anomalies, spoofing, scanning attempts	Reveals intent, data exfiltration, or malicious payloads

Key Metadata Elements

1. Source/Destination IP: Like home addresses
2. Ports: "Doors" on devices (e.g., 443 for secure web)
3. Protocol: Type of traffic (TCP for reliable, UDP for fast)
4. Timestamp: When packet was sent/received
5. Packet Size/Length: How much data
6. Flags: Special signals (e.g., SYN for starting connection)

Key Metadata Elements

No.	Time	Source	Destination	Protocol	Length	Info
40	0.160470	10.6.13.3	10.6.13.133	TCP	60	88 → 52428 [ACK] Seq=1 Ack=1910 Win=1049600 Len=0
41	0.163694	10.6.13.3	10.6.13.133	TCP	1514	88 → 52428 [ACK] Seq=1 Ack=1910 Win=1049600 Len=1460 [TCP PDU reassembled in 42]
42	0.163697	10.6.13.3	10.6.13.133	KRB5	471	TGS-REP
43	0.163760	10.6.13.133	10.6.13.3	TCP	60	52428 → 88 [ACK] Seq=1910 Ack=1878 Win=65280 Len=0
44	0.163794	10.6.13.133	10.6.13.3	TCP	60	52428 → 88 [FIN, ACK] Seq=1910 Ack=1878 Win=65280 Len=0
45	0.164009	10.6.13.3	10.6.13.133	TCP	60	88 → 52428 [ACK] Seq=1878 Ack=1911 Win=1049600 Len=0
46	0.164010	10.6.13.3	10.6.13.133	TCP	60	88 → 52428 [RST, ACK] Seq=1878 Ack=1911 Win=0 Len=0
47	0.165092	10.6.13.133	10.6.13.3	TCP	1514	52427 → 389 [ACK] Seq=351 Ack=2837 Win=65280 Len=1460 [TCP PDU reassembled in 48]
48	0.165094	10.6.13.133	10.6.13.3	LDAP	732	bindRequest(11) "<ROOT>" sasl
49	0.165096	10.6.13.3	10.6.13.133	TCP	60	389 → 52427 [ACK] Seq=2837 Ack=2489 Win=1049600 Len=0
50	0.166729	10.6.13.3	10.6.13.133	LDAP	265	bindResponse(11) success
51	0.167445	10.6.13.133	10.6.13.3	LDAP	129	SASL GSS-API Privacy: payload (11 bytes)
52	0.167602	10.6.13.3	10.6.13.133	TCP	60	389 → 52427 [RST, ACK] Seq=3048 Ack=2564 Win=0 Len=0
53	0.167634	10.6.13.133	10.6.13.3	TCP	60	52427 → 389 [FIN, ACK] Seq=2564 Ack=3048 Win=65280 Len=0
54	0.167635	10.6.13.3	10.6.13.133	TCP	60	389 → 52427 [RST] Seq=3048 Win=0 Len=0
55	0.444345	10.6.13.133	224.0.0.252	LLMNR	75	Standard query 0x3369 ANY DESKTOP-5AVE44C
56	0.613251	Intel_ac:97:df	Broadcast	ARP	60	Who has 169.254.48.3? (ARP Probe)
57	0.862738	10.6.13.133	10.6.13.255	NBNS	110	Registration NB MASSFRICTION<00>
58	0.862740	10.6.13.133	10.6.13.255	NBNS	110	Registration NB DESKTOP-5AVE44C<00>

Why Extract Metadata?

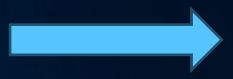
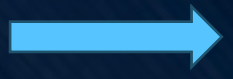
Security: Detect unusual patterns (e.g., many connections from one IP)

Analytics: Understand usage (e.g., most visited sites)

Forensics: Investigate incidents without full data

Basic Networking Concepts OSI Model

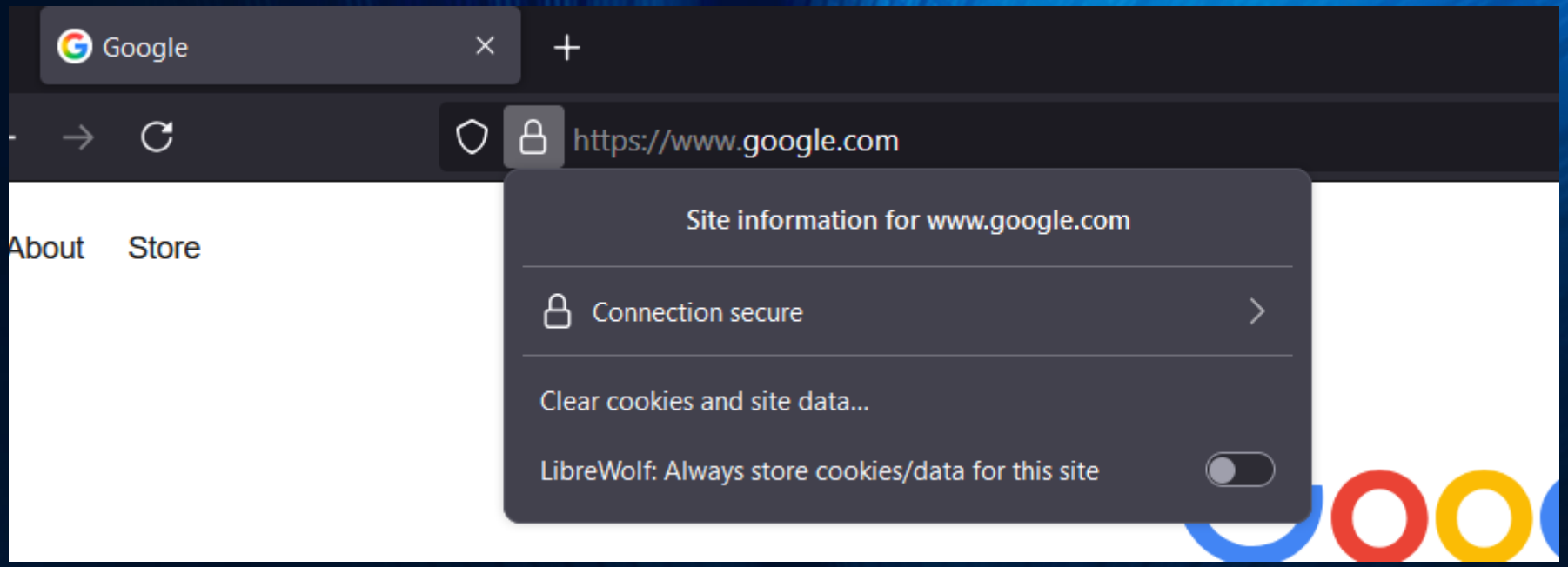
OSI Layer	Metadata Examples	Purpose of Metadata
Layer 1: Physical	Bit timing, voltage levels, modulation scheme	Defines how bits are physically transmitted over the medium
Layer 2: Data Link	MAC addresses, frame type, CRC, VLAN tags	Local delivery, error detection, and framing
Layer 3: Network	Source/Destination IP, TTL, Protocol ID	Routing, addressing, and packet forwarding
Layer 4: Transport	Source/Destination Port, Sequence/ Ack numbers, Flags (SYN, ACK, FIN)	Session management, reliability, flow control
Layer 5: Session	Session ID, authentication tokens	Establishing, maintaining, and terminating sessions
Layer 6: Presentation	Encoding type, compression method, encryption scheme	Data format negotiation, security, and translation
Layer 7: Application	HTTP headers, SMTP commands, DNS query types	Application-specific control info (e.g., content type, method)

Basic Networking Concepts OSI Model

TCP	UDP
Transmission Control Protocol	User Datagram Protocol
Connection-Oriented	Connectionless
Slower	Faster
20 Bytes Header	8 Bytes Header
Packet Reorder Mechanism	No Reorder Mechanism
Uses Acknowledgement	No Acknowledgement
Error Checking and Recovery	Basic Error Checking
Guaranteed delivery, reliable	No guarantee, unreliable
Uses SSL/TLS for security	Uses DTLS for security
HTTP, HTTPS, FTP etc.	DHCP, TFTP, SNMP etc.

HTTP vs HTTPS



TLS 1.3 goes further:

- **Encrypts more of the handshake,** hiding details like the server's certificate from passive observers (unless using techniques like TLS fingerprinting).
- Makes **passive monitoring even harder**.
- Tools like deep packet inspection (DPI) become less effective.

Tools for Metadata Extraction

- Wireshark: Graphical, user-friendly for beginners
- tcpdump: Command-line, lightweight
- tshark: Wireshark's command-line version

Getting started with WireShark

Techniques for Extraction –
- Capture traffic: "Sniff" packets on your network
- Filter: Use queries like "ip.src == 192.168.1.1" to focus
- Export: Save metadata to CSV for analysis

Let’s see it in action

DEMO

Privacy Consideration

- - Important: Only analyze your own traffic or with permission
- - Privacy laws: GDPR, CCPA – metadata can still identify people
- - Best practices: Anonymize data, secure captures

Key takeaways

- Metadata is essential info from packet headers
- Tools like Wireshark make extraction easy
- Always prioritize ethics and privacy

Resources for Further Learning -

- Malware PCAP analysis
- Github repo - Post Exploitation Techniques
- Wireshark PCAP
- Youtube – Chris Greer
- Wireshark tutorials
- tShark Manual

- Get your files here …

Have any questions ? Reach out …

Mob – 88 514 584 52

LinkedIn – Linkedin.com/in/teamblue