1.11 Data interception

Data interception

In this section, we'll look at data interception and traffic analysis. It is possible to highlight three types of traffic analysis: protocol analysis, packet analysis and flow analysis. Let's see the protocol analysis first.

Protocol analysis

To start analysing traffic, first, we need to understand protocols.

Protocol

A set of rules or instructions that determine how 'actors' act or interact in a given situation.

Examples include:

1. a recipe for making a cake
2. the procedure for boarding a plane
3. the steps for buying a movie ticket.

A communication protocol is a system of rules that allow two or more entities of a communication system to transmit information. Once in possession of the information, we might want to understand the semantics of the information being transmitted to be able to interpret it. From a security point of view, analysing protocols might allow the discovery of potential vulnerabilities that an attacker can use.

There are tons of different network protocols such as IP, TCP, UDP, HTTP, HTTPS, SMTP and others. The specification of many of them is public, but some are proprietary and may need specialist tools or reverse engineering to access. We need the protocol to understand the semantics of the information being transmitted to be able to interpret it. Support for public protocols is usually implemented in tools like Wireshark, which shows headers, flags, content in a more user-friendly and navigable way.

Once we have decided that we are interested in a specific protocol, we can capture all the packets that belong to that specific protocol and begin a more fine-grained analysis: packet analysis.

Packet analysis

Network professionals use packet analysis to monitor the health of a network and security professionals use it to conduct passive network vulnerability assessments. Passive in this context means inspection without altering the contents. On the other hand, attackers can use the same techniques and tools to steal information such as passwords.

However, in packet analysis the word packet is misleading. What is actually captured and analysed are the **frames**. Frames are what carry packets in a local network, so during packet analysis results frame are always mentioned and frame details are always provided in addition to the packet payload. This diagram shows the relationship between data, packets and frames:

A diagram showing that data are contained within segments which carry the TCP protocol and port number. Segments are then contained in packets which hold the sender and receiver IP addresses. Finally, packets are contained in frames which hold the sender and receiver MAC addresses.

There are several techniques that we can use to analyse the packets we collected.

- **Pattern matching:**
  Identify packets of interest by matching specific values within the packet capture.
- **Parsing protocol fields:**
  Extract the contents of particular protocol fields.
- **Packet filtering:**
  Separate packets based on the values of fields in protocol metadata.

Flow analysis

Another technique used for data interception is **flow analysis**. Flow analysis is the practice of examining related groups of packets in order to:

- identify patterns such as repeated communications
- isolate suspicious activity and discard irrelevant data
- analyse higher-layer protocols, for example reconstructing segment TCP packets and getting the full picture of the protocol encapsulated in it: HTTP, SSL, etc
- extract data such binary files to be analysed further.

There are several flow analysis techniques.

- **List conversations and flows:**
  List all conversations and/or flows within a packet capture or only specific flows based on their characteristics.
- **Export a flow:**
  Isolate a flow or multiple flows, and store the flow(s) of interest to disk for further analysis.
- **File and data carving:**
  Extract files or other data of interest from the reassembled flow.

You will have an opportunity to use some of these techniques and tools in the virtual labs.