Mobile Networks and GDPR

DISCLAIMER : NEITHER THE WHOLE NOR ANY PART OF THE INFORMATION CONTAINED IN, MAY BE ADAPTED OR REPRODUCED IN ANY MATERIAL OR ELECTRONIC FORM WITHOUT THE PRIOR WRITTEN CONSENT OF OMNIPACKET. THE INFORMATION CONTAINED IN THIS DOCUMENT IS SUPPLIED ON AN AS-IS BASIS AND NO WARRANTY AS TO IT'S SUITABILITY FOR ANY PARTICULAR PURPOSE IS EITHER MADE OR IMPLIED. OMNIPACKET WILL NOT ACCEPT ANY CLAIM FOR DAMAGES HOWSOEVER ARISING AS A RESULT OF USE OF ANY INFORMATION PROVIDED HERE.

Introduction

With GDPR becoming a law in 2018 the EU based companies and companies offering goods or services within the region have been working on ensuring the compliance. Most have learned the task is a difficult one. Mobile network operators are no exception.

Sharing Network Captures

Companies have to have a consistent and reliable way of sharing the captured network data across organizational and geographical boundaries while preserving the security and privacy of the users. The procedures for ensuring legal compliance with GDPR regulations in EU and NISTIR 8053 in US have to be put in place and adhered to. This is important not only from the operational perspective, but also to protect a brand reputation and, in light of the GDPR’s potential penalties, the bottom line.

Technical aspects of pcap data management

Computer networks are at the very core of most modern businesses. Packet Capture (pcap) files are de-facto industry standard for the networks data capture, storage and exchange. From business development to engineering, diagnostics, maintenance and support, pcap files are widely used everywhere.

Selectively scrambling the captured network traffic while fully preserving its binary integrity is a complex task due to a great variety of network protocols and the intricacy of packets' intra and inter-dependencies. The existing approaches to anonymizing privacy information from captured network traffic break binary integrity of captured packets, making it hard to analyze the anonymized data.

While most networks carry user data in need of anonymizing, mobile networks are “even more so” as they carry users unique ids (phone numbers, etc.) and location information. Below we’ll briefly describe Omnipacket approach to pcap anonymization using Mobile Core Networks as the example.

Anonymization Process

Mobile Core networks created over the last 30 years were designed with little concern over privacy. User data are spread across many data fields at many stack layers and often travels over Mobile Core networks un-encrypted. There is no industry standard list of the privacy related data to be anonymized and organizational requirements for the anonymization vary greatly.

Some of the data fields that are commonly require anonymization are: MSISDN, IMSI, IMEI, IP addresses, location info, SMS content, etc. In practice, a customer may request any field at any networking stack layer to be anonymized. To quickly address changing requirements, Omnipacket has created a universal packet editor (more on that below).

The steps involved in the process using MSISDN number anonymization as the example:

1. Parse all captured packets as we have to check them all for instances of MSIDN data.
2. Decide whether a decoded packet may have the data of interest and if so, find the exact place where MSISDN data resides.
3. Generate the MSISDN number Y to replace the original MSISDN number X, and make sure next time we find MSISDN X, we’ll replace it with MSISDN Y.
4. Re-encode the packet with Y value of MSISDN while making sure the rest of the data remained the same and the packet remains syntactically correct.

Step 1:

This step requires a packet decoder. Wireshark(TM) is the industry standard for packet decoding, however its dissectors are designed for decoding only, and couldn’t be used to encode the modified packets back as required by Step 4. Omnipacket has developed software architecture which supports both encoding and decoding.

Step 2:

Captured mobile network traffic is often a mixture of data carried over different networking stacks and interfaces. Private user data can be found in many of them, at different stacks layers, sometimes in unexpected spots. The search code is non-trivial and requires a detailed knowledge of the relevant 3GPP specs, and details of the implementations by various vendors.

Step 3:

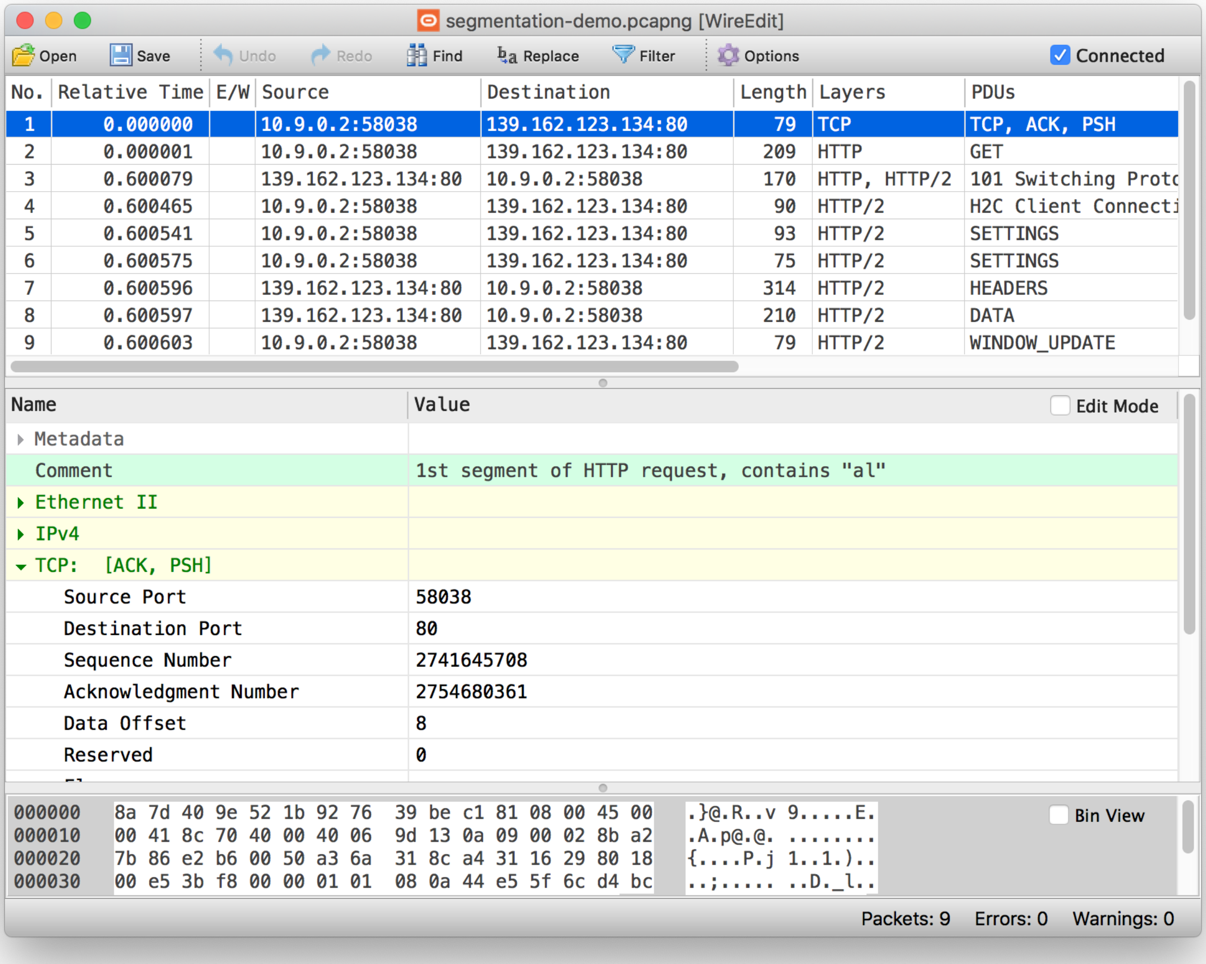
The algorithm makes sure every time we find MSISDN number X, we’ll replace it with Y.

Step 4:

We have to rebuild the anonymized packet so that all the data except MSISDN remains unchanged and the anonymized packet can be dissected with Wireshark(TM) or similar tools without errors.

Omnipacket Universal Packet Editor

Steps 1 and 4 taken together are supported by Omnipacket universal packet editor. WireEdit by Omnipacket is a desktop version of the editor. One can think about WireEdit as Microsoft Word(TM) for network packets. SafePcap uses the same packet editing engine.



No.	Relative Time	E/W	Source	Destination	Length	Layers	PDUs
1	0.000000		10.9.0.2:58038	139.162.123.134:80	79	TCP	TCP, ACK, PSH
2	0.000001		10.9.0.2:58038	139.162.123.134:80	209	HTTP	GET
3	0.600079		139.162.123.134:80	10.9.0.2:58038	170	HTTP, HTTP/2	101 Switching Protc
4	0.600465		10.9.0.2:58038	139.162.123.134:80	90	HTTP/2	H2C Client Connecti
5	0.600541		10.9.0.2:58038	139.162.123.134:80	93	HTTP/2	SETTINGS
6	0.600575		10.9.0.2:58038	139.162.123.134:80	75	HTTP/2	SETTINGS
7	0.600596		139.162.123.134:80	10.9.0.2:58038	314	HTTP/2	HEADERS
8	0.600597		139.162.123.134:80	10.9.0.2:58038	210	HTTP/2	DATA
9	0.600603		10.9.0.2:58038	139.162.123.134:80	79	HTTP/2	WINDOW_UPDATE

Name	Value
▸ Metadata	
Comment	1st segment of HTTP request, contains "al"
▸ Ethernet II	
▸ IPv4	
▾ TCP: [ACK, PSH]	
Source Port	58038
Destination Port	80
Sequence Number	2741645708
Acknowledgment Number	2754680361
Data Offset	8
Reserved	0

```
000000   8a 7d 40 9e 52 1b 92 76  39 be c1 81 08 00 45 00   .}@.R..v 9.....E.
000010   00 41 8c 70 40 00 40 06  9d 13 0a 09 00 02 8b a2   .A.p@.@. ........
000020   7b 86 e2 b6 00 50 a3 6a  31 8c a4 31 16 29 80 18   {....P.j 1..1.)..
000030   00 e5 3b f8 00 00 01 01  08 0a 44 e5 5f 6c d4 bc   ..;..... ..D._l..
```

Packets: 9 Errors: 0 Warnings: 0

SafePcap customers often have different ideas on which data carried by network packets have to be anonymized. Having a universal packet editing engine allows quick accommodation of evolving customer requirements.

Finding the data to anonymize

The difficulty in implementing this step is due to the fact that user data is carried by many networking stacks at different layers/fields. Omnipacket engineering team have accomplished the laborious task of investigating mobile core protocols and mapping the common privacy data fields locations. In other words, SafePcap not only knows what data to look for but also where to look for it. New network stacks and data fields are added to the anonymization engine's map as the 3GPP specs and RFCs become available.

Anonymization Algorithm

When a data field (MSISDN for example) to be anonymized is found by SafePcap, and the value of it is X, the decision has to be made what algorithm(s) should be used to generate a new value (Y) to replace X with. Different users have somewhat different ideas how this should be done. A common requirement for the anonymization algorithms is a consistency. Data field named A with value X should always be anonymized to the value Y of the same type. For example, a MSISDN number X should be converted to MSISDN number Y. Most agree that the number of digits should remain the same as well.

One of the frequent requirements is maintaining the consistency and congruence of all the sessions. For example each specific MSISDN has to be obfuscated in the same way for all the sessions belonging to that MS subscriber together with all the associated info (MSISDN, IMEI,…). Additionally, some of the users need to maintain this consistency across multiple pcap files belonging to the same original capture.

Attention is also required when choosing the substitution algorithm. Some users need to replace MSISDN number X by a randomly generated number Y consistently across a pcap file. Other cases require to replace MSISDN number only in a partial way, for example by keeping the country code intact. Yet another case required SafePcap instead of using a random number generation to take the phone numbers and other privacy related values from a data base and use it for the substitution.

IP-address subnet consistency is a desirable feature as well. When anonymizing two different IP addresses, the network mask prefix bits common for two IP addresses are converted to the same bits in the anonymized addresses. As a result subnets, even unknown, are preserved.

While in most cases the anonymization algorithm has to be non-reversible, there may be instances where the reversibility of the processed files is desired.

SafePcap is an ultimate pcap anonymization solution for all networking stacks, designed to quickly accommodate evolving customer requirements.