

Pseudonymisation of SS7 Identifiers by Random Tables

Charles Beumier Thibault Debatty
Royal Military Academy, Brussels, Belgium

1. Pseudonymisation

- = Technique for data privacy protection replacing values by pseudonyms to hide identity but keep distinctiveness
- ▶ GDPR imposes to hide subscriber identifiers in data
 - ▶ Our data analysis needs uniqueness of identifiers
 - ▶ Method: Look Up Table to replace identifiers with unique random values

2. Organisation

- ▶ Binding decision from BIPT to MNOs for data
- ▶ Data collection and pseudonymisation by MNOs
- ▶ Data transfer to RMA by hardware-encrypted disk and storage on secured server

3. SS7

- ▶ *Signalling System No. 7* is a stack of protocols between equipment in the 2G & 3G network core.
- ▶ Still in use for 2G/3G devices and where 4G & 5G coverage lacks
- ▶ Possible SS7 attacks: Subscriber location, SMS or Call interception, Denial of service, Fraud
 - Project DAP20-01: "Security issues related to SS7"

4. SS7 Identifiers

- ▶ SS7 traffic is composed of messages exchanged between network devices, containing:
 - Time, Operation Code, message size, ...
 - SS7 Identifiers for Subscribers or 2G/3G equipment

Identifier	Object	Format Land MNO User	Example
Global Title	Equipment, routing	e.164 CC NDC SN	32 475 000006
Global Title	Subscriber, routing	e.214 CC NDC MSIN	32 475 1234567890
IMSI	Subscriber (internal)	e.212 MCC MNC MSIN	206 01 1234567890
MSISDN	Subscriber (public)	e.164 CC NDC SN	32 475 123456

- ▶ Subscriber identifiers (IMSI, MSISDN) must be pseudonymised
- ▶ Global Titles also, as they may contain subscriber identifiers
 - Other fields with Global Titles must be pseudonymised



5. Random Table for last 5 digits

- ▶ Mapping of 5-digit -> 5-digit (all unique values)
 - 5 last digits, because in user part of identifiers
 - First digits (Country & Operator) unchanged, since needed
 - Memory limitation (100.000 entries for 5-digit)
- ▶ $99999 \times 99998 \times 99997 \times \dots \times 2 \times 1$ possible tables
 - ▶ One known table entry does not leak other entries

6. Implementation

- ▶ Algorithm forcing Change & Unicity
 - Init: for i in range(100.000): $table[i] = i$
 - Swap: $swap(table[i], table[random])$
 - Force change: $table[i] \neq i$, so pseudonym $\neq$ original
- ▶ One different Random Table for each identifier
 - Risk: Reidentification by crossing info with other databases
 - Possible correlation IMSI - MSISDN (from black market or in some messages)
 - Global Title is an equipment number OR an IMSI/MGT !

7. Testing Pseudonymisation

- ▶ Compare original and pseudonymised files
 - Files only differ at pseudonymised 5-digit
 - Many files since huge data (1 TB)
- ▶ Fill tables thanks to (original, pseudonym) 5-digit
- ▶ Test FAILS if any entry of any table:
 - gets a second value (missing unicity)
 - equals its index (missing a value change)

Glossary

BIPT	Belgian Institute for Postal Services and Telecommunications
GDPR	General Data Protection Regulation
IMSI	International Mobile Subscriber Identity
MNO	Mobile Network Operator
MSISDN	Mobile Station International Subscriber Directory Number

References

[1] The project "Security Issues related to SS7 and Diameter" is funded by MDN / RHID [DAP20-01, Jul2020-Jun2025].

