Qualcomm

Qualcomm Technologies, Inc.

DIRBS Operator Data Requirements

Formatting and Providing Operator Data Dumps

80-GD079-4 Rev. A

April 30, 2018

Revision history

Revision	Date	Description
A	April 2018	Initial release

Contents

1 Introduction

This document provides mobile network operators with information on how to format and provide data dumps for ingestion and analysis by DIRBS.

1.1 References and standards

Table 1-1 References and standards

Document	URL
Telecommunication management; Charging management; Circuit Switched (CS) domain charging, 3GPP TS 32.250	https://portal.3gpp.org/
Telecommunication management;Charging management;Packet Switched (PS) domain charging, 3GPP TS 32.251	https://portal.3gpp.org/

2 Required Data Fields

All necessary fields are available from output charging data records (CDRs). Table 2-1 shows field details and formatting.

Table 2-1 Required data fields

Field	CDR source fields	Format
Date	▪ Converted from date portion of the following, aggregated by local time: ▫ Seizure Time/Answer Time (Record Types 0, 1, 87) – Seizure Time is used for unsuccessful calls. – Answer Time is used for successful calls. – Use Answer Time if available, otherwise Seizure Time. ▫ Event Timestamp (Record Types 6-7, 21-23, 25, 28, 93-94, and IMEI Observation Ticket) ▫ Record Opening Time (Record Types 18, 20, 84-85, 96)	YYYYMMDD* (e.g. 20160423) *Aggregated by local time.
IMEI	▪ Served IMEI	14-16 digits* (e.g. 013845000153547) *Includes any leading zeros. If an IMEI starts with one or more zero digits, these leading zeroes must not be stripped off; doing so would completely change the IMEI value.
IMSI	▪ Originator IMSI (Record Type 93) or Recipient IMSI (Record Type 94) ▪ Served IMSI (otherwise)	14-15 digits
MSISDN	▪ Originator MSISDN (Record Type 93) or Recipient MSISDN (Record Type 94) ▪ Served MSISDN (otherwise)	Up to 15 digits* (e.g. 18583551234) *In E.164 format, i.e., international telephone number, including country code but excluding any international call prefixes, spaces, or separators.

Field	CDR source fields	Format
RAT	- System Type (Record Types 0-1, 6-7, 23, 25) - RAT Type (Record Types 18, 20-22, 28, 84-85, 93-94, 96) - At the time of writing, defined RAT values include: 001 – UTRAN (3G), 007 – Virtual 002 – GERAN (2G), 101 – IEEE 802.16 (WiMAX) 003 – WLAN, 102 – 3GPP2 eHRPD (3.5G) 004 – GAN 103, 3GPP2 HRPD (3G) 005 – HSPA Evolution (3.5G), 104 – 3GPP2 1xRTT (2G) 006 – E-UTRAN (4G), 105 – 3GPP2 UMB (4G)	Pipe-separated list of 3 digit codes, with leading zeroes intact

Data can be sourced from different fields in different types of CDRs produced in the operator network, e.g., SMS, packet data, voice call, etc. The CDR source fields identify which field in each of the possible CDR types (identified by 3GPP Record Type value) contains the relevant information.

The intention with listing many different record types is to obtain a complete view, capturing as many IMEIs as possible, regardless of the kind of chargeable activity in which they engaged. Once data is aggregated, the different source record types and fields that were used will not be apparent.

3 Excluded CDRs

In general, multiple CDR Record Types are intended to contribute to the final MNO data dump (see Table 2-1). However, some CDRs may be generated for IMEIs that are blocked by the EIR. If these are included by the MNO, they may appear incorrectly as blacklist violations (activity by devices which should have been blocked) in subsequent reporting.

The following cases should be excluded from the file submitted to DIRBS:

- **Emergency calls:** Emergency calls are allowed regardless of blacklist status. They may be identified by the Emergency Calls teleservice in the Basic Service field of:

  □ A Mobile Originated Call or MSC-SRVCC Record (see Telecommunication management; Charging management; Circuit Switched (CS) domain charging (3GPP TS 32.250)), or

  □ IMSI Unauthenticated Flag in PS CDR (see Telecommunication management;Charging management;Packet Switched (PS) domain charging (3GPP TS 32.251))

- **CDRs for attempts blocked by the EIR:** In case a CDR is generated for an attempt by a UE that is actually unsuccessful due to EIR blacklisting, such records may be identified by Illegal Equipment in the Diagnostics field, or by correlation with the associated blacklisted IMEI Observation ticket using the Call Reference field (see *Telecommunication management; Charging management; Circuit Switched (CS) domain charging* (3GPP TS 32.250)).

4 Data Aggregation from CDRs

Each row in an operator data dump represents an aggregation of CDR fields comprised of:

■ A distinct Date-IMEI-IMSI-MSISDN combination

■ A list of distinct RAT values used by that combination

This list of distinct RAT values is pipe delimited, i.e., a concatenated list of values with each value separated by a pipe ('|') character. Blank or missing fields shall be considered as distinct values and included as such.

Table 4-1 shows a set of example fields before aggregation.

Table 4-1 Pre-aggregation

Date	IMEI	IMSI	MSISDN	RAT
20160130	35780502398494	310150123456789	18585551234	001
20160130	35780502398494	310150123456789	18585551234	002
20160130	35780502398494	310150123456789	18585551234	006
20160130		310150123456789	18585551234	001
20160131	35780502398494	310150123456789	18585551234	001
20160131	35780502398494	310150123456789	18585551234	002
20160131	35780502398494	310150123456790		001
20160131	35780502398494	310150123456790		001

Based on these example fields, Table 4-2 shows the data dump rows created after aggregation

Table 4-2 Post-aggregation

Date	IMEI	IMSI	MSISDN	RAT
20160130	35780502398494	310150123456789	18585551234	001|002|006
20160130		310150123456789	18585551234	001
20160131	35780502398494	310150123456789	18585551234	001|002
20160131	35780502398494	310150123456790		001

5 Exporting Aggregated Data to CSV

Data shall be exported to a CSV text file using UTF-8 character encoding.

To ensure correct import, a header line shall be included in the CSV that identifies the fields being provided. Each record (including the header line) shall be located on a separate line with a CR/LF ending (/r/n). Data fields in each line shall be separated by a comma character.

For more information on CSV format, see www.ietf.org/rfc/rfc4180.txt.

Example CSV formatted data

```
Date,IMEI,IMSI,MSISDN,RAT
20160130,35780502398494,310150123456789,18585551234,001|002|006
20160130,,310150123456789,18585551234,001
20160131,35780502398494,310150123456789,18585551234,001|002
20160131,35780502398494,310150123456790,,001
```

6 Transferring Data to DIRBS

Each operator shall securely upload data to DIRBS during their assigned time window using their provided credentials. Time windows and credentials shall be provided to each operator by the DIRBS operational entity for that country.

Prior to upload, operators should validate their data format using the provided schema and open source tool (identified in the schema). Performing this validation step can help ensure successful processing of the data and cut down on roundtrips.

Files transferred by each operator shall be zipped to provide efficient transfer and enable detection of corruption due to network connection failures.

Table 6-1 shows the ZIP file details.

Table 6-1 ZIP file details

ZIP file name	The ZIP file name shall be comprised of the operator name followed by start and end dates in YYYYMMDD format. The dates shall be in local time and define the date range for the data inclusively, i.e., data in the file shall include both start and end dates: OperatorName_OptionalRegion_StartDate_EndDate.zip (e.g. Foo Wireless_Zone4_20160101_20160131.zip)
ZIP file contents	Each ZIP file shall contain only one CSV file, with the exception of the file extension, the CSV filename shall be the same as the ZIP filename: OperatorName_OptionalRegion_StartDate_EndDate.csv (e.g. Foo Wireless_Zone4_20160101_20160131.csv)
File security	Files shall not contain passwords. File security is accomplished via secure file transfer. No encryption is applied to files during storage on DIRBS.

Once received, each file is moved to the incoming folder in that operator's home directory on DIRBS for import and processing.

7 Data Validation by DIRBS

DIRBS regularly monitors for newly uploaded data from operators. When a new file is detected, it will be validated by DIRBS.

If validation fails, an alert will be generated and the DIRBS operational entity for that country will contact the operator to initiate re-upload of a valid data file.

NOTE: While schema validation can be checked by an operator prior to upload, such validation is only a subset of the validation performed by DIRBS, which may include comparison against historical metadata, identification of invalid rows, flagging of discrepancies, and application of import failure thresholds. Validation issues may be included in operator and/or audit reporting and may result in a validation failure alert, depending on the type and/or severity of such issues.