DIRBS Core Release 9.0.0

User Guide

80-GD079-2 Rev. E

30/09/2018Revision history

|  |  |  |
| --- | --- | --- |
| Revision | Date | Description |
| A | January 2018 | Initial release |
| B | April 2018 | Updates for DIRBS 7.0.0 release |
| C | May 2018 | Updates for DIRBS 8.0.0 release |
| D | July 2018 | Updates for DIRBS 8.0.1 release |
| E | September 2018 | Updates for DIRBS 9.0.0 release |

Contents

[1 Introduction 6](#__RefHeading___Toc16372_948293803)

[2 Configuring DIRBS Core 8](#__RefHeading___Toc16374_948293803)

[2.1 Copying the config file 8](#__RefHeading___Toc16376_948293803)

[2.2 Database details 8](#__RefHeading___Toc16378_948293803)

[2.3 Regional settings 9](#__RefHeading___Toc16380_948293803)

[2.3.1 Operator settings 9](#__RefHeading___Toc16382_948293803)

[2.4 Data purging 10](#__RefHeading___Toc16384_948293803)

[2.5 List generation settings 10](#__RefHeading___Toc16386_948293803)

[2.6 Report generation settings 11](#__RefHeading___Toc16388_948293803)

[2.7 Multiprocessing settings 11](#__RefHeading___Toc16390_948293803)

[2.8 Operator data import validation thresholds 12](#__RefHeading___Toc16392_948293803)

[2.9 Historic thresholds 13](#__RefHeading___Toc16394_948293803)

[2.10 Classification conditions 13](#__RefHeading___Toc16396_948293803)

[2.11 Logging settings 14](#__RefHeading___Toc16398_948293803)

[2.12 StatsD settings 15](#__RefHeading___Toc16400_948293803)

[2.13 Data catalog settings 16](#__RefHeading___Toc16402_948293803)

[3 Operating DIRBS Core 17](#__RefHeading___Toc16404_948293803)

[3.1 Data in the Docker instance 17](#__RefHeading___Toc16406_948293803)

[3.2 Managing database schema – dirbs-db 17](#__RefHeading___Toc16408_948293803)

[3.2.1 Creating DIRBS Core PostgreSQL roles – dirbs-db install\_roles 18](#__RefHeading___Toc16410_948293803)

[3.2.2 Installing schema into a clean database – dirbs-db install 18](#__RefHeading___Toc16412_948293803)

[3.2.3 Checking status of the database schema – dirbs-db check 19](#__RefHeading___Toc16414_948293803)

[3.2.4 Upgrading the database schema – dirbs-db upgrade 19](#__RefHeading___Toc16416_948293803)

[3.3 Importing data into DIRBS Core – dirbs import 19](#__RefHeading___Toc16418_948293803)

[3.3.1 GSMA TAC DB – dirbs-import gsma\_tac 21](#__RefHeading___Toc16420_948293803)

[3.3.2 operator data dumps – dirbs-import operator 22](#__RefHeading___Toc16422_948293803)

[3.3.3 golden list data – dirbs-import golden\_list 24](#__RefHeading___Toc16424_948293803)

[3.3.4 Local stolen list data – dirbs-import stolen\_list 25](#__RefHeading___Toc16426_948293803)

[3.3.5 pairing list data – dirbs-import pairing\_list 27](#__RefHeading___Toc16428_948293803)

[3.3.6 registration list data – dirbs-import registration\_list 28](#__RefHeading___Toc16430_948293803)

[3.4 Automating import of new files 29](#__RefHeading___Toc16432_948293803)

[3.4.1 Sample Makefiles and Jenkins 29](#__RefHeading___Toc16434_948293803)

[3.5 Classification of IMEIs – dirbs-classify 30](#__RefHeading___Toc16436_948293803)

[3.6 Generating lists – dirbs-listgen 33](#__RefHeading___Toc16438_948293803)

[3.7 Generating DIRBS reports – dirbs-report 35](#__RefHeading___Toc16440_948293803)

[3.7.1 dirbs-report directory structure 37](#__RefHeading___Toc16442_948293803)

[3.8 Accessing the API server 38](#__RefHeading___Toc16444_948293803)

[3.8.1 Data catalog API (Version 1.0) 39](#__RefHeading___Toc16446_948293803)

[3.8.2 Data catalog API (Version 2.0) 40](#__RefHeading___Toc16448_948293803)

[3.8.3 Job metadata API (Version 1.0) 40](#__RefHeading___Toc16450_948293803)

[3.8.4 Job Metadata API (Version 2.0) 41](#__RefHeading___Toc16452_948293803)

[3.8.5 TAC API (Version 1.0) 42](#__RefHeading___Toc16454_948293803)

[3.8.6 TAC API (Version 2.0) 42](#__RefHeading___Toc16456_948293803)

[3.8.7 IMEI API (Version 1.0) 42](#__RefHeading___Toc16458_948293803)

[3.8.8 IMEI API (Version 2.0) 43](#__RefHeading___Toc16460_948293803)

[3.8.9 MSISDN API (Version 1.0) 43](#__RefHeading___Toc16462_948293803)

[3.8.10 MSISDN API (Version 2.0) 43](#__RefHeading___Toc16464_948293803)

[3.8.11 Version API (Version 1.0) 44](#__RefHeading___Toc16466_948293803)

[3.8.12 Version API (Version 2.0) 44](#__RefHeading___Toc16468_948293803)

[3.9 Pruning old data 44](#__RefHeading___Toc16470_948293803)

[4 Understanding DIRBS Reports 46](#__RefHeading___Toc16472_948293803)

[4.1 Standard reports 46](#__RefHeading___Toc16474_948293803)

[4.1.1 Country report 47](#__RefHeading___Toc16476_948293803)

[4.1.2 Operator reports 66](#__RefHeading___Toc16478_948293803)

[4.2 Condition IMEI overlaps reports 67](#__RefHeading___Toc16480_948293803)

[4.3 GSMA not found reports 67](#__RefHeading___Toc16482_948293803)

[4.4 Stolen violations reports 67](#__RefHeading___Toc16484_948293803)

[4.5 Top duplicates reports 68](#__RefHeading___Toc16486_948293803)

[5 Understanding DIRBS Lists 69](#__RefHeading___Toc16488_948293803)

[5.1 Blacklist 69](#__RefHeading___Toc16490_948293803)

[5.1.1 Full blacklist 69](#__RefHeading___Toc16492_948293803)

[5.1.2 Delta blacklist 70](#__RefHeading___Toc16494_948293803)

[5.2 Notifications lists 71](#__RefHeading___Toc16496_948293803)

[5.2.1 Full notification list 71](#__RefHeading___Toc16498_948293803)

[5.2.2 Delta notification lists 72](#__RefHeading___Toc16500_948293803)

[5.3 Exceptions lists 73](#__RefHeading___Toc16502_948293803)

[5.3.1 Full exceptions list 73](#__RefHeading___Toc16504_948293803)

[5.3.2 Delta exceptions list 73](#__RefHeading___Toc16506_948293803)

[6 Frequently Asked Questions 75](#__RefHeading___Toc16508_948293803)

[6.1 How does duplicate averaging work? 75](#__RefHeading___Toc16510_948293803)

[6.2 Reported error during dirbs-classify or dirbs-import 75](#__RefHeading___Toc16512_948293803)

[6.3 Reported error during dirbs-import 76](#__RefHeading___Toc16514_948293803)

[1.1 Understanding gsma\_not\_found Reporting Body Index delay configuration 76](#__RefHeading___Toc16516_948293803)

[1.2 Duplicate and conflicting rows in non-operator imports 77](#__RefHeading___Toc16518_948293803)

[1.2.1 Key and metadata columns 77](#__RefHeading___Toc16520_948293803)

[6.3.2 Problems 78](#__RefHeading___Toc16522_948293803)

[1.2.2 Options for resolving a conflicting row problem 80](#__RefHeading___Toc16524_948293803)

[6.4 DIRBS Amnesty feature 80](#__RefHeading___Toc16526_948293803)

[1.2.3 Enabling and configuring amnesty in .dirbs.yml 81](#__RefHeading___Toc16528_948293803)

[6.4.1 Eligibility and notifications 82](#__RefHeading___Toc16530_948293803)

[6.4.2 Stolen, paired, and golden IMEI interaction 82](#__RefHeading___Toc16532_948293803)

Figures

[Figure 1-1 DIRBS Core 7](#_Toc514405158)

[Figure 1-2 DIRBS input/output 8](#_Toc514405159)

[Figure 3-1 Importing data into DIRBS 21](#_Toc514405160)

[Figure 4-1 Country report main page – HTML 47](#_Toc514405161)

[Figure 4-2 Identifier counts 48](#_Toc514405162)

[Figure 4-3 Identifier trends 48](#_Toc514405163)

[Figure 4-4 Compliance breakdown 49](#_Toc514405164)

[Figure 4-5 IMEI compliance trends 49](#_Toc514405165)

[Figure 4-6 Conditions breakdown 51](#_Toc514405166)

[Figure 4-7 Condition combinations 52](#_Toc514405167)

[Figure 4-8 Blacklist and blacklist violations 52](#_Toc514405168)

[Figure 4-9 Top models: counts 53](#_Toc514405169)

[Figure 4-10 Top models: gross adds 54](#_Toc514405170)

[Figure 7-1 Duplicate averaging 78](#_Toc514405171)

[Figure 6-2 RBI delay 80](#_Toc514405172)

Tables

[Table 2-1 Configuring database settings 9](#_Toc514405173)

[Table 2-2 Configuring regional settings 10](#_Toc514405174)

[Table 2-3 Operator settings 10](#_Toc514405175)

[Table 2-4 Data retention settings 11](#_Toc514405176)

[Table 2-5 List generation settings 11](#_Toc514405177)

[Table 2-6 Report generation settings 12](#_Toc514405178)

[Table 2-7 Multiprocessing settings 12](#_Toc514405179)

[Table 2-8 Import validation thresholds 13](#_Toc514405180)

[Table 2-9 Historic threshold 14](#_Toc514405181)

[Table 2-10 Conditions settings 14](#_Toc514405182)

[Table 2-11 Logging settings 15](#_Toc514405183)

[Table 2-12 StatsD Settings 16](#_Toc514405184)

[Table 2-13 Data catalog settings 17](#_Toc514405185)

[Table 3-1 gsma\_tac fields and format 22](#_Toc514405186)

[Table 3-2 operator data fields and format 24](#_Toc514405187)

[Table 3-3 Validation checks 24](#_Toc514405188)

[Table 3-4 golden list MD5 pre-hashed fields and format 26](#_Toc514405189)

[Table 3-5 golden list fields and format 26](#_Toc514405190)

[Table 3-6 stolen list fields and format 28](#_Toc514405191)

[Table 3-7 pairing list fields and format 29](#_Toc514405192)

[Table 3-8 registration list fields and format 30](#_Toc514405193)

[Table 3-9 Implemented dimensions and parameters 32](#_Toc514405194)

[Table 3-10 DIRBS Core lists 36](#_Toc514405195)

[Table 3-11 Report types 37](#_Toc514405196)

[Table 3-12 Data catalog API 42](#_Toc514405197)

[Table 3-13 Job metadata API 43](#_Toc514405198)

[Table 3-14 Prune commands 44](#_Toc514405199)

[Table 4-1 Blacklist information 55](#_Toc514405200)

[Table 4-2 Compliance breakdown 57](#_Toc514405201)

[Table 4-3 Condition combinations 58](#_Toc514405202)

[Table 4-4 Conditions breakdown 59](#_Toc514405203)

[Table 4-5 Historic IMEI, IMSI, MSISDN and triplet counts 63](#_Toc514405204)

[Table 4-6 IMEI/IMSI and IMSI/IMEI overloading 64](#_Toc514405205)

[Table 4-7 Daily counts for IMEIs, IMSIs and MSISDNs 64](#_Toc514405206)

[Table 4-8 Top models 66](#_Toc514405207)

[Table 4-9 Monthly counts 67](#_Toc514405208)

[Table 5-1 Blacklist event types 73](#_Toc514405209)

[Table 5-2 Notification list event types 75](#_Toc514405210)

[Table 5-3 Exceptions list change types 77](#_Toc514405211)

[Table 6-1 Key and metadata columns 80](#_Toc514405212)

[Table 6-2 Stolen list 81](#_Toc514405213)

[Table 6-3 Stolen list after normalization 81](#_Toc514405214)

[Table 6-4 Conflicting rows 82](#_Toc514405215)

[Table 6-5 Future DIRBS Core registration list 82](#_Toc514405216)

[Table 6-6 Normalized delta file 83](#_Toc514405217)

[Table B-1 YML sample configuration 96](#_Toc514405218)

# Introduction

The Device Identification, Registration & Blocking System (DIRBS) is a country-wide system deployed in cooperation between the country regulator, operators in that country, and a technology partner that supports deployment. The system checks, identifies, and discourages non-compliant devices by verifying the installed base of devices currently active in a market and continuing to monitor as new devices are activated.

DIRBS can verify that:

* Devices have properly allocated identifiers and type approval
* Devices are not duplicated or stolen
* Device importation takes place through legal channels

DIRBS consists of the DIRBS Core and a set of DIRBS Interface Systems that interface with DIRBS Core (see Figure 1 -1 and ).

1. DIRBS Interface Systems may be developed by third-party technology partner(s).



Figure 1-1 DIRBS Core



Figure 1-2 DIRBS input/output

# Configuring DIRBS Core

A sample config file, config.yml, was distributed with this release. When DIRBS Core scripts run, they look for a file in ~/.dirbs.yml. If that file is not found, look in the system location /opt/dirbs/etc/config.yml.

For the complete sample .yml file, see Appendix .

## Copying the config file

**To run a configuration using Docker,** edit ./etc/config.yml, and rebuild the Docker images.

If you are not using Docker:

1. Copy dirbs config to the appropriate location with:

cp /opt/dirbs/etc/config.yml ~/.dirbs.yml

## Database details

Table 2 -1 lists settings of the postgresql section in the sample config. These values are overridden by the environment variable if configured.

Table 2-1 Configuring database settings

|  |  |  |
| --- | --- | --- |
| Config file setting name | Environment variable name | Function |
| database | DIRBS\_DB\_DATABASE | Database name (an empty database on the first run) |
| host | DIRBS\_DB\_HOST | Host that the PostgreSQL server runs on |
| port | DIRBS\_DB\_PORT | PostgreSQL port if not running on standard port of 5432 |
| user | DIRBS\_DB\_USER | Username (user should have all privileges specified on the database) |
| password | DIRBS\_DB\_PASSWORD | Password used to connect to the database |

1. The database password can be configured in several ways and are dependent on the level of security required in your deployed system. The password may be configured in the following ways:

* In clear text in the .dirbs.yml file by configuring the password setting.
* The user’s .pgpass file located in their home directory.
* The DIRBS\_DB\_PASSWORD environment variable is set. This overwrites values configured in the .dirbs.yml and the .pgpass file
* The password can be provided in the dirbs command-line option as –db-password-prompt.

## Regional settings

DIRBS Core uses regional settings for reporting and input validation. Table 2 -2 lists settings of the region section in the sample config.

Table 2-2 Configuring regional settings

|  |  |  |
| --- | --- | --- |
| Config file setting name | Value, range | Function |
| name | String | Name used for the country level report |
| import\_msisdn\_data | True/False | Whether or not MSISDN data is present and should be imported for this region |
| import\_rat\_data | True/False | Whether or not RAT data is present and should be imported for this region |
| country\_codes | Digits | * List of country codes for the region. * Used to validate MSISDNs in operator data dumps |
| exempted\_device\_types | String | * Exempted\_device\_types contains a list of GSMA device types that do not require registration in this country. * Specifiying a list of device types here will mean that the not\_in\_registration\_list classification dimension will ignore IMEIs whose TACs correspond to the listed device types. They will also be ignored in the IMEI API's realtime registration check. |
| operators | See Section 2.3.1 | See Section 2.3.1 |

### Operator settings

The region config has an operators section. Table 2 -3 lists the settings for each operator section in the sample config.

Table 2-3 Operator settings

|  |  |  |
| --- | --- | --- |
| Setting | Value, range | Function |
| id | String | Operator ID (short string used to uniquely id operator) |
| name | String | Longer name for the operator, used in reports |
| mcc\_mnc\_pairs | Digits | MCC/MNC pairs for the operator   * Validate IMSIs during operator data import * Determine which operator notifications about an offending subscriber should be sent to * Determine which operators excepted IMEI-IMSI pairings should be sent to |

## Data purging

The sample config has a data\_retention section that defines the number of months operator data can remain before it is deleted from the system (see Error: Reference source not found).

1. The DIRBS Core system does not delete CSV data off disk. Purging only applies to data in the database. It is recommended that an operator of the DIRBS Core system write a small job to remove old CSV data if this is also required by regulation.

Table 2-4 Data retention settings

|  |  |  |
| --- | --- | --- |
| Setting | Value, range | Function |
| months\_retention | Integer | * The number of months from the start of the current month that DIRBS Core retains data about a triplet seen in its DB. * After this time, the triplet will be erased from the seen\_triplet table. The IMEI continues to be stored after this date as it is needed for continued list generation, etc. All references to IMSI and MSISDN will be pruned after this date. * If this value is set to two months and the current date is March 29, only the data in December or earlier will be purged from the database when the dirbs-prune command is run. |

## List generation settings

The sample config has a list\_generation section containing a list of configurable settings related to list generation (see Table 2 -5).

Table 2-5 List generation settings

|  |  |  |
| --- | --- | --- |
| Setting | Value, range | Function |
| lookback\_days | Integer | The number of days that DIRBS Core looks back through data from the current date to determine IMSIs/MSISDNs which were associated with the notifiable IMEIs. |
| restrict\_exceptions  \_list\_to\_blacklisted\_imeis | True/False | * If true, the exception list contains only those IMEI-IMSI pairs where the IMEI is on the blacklist. * By default, all IMEI-IMSI pairs of the pairing list are output to the exception list. |
| generate\_check\_digit | True/False | * If true, generates a check digit for IMEIs during list generation. * Check digit will only be added to "valid IMEIs". |
| output\_invalid\_imeis | True/False | * If true, outputs only "valid" IMEIs. * Valid IMEIs start with 14 digits (they will have 15 digits if the check digit append has been enabled). |

## Report generation settings

The sample config has a report\_generation section containing a list of configurable settings related to report generation process (see Table 2 -6).

Table 2-6 Report generation settings

|  |  |  |
| --- | --- | --- |
| Setting | Value, range | Function |
| blacklist\_violations\_grace\_period\_days | Integer | Used by blacklist violations and stolen list violations reports to give the MNO some processing time (in days) before an IMEI appearing on the network is considered a violation. |

## Multiprocessing settings

The sample config has a multiprocessing section that outlines how DIRBS Core uses multiple cores and database connections to speed up processing (see Table 2 -7).

Table 2-7 Multiprocessing settings

|  |  |  |
| --- | --- | --- |
| Setting | Value, range | Function |
| max\_local\_cpus | Integer | * Maximum number of local processing blade workers to achieve DIRBS Core tasks * Useful for pre-validation of large operator import jobs where we can run multiple instances of the pre-validator in parallel on different parts of the file * Default is to use half of the available CPUs in the system |
| max\_db\_connections | Integer | * Maximum number of database connections to parallelize DIRBS Core tasks   + PostgresSQL 9.6 has support for parallelizing tasks internally and this setting does not affect parallelization for a single connection * When PostgresSQL cannot parallelize a single query by itself, we use this number of workers to issue multiple queries at once on different connections   + Generally scales very well, safe to set this high   + Should probably be set to roughly the number of disks in your RAID array |
| max\_db\_writers | Integer | * Number of I/O write-intensive DB connections to use at once   + Should be lower than max\_db\_connections   + For RAID-10, this should be set to about half of max\_db\_connections to indicate the paired nature of that type of RAID array |

## Operator data import validation thresholds

The sample config has an operator\_threshold section containing a list of validation thresholds for operator data. These thresholds are used when importing operator data using the dirbs-import command. Table 2 -8 lists the settings for these thresholds and their functions.

Table 2-8 Import validation thresholds

|  |  |  |
| --- | --- | --- |
| Setting | Value, range | Function |
| null\_imei\_threshold  null\_imsi\_threshold  null\_msisdn\_threshold  null\_rat\_threshold | Decimal, 0.0-1.0 | * Proportion of entries in the data allowed to have one or more of IMSI, MSISDN, IMEI, or RAT as null * null\_msisdn\_threshould is ignored if import\_msisdn\_data is set to False * null\_rat\_threshold is ignored if import\_rat\_data is set to False |
| null\_threshold | Decimal, 0.0-1.0 | * Proportion of the entries in the data allowed to have any column equal to NULL. * Only includes columns enabled in the import (MSISDN and RAT may be excluded) |
| unclean\_imei\_threshold | Decimal, 0.0-1.0 | Proportion of non-NULL IMEIs in the data allowed to not start with 14 digits |
| unclean\_imsi\_threshold | Decimal, 0.0-1.0 | Proportion of non-NULL IMSIs in the data allowed to not be 14-15 digits |
| unclean\_threshold | Decimal, 0.0-1.0 | Proportion of entries in the data allowed to have either an unclean IMEI or an unclean IMSI |
| out\_of\_region\_imsi\_threshold | Decimal, 0.0-1.0 | Proportion of non-NULL IMSIs in the data allowed to have an MCC that does not match the configured region |
| out\_of\_region\_msisdn\_threshold | Decimal, 0.0-1.0 | * Proportion of non-NULL MSISDNs in the data allowed to have a CC that does not match the configured region * Ignored if MSISDN disabled |
| out\_of\_region\_threshold | Decimal, 0.0-1.0 | * Combined proportion of entries in the data allowed to have either a CC (IMSI) or MCC (MSISDN) that does not match the configured region * Ignored if MSISDN disabled (same as the out-of-region IMSI check) |
| non\_home\_network\_threshold | Decimal, 0.0-1.0 | Proportion of entries in the data allowed to have an IMSI not starting with one of the MCC-MNC prefixes associated with the operator the data is being imported for |
| historic\_imei\_threshold | Decimal, 0.0-1.0 | Minimum valid ratio of average daily IMEI count against historical daily IMEI count for a data dump to be considered valid |
| historic\_imsi\_threshold: | Decimal, 0.0-1.0 | Minimum valid ratio of average daily IMSI count against historical daily IMSI count for a data dump to be considered valid |
| historic\_msisdn\_threshold | Decimal, 0.0-1.0 | * Minimum valid ratio of average daily MSISDN count against historical daily MSISDN count for a data dump to be considered valid * Ignored if MSISDN disabled |

## Historic thresholds

The sample config has a historic\_thresholds section that can validate new import row count against previously imported data for the same importer. Table 2 -9 lists the settings for these thresholds and their functions.

Table 2-9 Historic threshold

|  |  |  |
| --- | --- | --- |
| Setting | Value, range | Function |
| gsma\_threshold:  import\_size\_variation\_absolute:  import\_size\_variation\_percent:  pairing\_list\_threshold:  import\_size\_variation\_absolute:  import\_size\_variation\_percent:  stolen\_list\_threshold:  import\_size\_variation\_absolute:  import\_size\_variation\_percent:  registration\_list\_threshold:  import\_size\_variation\_absolute:  import\_size\_variation\_percent:  golden\_list\_threshold:  import\_size\_variation\_absolute:  import\_size\_variation\_percent: | Integer  Decimal, 0.0-1.0  Integer  Decimal, 0.0-1.0  Integer  Decimal, 0.0-1.0  Integer  Decimal, 0.0-1.0  Integer  Decimal, 0.0-1.0 | * import\_size\_variation\_absolute: The most an import can decrease in absolute row count before it is rejected as invalid. By setting this variable to -1, this check will be disabled. * import\_size\_variation\_percent: The most an import can decrease in percentage row count before it is rejected as invalid.   + 0.75 indicates a new import must be at least 75% of the previous import's row count or it will be rejected. Therefore, setting this variable to 0 will disable this check. |

## Classification conditions

The sample config has a conditions section containing a list of conditions that classify IMEIs (see Table 2 -10).

Table 2-10 Conditions settings

|  |  |
| --- | --- |
| Setting | Function |
| label | * A name/ID/key for the condition * If label is changed, all previous classifications will be reset   + Likewise, if you change the dimensions but keep the condition label the same, existing classifications for that condition will be retained. |
| dimensions | * For dimension configuration details, see Section 3.5. * A list of dimensions whose intersection forms the IMEI set result for the condition. Each dimension can take parameters specific to the dimension being used. Additionally, they all accept an invert property, which NOTs the result of the dimension by taking the all-time observed IMEIs list and subtracting the set of IMEIs returned by this dimension. * Each dimension has:   + module: Python code module that implements condition.   + parameters: List of parameters supplied to dimension (available parameters depend on dimension).   + invert: Boolean stating whether results from this dimension should be inverted to produce a NOT result, i.e., inverting the GSMA Not Found dimension would return a list of all seen IMEIs that were found in the GSMA TAC DB. |
| grace\_period\_days | Number of days after DIRBS Core detects condition has been met by IMEI before it moves from the notification list to the blacklist. |
| blocking | * Boolean stating whether this condition contributes to list generation or is simply informational. * Information conditions can be used to try out new modules or to tweak parameters. |
| reason | Human-readable reason string summarizing this condition (to be used in notification lists and blacklist). |
| max\_allowed\_matching\_ratio | * The maximum ratio of IMEIs that can be matched before the condition fails because something is wrong.   + Example: Under normal circumstances, we might not expect more than 20% of all IMEIs to not be found in GSMA, so we can set this to 0.2. If more than this number of IMEIs match, we probably forgot to import the GSMA TAC DB, so we refuse to add the results of this condition to the classification\_state table. Other conditions will still be added to the DB but the overall status of the dirbs-classify job will be non-zero (failure). |

## Logging settings

The sample config has a logging section containing a list of configurable settings related to logging output (see Table 2 -11).

Table 2-11 Logging settings

|  |  |  |
| --- | --- | --- |
| Setting | Value, range | Function |
| level | String | * Minimum logging level to output log messages before * Valid values are:   + debug   + info   + warn   + error |
| format | [LogRecord Objects](https://docs.python.org/3/library/logging.html" \l "logging.LogRecord) | Format string for log messages output by the application  Example Format:  format: '%(asctime)s - %(name)s - %(levelname)s - %(message)s' |
| show\_statsd\_messages | True/False | * Determines whether or not log messages related to StatsD should be output * If enabled, all calls to StatsD will be logged for debugging |
| show\_sql\_messages | True/False | * Determines whether or not log queries made it to the SQL database * If enabled (True), all queries will be logged for debugging (can result in very large output) |
| show\_werkzeug\_messages | True/False | * Determines whether or not log internal Werkzeug messages from the TAC/IMEI APIs * Should almost always be set to False * If enabled (True), all TAC/IMEI API queries will be logged (can resilt in large output) |
| log\_directory | String | * Sets this directory to store the log files * Directory must exist and be writable |
| file\_prefix | String | Uncomment and set this value to prefix all log files created on this host with a prefix to distinguish them from other hosts |
| file\_rotation\_max\_bytes | Integer | * Sets the number of bytes before a logfile is rotated * If this or file\_rotation\_backup\_count is 0, rotation is disabled |
| file\_rotation\_backup\_count | Integer | Sets the number of old logs to keep |

## StatsD settings

The sample config has a statsd section containing a list of configurable settings related to forwarding application-defined metrics to a StatsD server for aggregation (see Table 2 -12).

Table 2-12 StatsD Settings

|  |  |  |
| --- | --- | --- |
| Setting | Value, range | Function |
| hostname | String | * StatsD server hostname * Overridden by the environment variable DIRBS\_STATSD\_HOST |
| port | Integer | * UDP port that StatsD server listens on for metrics * Overridden by environment variable DIRBS\_STATSD\_PORT |
| prefix | True/False | * Prefix for all metrics collected from this instance   + Useful if there are multiple hosts or environments sending data to the same StatsD server and you want to differentiate them * Overridden by the environment variable DIRBS\_ENV |

## Data catalog settings

The sample config has a catalog section containing a list of configurable settings related to the data cataloging process (see Table 2 -13).

Table 2-13 Data catalog settings

|  |  |  |
| --- | --- | --- |
| Setting | Value, range | Function |
| file\_type | String | * Type of files contained within the specified paths * Should match the keyword specified during dirbs-import, e.g. operator, gsma\_tac, etc. |
| paths | Integer | * Directories and/or files to be harvested   + Sub-directories within the listed path are not traversed automatically   + Should be listed separately if files within them must be cataloged. * Multiple paths can be defined for each file type and the path used should be absolute and globally unique |
| schema\_filename: | String | * Schema file for data pre-validation (if enabled) * Multiple prospectors can be defined for the same file\_type if files exist across multiple schema versions |
| perform\_prevalidation | String, True/False | * Set to true if pre-validation should be performed on the data files * Enabling this can slow down the process if there are a lot of uncataloged files |

# Operating DIRBS Core

1. **If you are not using Docker**, the following commands only work when the DIRBS Core virtual environment is activated. This must be done each time you log into the machine or start a new shell.

To activate the virtual environment, run:

source <install\_path>/bin/activate

## Data in the Docker instance

If you run the Data Processing blade image, the /data directory is marked as a persistent data volume in Docker. Data in this directory will persist after a container has been destroyed.

For more information on data volumes, see <https://docs.docker.com/engine/tutorials/dockervolumes/>.

When the Data Processing blade image is created, the entry point script (entrypoint.sh) populates the /data directory with the required folder structure, along with the correct permissions. This script reads a list of operators from the DIRBS\_OPERATOR environment variable and creates folders for each if they are missing.

There are a couple options to get data into the container:

* Bind-mount /data to a directory on the host machine using the –v option to docker run. If using this approach, ensure that the directory has the correct permissions and can be written to by the docker user. The entry point creates the requisite folders in this directory so it can be empty to begin with.
* Use commands scp or ssh to copy data into the container. Once it is in there, it will persist across container builds due to the persistent nature of the data volume.

## Managing database schema – dirbs-db

dirbs-db manages the database schema version deployed to the PostgreSQL server.

For information on the dirbs-db command and its available subcommands, run:

dirbs-db –-help

|  |
| --- |
| Usage: dirbs-db [OPTIONS] COMMAND [ARGS]...  DIRBS script to intiliaze, configure and upgrade the PostgreSQL schema.  Options:  --version Show the version and exit.  -v, --verbose Print debug console output - file output is  unaffected.  --db-password-prompt If set, will prompt the user for a PostgreSQL  password rather than reading from config.  --db-user TEXT The PostgreSQL DB database user to connect as.  --db-name TEXT The PostgreSQL DB database name to connect to.  --db-port INTEGER The PostgreSQL DB port to connect to.  --db-host TEXT The PostgreSQL DB host to connect to.  --statsd-prefix TEXT The environment prefix to prepend to all StatsD  metrics.  --statsd-port INTEGER The StatsD port to connect to on the configured host.  --statsd-host TEXT The StatsD host to send metrics to.  --help Show this message and exit.  Commands:  check Checks whether DB schema matches software DB...  fix\_whitespace\_triplets Fixes whitespace issues in seen\_triplets...  install Installs latest schema on clean DB instance.  install\_roles Creates DIRBS Core PostgreSQL base roles if...  upgrade Upgrades the current DB schema to the version... |

### Creating DIRBS Core PostgreSQL roles – dirbs-db install\_roles

1. **Roles must be created and installed prior to running the** dirbs-db install **and** dirbs-db upgrade **commands. For detailed instructions on installing and configuring a new database, see *DIRBS Core Release 5.2.0 Installation Guide* (80-GD079-1).**

dirbs-db --db-user <username> --db-password-prompt install\_roles

where

* <username> is the name of the user with the CREATEROLE privilege

### Installing schema into a clean database – dirbs-db install

To install the DIRBS Core schema into a clean database, run:

dirbs-db --db-user <username\_of\_power\_user> --db-password-prompt install

This command only works on a clean database. You can force-install the schema into a non-clean database using –force flag, but this is dangerous as it may leave the database in an inconsistent state where future migration scripts fail.

For help on the options available to dirbs-db install, run:

dirbs-db install –-help

|  |
| --- |
| Usage: dirbs-db install [OPTIONS]  Installs latest schema on clean DB instance.  Options:  --help Show this message and exit. |

### Checking status of the database schema – dirbs-db check

dirbs-db check displays which schema version is currently deployed to the PostgreSQL database, and which version is required by the installed code.

To run this check, run:

dirbs-db check

### Upgrading the database schema – dirbs-db upgrade

It is recommended to take a database backup before attempting these steps in case something goes wrong. Downgrades are not possible.

dirbs-db upgrade upgrades the currently installed database schema in PostgreSQL to the version required by the installed software.

To run the upgrade, run:

dirbs-db --db-user <power\_user> --db-password-prompt upgrade

where

* <power\_user> is a user that has been GRANT’ed the dirbs\_core\_poweruser role

The upgrade script determines which version of the schema is required and automatically runs SQL migration scripts to upgrade the schema.

## Importing data into DIRBS Core – dirbs import

This section describes how to import data into the DIRBS Core using dirbs-import functionality.

While the DIRBS System does not enforce a specific data import order, Figure 3 -3 describes the recommended steps involved in importing data.



Figure 3-3 Importing data into DIRBS

During normal system operation, it is expected that jobs are scheduled and configured to periodically import these data files.

Depending on the analysis that you are performing you may or may not require the import of the files listed in .

For information on the dirbs-import and its available subcommands, run:

dirbs-import –-help

|  |
| --- |
| Usage: dirbs-import [OPTIONS] COMMAND [ARGS]...  DIRBS script to import data into DIRBS Core PostgreSQL database.  Options:  --version Show the version and exit.  -v, --verbose Print debug console output - file output is  unaffected.  --db-password-prompt If set, will prompt the user for a  PostgreSQL password rather than reading from  config.  --db-user TEXT The PostgreSQL DB database user to connect  as.  --db-name TEXT The PostgreSQL DB database name to connect  to.  --db-port INTEGER The PostgreSQL DB port to connect to.  --db-host TEXT The PostgreSQL DB host to connect to.  --statsd-prefix TEXT The environment prefix to prepend to all  StatsD metrics.  --statsd-port INTEGER The StatsD port to connect to on the  configured host.  --statsd-host TEXT The StatsD host to send metrics to.  --max-db-writers INTEGER The maximum write-intensive DB connections  to use concurrently during this job.  --max-db-connections INTEGER The maximum DB connections to use  concurrently during this job.  --max-local-cpus INTEGER The maximum number of local CPUs to use  concurrently during this job.  --batch-size INTEGER Size of batches to import into DB, in lines.  --no-cleanup If set, intermediate split data files and  the staging table will not be deleted so  that they can inspected.  --extract-dir DIRECTORY Directory to extract contents of .zip file  into (same directory as input file by  default).  --prevalidator-path PATH The path to the CSV pre-validator  executable.  --prevalidator-schema-path DIRECTORY  The path to the directory where the CSV pre-  validator schema are stored.  --help Show this message and exit.  Commands:  golden\_list Import the Golden list data found in INPUT...  gsma\_tac Import the GSMA TAC DB data found in INPUT...  operator Import the CSV operator data found in INPUT...  pairing\_list Import the Pairing List data found in INPUT...  registration\_list Import the Registration list data found in...  stolen\_list Import the Stolen List data found in INPUT |

### GSMA TAC DB – dirbs-import gsma\_tac

To import a .zip version of the GSMA TAC database, run:

dirbs-import gsma\_tac <gsma\_zip\_file>

The .zip file is expected to contain a .txt file where the columns are pipe-separated. It is run through the CSV pre-validator GoldenListSchemaData.csvs located at /opt/dirbs/etc/schema to ensure it conforms to the expected format by DIRBS Core.

Table 3 -14 lists the expected header columns and format. These fields can be in any order and case.

Table 3-14 gsma\_tac fields and format

|  |  |  |
| --- | --- | --- |
| Field | Mandatory (M)  Optional (O) | Expected format |
| tac | M1 | Integer, Length 8 |
| manufacturer | O | String, Length (1-128) 2 |
| model name | O | String, Length (1-1024) 2 |
| bands | O | String, Length (1-4096) 2 |
| allocation date | O | * Day-Month-Year format2, e.g., 26-Apr-2016   + Day: 0 – 31 (must correspond to days in month/year)   + Month: Jan, Feb, Mar, Apr, May, Jun, Jul, Aug, Sept, Oct, Nov, Dec   + Year: 19XX, 20XX |
| optional fields | O | String, Length (1-4096) 2 |
| 1 The tac field must not be empty.  2 Empty string and whitespace allowed. | | |

The GSMA TAC DB importer can accept additional columns, which will be imported into an optional\_fields column in the DB and returned by the TAC API without changes. No validation or processing takes place on these optional fields.

The GSMA TAC DB also performs a historic validation check as specified in section 2.9. If this check fails, the import is rejected. This check can be disabled with --disable-historic-check.

For help on all the options available to dirbs-import gsma\_tac, run:

dirbs-import gsma\_tac –-help

|  |
| --- |
| Usage: dirbs-import gsma\_tac [OPTIONS] INPUT\_FILE  Import the GSMA TAC DB data found in INPUT into the PostgreSQL database.  Options:  --disable-historic-check Skip checking the size of this import against  the currently stored data.  --disable-duplicates-check Skip checking for duplicate rows in this file  and failing if there any.  --help Show this message and exit. |

### operator data dumps – dirbs-import operator

To import a .zip version of an operator data dump, run:

dirbs-import operator <operator\_id> <operator\_zip\_file>

The .zip file is expected to contain a ‘,’ comma separated .csv file containing the operator data. It is run through the CSV pre-validator OperatorImportSchema\_v2.csvs located at /opt/dirbs/etc/schema to ensure it conforms to the expected format by DIRBS Core.

Table 3 -15 lists the expected header columns and format. These fields can be in any order and case.

Table 3-15 operator data fields and format

|  |  |  |
| --- | --- | --- |
| Field | Mandatory (M)  Optional (O) | Expected format |
| date | M | * yyyymmdd format starting with year 2000, e.g., 20160703   + yy: 20XX, where X=0-9   + mm: 00-12   + dd: 00 – 31 (must correspond to days in month/year) |
| imei | O | * Digits, Length (1 – 16) 1   + Valid digits: 0-9,A-F,a-f,\*,#   + 14 leading digits for good IMEI Records(check digit/software version is stripped on import) |
| imsi | O | * Digits, Length (1-15) 1   + Valid digits: 0-9 |
| msisdn | O | * Digits, Length (1-15) 1   + E.164 format |
| rat | O | * Digits, Length (3) 1   + Within range of 001-007 or 101-105 |
| 1 Empty string and whitespace allowed but will be stored as NULL. | | |

operator\_id is a string or number that uniquely identifies the operator which must match one of the operator IDs in the config .yml file. Data import will fail and will not be imported if the Operator ID does not match.

1. **If operatorIDs are modified/replaced after successfully importing operator data, that data will still be included in the country-level reporting and in blacklist generation. However, notification lists will not be generated for the previously replaced operatorID.**

The default behavior of the operator data importer expects data to contain RAT information. If a data dump does not contain RAT information, it can be imported using the --disable-rat-import.

Validation checks

Operator data performs multiple validation checks during import. Table 3 -16 lists the validation checks and their functions.

Table 3-16 Validation checks

|  |  |
| --- | --- |
| Check | Function |
| CSV pre-validation | Ensures input CSV conforms to the operator data schema |
| Filename checks | Ensures .zip file conforms to the required filename format (<operator\_id>\_<startdate>\_<enddate>.zip) and that the .csv filename within the .zip also conforms to that filename format |
| Date checks | Ensures connection\_date field in the CSV data falls within the date range specified by the filename |
| Null checks | Ensures CSV data does not contain too many rows with blanks for IMEIs, IMSIs, MSISDNs, and/or RAT values |
| Clean checks | Ensures CSV data does not contain too many rows with invalid characters in the IMEI, IMSI or MSISDN |
| Leading zero checks | Ensures leading zeroes have not been stripped from the IMEIs in the CSV data |
| Region checks | Ensures that not too many rows contain IMSIs or MSISDNs with out-of-region CC and MCC values |
| Home network checks | Ensures that not too many rows contain IMSIs with an MCC-MNC prefix not associated with the operator |
| Historic checks | * Ensures data is consistent with previous imports from the same operator * Performed based on previously generated reports   + If reporting was never performed, historic checks will not be performed   + These checks compare the average daily counts for IMEI, IMSIs, and MSISDNs against historical counts |

Most of these checks can be disabled with command-line options. For help on all the options available to dirbs-import operator, run:

dirbs-import operator –-help

|  |
| --- |
| Usage: dirbs-import operator [OPTIONS] OPERATOR\_ID INPUT\_FILE  Import the CSV operator data found in INPUT into the PostgreSQL database.  OPERATOR\_ID is an ID up to 16 characters to unique identify the operator.  Options:  --disable-leading-zero-check Skip checking if the import data appears to  have lost leading zeros.  --disable-null-check Skip checking the ratio of IMSIs, MSISDNs,  IMEIs and RATs that are NULL.  --disable-clean-check Skip checking the ratio of IMEIs and IMSIs  that are the wrong length or contain invalid  characters.  --disable-region-check Skip checking the ratio of MSISDNs and IMSIs  that have out of region cc and mcc values.  --disable-home-check Skip checking the ratio of and IMSIs that have  out of region mcc and mnc pair values.  --disable-msisdn-import Skip importing MSISDN field even if it does  exist in input data.  --disable-rat-import Skip importing RAT field if it does not exist  in input data.  --disable-historic-check Skip checking the size of this import against  the currently stored data.  --help Show this message and exit. |

### golden list data – dirbs-import golden\_list

To import a .zip version of the complete golden list, run:

dirbs-import golden\_list <golden\_list\_zip\_file>

The golden list identifies IMEIs of high-ranking officials to be excluded from being blocked.

* 1. Any IMEIs added to the golden list will never be blocked.

The .zip file is expected to contain a .csv file containing the list of golden list data. It is run through the CSV pre-validator GoldenListSchemaPreHashedData.csvs located at /opt/dirbs/etc/schema to ensure it conforms to the expected format by DIRBS Core.

The golden list can be imported with MD5 pre-hashed IMEIs or non-hashed IMEI. When hashing the IMEIs, it is expected to hash a 14-digit IMEI (see Section 1.2)

Table 3 -17 and Table 3 -18 list the expected header columns and format of a golden list.

Table 3-17 golden list MD5 pre-hashed fields and format

|  |  |  |
| --- | --- | --- |
| Field | Mandatory (M)  Optional (O) | Expected format |
| golden\_imei | M | Hex String, Length (32), MD5 Encrypted  Valid characters: 0-9,A-F,a-f,\*,# |

Table 3-18 golden list fields and format

|  |  |  |
| --- | --- | --- |
| Field | Mandatory(M)  Optional(O) | Expected format |
| golden\_imei | M | Digits, Length(1 – 16)  Valid digits: 0-9,A-F,a-f,\*,#  Do not remove leading zeros |

For help on all the options available to dirbs-import golden list, run:

dirbs-import golden\_list –-help

|  |
| --- |
| Usage: dirbs-import golden\_list [OPTIONS] INPUT\_FILE  Import the Golden list data found in INPUT into the PostgreSQL database.  NOTE: Use caution when adding entries to the Golden list, as any IMEIs  added to this list will never be blocked.  Options:  --disable-historic-check Skip checking the size of this import against  the currently stored data.  --pre-hashed TEXT DANGEROUS: The input file contains normalized  IMEIs that have already been hashed using the  MD5 algorithm. If IMEIs have not been  normalized or hashed according to DIRBS Core  rules, the IMEIs in the imported list may not  be correctly excluded from being blocked.  --disable-delta-adds-check If in delta mode, disable verification that  adds in delta list are not already in DB.  --disable-delta-removes-check If in delta mode, disable verification that  removes in delta list are already in DB.  --disable-delta-updates-check If in delta mode, disable verification that  updates in delta list are already in DB.  --delta Switch to delta import mode.  --help Show this message and exit. |

#### golden list delta import

The delta list import feature was designed to allow regulators to supply changes to lists in a file rather than the complete file every time. These changes include ‘add’, ‘remove’.

The golden list delta import functionality can be invoked by the command line option:

dirbs-import golden\_list <golden\_list\_delta\_zip\_file>

The .zip file is expected to contain a .csv file containing the delta golden list data. It is run through the CSV pre-validator GoldenListDeltaSchemaData.csvs located at /opt/dirbs/etc/schema to ensure it conforms to the expected format by DIRBS Core.

The golden\_imei field is used as a key column to uniquely identify an entry in the list   
(see Section 1.2).

Sample delta .csv file

golden\_imei,change\_type

622222222222222,add

633333333333333,remove

### Local stolen list data – dirbs-import stolen\_list

To import a .zip version of the local stolen list, run:

dirbs-import stolen\_list <stolen\_list\_zip\_file>

The stolen list inputs IMEIs of stolen devices and the reported stolen date.

The .zip file is expected to contain a .csv file containing the list of data. It is run through the CSV pre-validator StolenListSchema.csvs located at /opt/dirbs/etc/schema to ensure it conforms to the expected format by DIRBS Core (see Section 1.2).

Table 3 -19 lists the expected header columns and format. These fields can be in any order and case.

Table 3-19 stolen list fields and format

|  |  |  |
| --- | --- | --- |
| Field | Mandatory (M)  Optional (O) | Expected format |
| imei | M | * Digits, Length (1 – 16)   + Valid digits: 0-9,A-F,a-f,\*,#   + 14 leading digits for good IMEI Records (check digit/software version is stripped on import) |
| reporting\_date | M | * yyyymmdd format1, e.g., 20160703   + Starting with year 2000   + yy: 20XX, where X=0-9   + mm: 00-12   + dd: 00 – 31 (must correspond to days in month/year) |
| 1 Empty string allowed, whitespace not allowed. | | |

For help on all the options available to dirbs-import stolen\_list, run:

dirbs-import stolen\_list –help

|  |
| --- |
| Usage: dirbs-import stolen\_list [OPTIONS] INPUT\_FILE  Import the Stolen List data found in INPUT into the PostgreSQL database.  Options:  --disable-historic-check Skip checking the size of this import against  the currently stored data.  --disable-delta-adds-check If in delta mode, disable verification that  adds in delta list are not already in DB.  --disable-delta-removes-check If in delta mode, disable verification that  removes in delta list are already in DB.  --disable-delta-updates-check If in delta mode, disable verification that  updates in delta list are already in DB.  --delta Switch to delta import mode.  --help Show this message and exit. |

#### stolen list delta import

The delta list import feature was designed to allow regulators to supply changes to lists in a file rather than the complete file every time. These changes include ‘add’, ‘remove’ or ‘update’.

The stolen list delta import functionality can be invoked by the command line option:

dirbs-import stolen\_list <stolen\_list\_delta\_zip\_file>

The .zip file is expected to contain a .csv file containing the delta stolen list data. It is run through the CSV pre-validator StolenListDeltaSchemaData.csvs located at /opt/dirbs/etc/schema to ensure it conforms to the expected format by DIRBS Core.

The imei field is used as a key column to uniquely identify an entry in the list (see Section 1.2).

Sample delta .csv file

imei,reporting\_date,change\_type

622222222222222,20180909,add

633333333333333,20180909,remove

644444444444444,20180909,update

### pairing list data – dirbs-import pairing\_list

To import a .zip version of the complete pairing list, run:

dirbs-import pairing\_list <pairing\_list\_zip\_file>

The pairing list inputs IMEI-IMSI pairs that will be excluded from blocking.

The .zip file is expected to contain a .csv file containing the list of IMEI-IMSI pairs. It is run through the CSV pre-validator PairingListSchema.csvs located at /opt/dirbs/etc/schema to ensure it conforms to the expected format by DIRBS Core   
(see Section 1.2).

Table 3 -20 lists the expected header columns and format. These fields can be in any order and case.

Table 3-20 pairing list fields and format

|  |  |  |
| --- | --- | --- |
| Field | Mandatory (M)  Optional (O) | Expected format |
| imei | M | Digits, Length (1 – 16)  Valid digits: 0-9,A-F,a-f,\*,#  14 leading digits for good IMEI Records(check digit/software version is stripped on import) |
| imsi | M | Digits, Length (1-15)  Valid digits: 0-9 |

For help on all the options available to dirbs-import pairing\_list, run:

dirbs-import pairing\_list –-help

|  |
| --- |
| Usage: dirbs-import pairing\_list [OPTIONS] INPUT\_FILE  Import the Pairing List data found in INPUT into the PostgreSQL database.  Options:  --disable-historic-check Skip checking the size of this import against  the currently stored data.  --disable-delta-adds-check If in delta mode, disable verification that  adds in delta list are not already in DB.  --disable-delta-removes-check If in delta mode, disable verification that  removes in delta list are already in DB.  --disable-delta-updates-check If in delta mode, disable verification that  updates in delta list are already in DB.  --delta Switch to delta import mode.  --help Show this message and exit. |

#### pairing list delta import

The delta list import feature was designed to allow regulators to supply changes to lists in a file rather than the complete file every time. These changes include ‘add’ or ‘remove’.

The pairing list delta import functionality can be invoked by the command line option:

dirbs-import pairing\_list <pairing\_list\_delta\_zip\_file>

The .zip file is expected to contain a .csv file containing the delta pairing list data. It is run through the CSV pre-validator PairingListDeltaSchemaData.csvs located at /opt/dirbs/etc/schema to ensure it conforms to the expected format by DIRBS Core.

The imei, imsi fields are a key column to uniquely identify an entry in the list   
(see Section 1.2).

Sample delta .csv file

imei,imsi,change\_type

00333333333333,003333333333333,add

012222222222222,012222222222222,add

02333333333333,023333333333333,remove

### registration list data – dirbs-import registration\_list

To import a .zip version of the complete import list, run:

dirbs-import registration\_list <registration\_list\_zip\_file>

The purpose of the pairing list is to input IMEIs that have been registered.

The .zip file is expected to contain a .csv file containing the list of approved IMEIs. It is run through the CSV pre-validator RegistrationListSchema.csvs located at /opt/dirbs/etc/schema to ensure it conforms to the expected format by DIRBS Core   
(see Section 1.2).

Table 3-21 registration list fields and format

|  |  |  |
| --- | --- | --- |
| Field | Mandatory (M)  Optional (O) | Expected format |
| approved\_imei | M | Digits, Length (1 – 16)  Valid digits: 0-9,A-F,a-f,\*,# |
| make | O | String, Length (1-4096) 2 |
| model | O | String, Length (1-4096) 2 |
| status | O | String, Length (1-4096) 2 |
| model\_number | O | String, Length (1-4096) 2 |
| brand\_name | O | String, Length (1-4096) 2 |
| device\_type | O | String, Length (1-4096) 2 |
| radio\_interface | O | String, Length (1-4096) 2 |

For help on all the options available to dirbs-import registration\_list, run:

dirbs-import registration\_list –-help

|  |
| --- |
| Usage: dirbs-import registration\_list [OPTIONS] INPUT\_FILE  Import the Registration list data found in INPUT into the PostgreSQL  database.  Options:  --disable-historic-check Skip checking the size of this import against  the currently stored data.  --disable-delta-adds-check If in delta mode, disable verification that  adds in delta list are not already in DB.  --disable-delta-removes-check If in delta mode, disable verification that  removes in delta list are already in DB.  --disable-delta-updates-check If in delta mode, disable verification that  updates in delta list are already in DB.  --delta Switch to delta import mode.  --help Show this message and exit. |

#### registration list delta import

The delta list import feature was designed to allow regulators to supply changes to lists in a file rather than the complete file every time. These changes include ‘add’ or ‘remove’.

The pairing list delta import functionality can be invoked by the command line option:

dirbs-import registration\_list <registration\_list\_delta\_zip\_file>

The .zip file is expected to contain a .csv file containing the delta registration list data. It is run through the CSV pre-validator RegistrationListDeltaSchemaData.csvs located at /opt/dirbs/etc/schema to ensure it conforms to the expected format by DIRBS Core.

The approved\_imei field is used as a key column to uniquely identify an entry in the list (see Section 6.5).

Sample delta .csv file

approved\_imei, make, model, status, model\_number, brand\_name, device\_type, radio\_interface, change\_type

10000000000000,,,,,,,add

10000000000001,,,,,,,remove

10000000000002,,,,,,,add

## Automating import of new files

### Sample Makefiles and Jenkins

The dirbs-import command works very well for a single file, but DIRBS Core also provides the ability to monitor a directory for new files and automatically import them. This functionality is provided by the standard UNIX utility make.

We have provided Makefiles in the release distributables under “etc/makefiles”. Once installed, these Makefiles are located at /opt/dirbs/etc/makefiles. There is a different Makefile for each type of import.

Whenever make is invoked, it looks for files where there is no corresponding .processed file, or where the source file is newer than the .processed file. For each file it finds that matches the previous criteria, dirbs-import imports the file and creates a .processed file.

This approach is very flexible and can be integrated with crontabs or with a more sophisticated approach using Jenkins. In both approaches, either a crontab entry or Jenkins job would be created for each type of import and for each operator.

Sample Makefile invocations

* To import GSMA data:

make -f /opt/dirbs/etc/makefiles/tac\_db\_import.mk GSMA\_HOME=/data/  
gsma\_tac all

* To import operator data with operator ID operator\_id:

make -f /opt/dirbs/etc/makefiles/operator\_import.mk OPERATOR\_ID=operator\_id all

* To import stolen\_list data:

make -f /opt/dirbs/etc/makefiles/stolen\_list\_import.mk STOLEN\_LIST\_HOME=/data/stolen\_list all

* To import pairing\_list data:

make -f /opt/dirbs/etc/makefiles/pairing\_list\_import.mk PAIRING\_LIST\_HOME=/data/pairing\_list all

* To import registration\_list data:

make -f /opt/dirbs/etc/makefiles/registration\_list\_import.mk REGISTRATION\_LIST\_HOME=/data/registration\_list all

* To import golden\_list data:

make -f /opt/dirbs/etc/makefiles/golden\_list\_import.mk GOLDEN\_LIST\_HOME=/data/golden\_list all

Jenkins

If you are using Jenkins to trigger the above Makefile invocations, Jenkins will not, by default, create a login shell, and the DIRBS Core virtualenv will not be activated. In this case, virtualenv activation must precede the call to make:

. /home/dirbs/dirbs-venv/bin/activate

This makes the total command, as run under Jenkins via an SSH slave, similar to:

. /home/dirbs/dirbs-venv/bin/activate && make –f /opt/dirbs/etc/makefiles/

stolen\_list\_import.mk STOLEN\_LIST\_HOME=/data/stolen\_list all

## Classification of IMEIs – dirbs-classify

The dirbs-classify command runs analysis on imported data, based on the configured conditions in the .yml configuration file. Analysis should be run prior to running dirbs-listgen and dirbs-reports.

A sample configuration for the conditions in this section is provided in Appendix B.

Table 3 -22 lists the implemented dimensions and their parameters in release 7.0.0.

Table 3-22 Implemented dimensions and parameters

|  |  |
| --- | --- |
| Asset | Function |
| gsma\_not\_found | Determines whether an IMEI is in the GSMA TAC database  **Note:** Do not use this condition if there is a live DRS enforcing GSAM not found. |
| stolen\_list | Matches IMEIs on the local stolen list |
| duplicate\_threshold | * Matches duplicate IMEIs where the number of triplets (IMEI/IMSI/MSISDN combinations) with IMEI exceeds the threshold over configurable period * Required parameters:   + threshold: Threshold of IMSIs that an IMEI must be seen with to be considered a duplicate (inclusive)   + period\_days or period\_months: Number of days or months in history to consider for duplicate analysis (only one of these can be specified) |
| duplicate\_daily\_avg | * Matches duplicate IMEIs where the average daily number of IMSIs seen with that IMEI over a configurable period exceeds a configurable threshold if that IMEI was seen on at least a configurable number of days during that period * Required parameters:   + threshold: Floating point number of daily average IMSIs that an IMEI must be seen with to be considered a duplicate (inclusive)   + period\_days or period\_months: Number of days or months in history to consider for duplicate analysis (only one of these can be specified)   + min\_seen\_days: Minimum number of days that an IMEI must be seen before it can be considered a duplicate (used to avoid averaging a small number of data points) |
| malformed\_imei | * Matches IMEIs containing a non-digit character in the first 14 characters * Matches IMEIs that are not 14 characters in length |
| not\_on\_registration\_list | Matches IMEIs that do not appear on the registration list |
| inconsistent\_rat | Matches IMEIs whose observed RAT on the network does not match model capabilities in GSMA TAC DB |
| used\_by\_dirbs\_subscriber | * Matches IMEIs seen with an IMSI belonging to a configured DIRBS operator (MCC-MNC match) * Can be used as part of a compound condition to specify different business rules when IMEI was seen with at least one local DIRBS subscriber * Required parameters:   + lookback\_days: Maximum number of days to look back when considering whether IMEI was seen with a DIRBS subscriber |
| used\_by\_international\_roamer | * Matches IMEIs seen with an IMSI where the MCC did not match one of the configured MCCs for the DIRBS country * Can be used as part of a compound condition to specify different business rules when IMEI was seen with at least one international roamer * Required parameters:   + lookback\_days: Maximum number of days to look back when considering whether IMEI was seen with an international roamer |
| used\_by\_local\_non\_dirbs\_roamer | * Matches IMEIs seen with an IMSI belonging to the DIRBS country but not a configured MCC-MNC * Intended to target an edge case where not all national operators might be analyzed DIRBS and/or where only certain regions were targeted * Can be used as part of a compound condition to define different business rules for these cases * Required parameters:   + lookback\_days: Maximum number of days to look back when considering whether IMEI was seen with a DIRBS subscriber |

To run classification using the conditions specified in the config file, use:

dirbs-classify

For help on all the options available to dirbs-classify, run:

dirbs-classify –-help

|  |
| --- |
| Usage: dirbs-classify [OPTIONS]  DIRBS script to classify IMEIs.  Iterates through all configured conditions and write to the  classification\_state table.  Options:  --conditions TEXT By default, dirbs-classify classifies on all  conditions. Specify a comma-separated list  of condition names if you wish to classify  only on those conditions. The condition name  corresponds to the label parameter of the  condition in the DIRBS configuration file.  --safety-check / --no-safety-check  DANGEROUS: Disables safety check that  ensures that no more than a certain ratio of  IMEIs will be classified.  --curr-date TEXT DANGEROUS: Sets current date in YYYYMMDD  format for testing. By default, uses system  current date.  --version Show the version and exit.  -v, --verbose Print debug console output - file output is  unaffected.  --db-password-prompt If set, will prompt the user for a  PostgreSQL password rather than reading from  config.  --db-user TEXT The PostgreSQL DB database user to connect  as.  --db-name TEXT The PostgreSQL DB database name to connect  to.  --db-port INTEGER The PostgreSQL DB port to connect to.  --db-host TEXT The PostgreSQL DB host to connect to.  --statsd-prefix TEXT The environment prefix to prepend to all  StatsD metrics.  --statsd-port INTEGER The StatsD port to connect to on the  configured host.  --statsd-host TEXT The StatsD host to send metrics to.  --max-db-connections INTEGER The maximum DB connections to use  concurrently during this job.  --max-local-cpus INTEGER The maximum number of local CPUs to use  concurrently during this job.  --help Show this message and exit. |

If a specific limited list of conditions must be run instead of all the conditions listed on the configuration file .dirbs.yml, the --conditions option can be used. Use the condition label to be run.

Example

dirbs-classify --conditions simple\_dimension

where the condition simple\_dimension is the label parameter of the condition in the DIRBS configuration or .dirbs.yml as shown below:

conditions:

- label: simple\_dimension

dimensions:

- module: gsma\_not\_found

grace\_period\_days: 30

blocking: true

reason: Violated simple dimension

max\_allowed\_matching\_ratio: 0.1

This command classifies all the IMEIs and stores the results in the database for list generation. It can be trivially scheduled using a crontab or Jenkins job to allow for daily classification.

## Generating lists – dirbs-listgen

List generation takes place after classification.

To run list generation, run:

dirbs-listgen <output\_dir>

where output\_dir is a directory where the various lists will be output. dirbs-listgen automatically creates a timestamp-based subdirectory under this directory. There is no need for this directory to be empty.

Running listgen with no explicit curr-date parameter will base the end of its lookback window off the most recent operator data date, rather than the current date.

Table 3 -23 lists the different types of lists created by dirbs-listgen.

Table 3-23 DIRBS Core lists

|  |  |
| --- | --- |
| List | Function |
| blacklist | * Lists IMEIs which have met a blocking condition and where the current date has exceeded the block date. * This list is distributed to all operators and is the same for each. |
| notification lists | * Lists IMEIs which have met a blocking condition where the current date is still within the grace period for the condition.   + Does not include any IMEI already on the blacklist. * For each IMEI, we generate subscriber triplets based on imported operator data. There is one row in the list for each triplet. * For each triplet, we determine who the home network is based on the IMSI and the configured MCC/MNC pairs for each configured operator. * If a triplet does not match any MCC/MNC pairing for a configured operator (roamers, etc.), we notify all operators whose data they have been seen in. * Each operator gets a different list containing their subscribers and any fallback triplets seen on their network. |
| exception lists | * Each operator gets a copy of the pairing list, split into per-operator exception lists based again on their IMSI and the configured MCC/MNC pairs for the configured operators. * If a pairing's IMSI does match any MCC/MNC pairing for a configured operator (roamers, etc.), the pairing is placed on every MNO exception list. |

For help on all the options available to dirbs-listgen, run:

dirbs-listgen -–help

|  |
| --- |
| Usage: dirbs-listgen [OPTIONS] OUTPUT\_DIR  DIRBS script to output CSV lists (blacklist, exception, notification) for  the current classification state.  Options:  --version Show the version and exit.  -v, --verbose Print debug console output - file output is  unaffected.  --db-password-prompt If set, will prompt the user for a PostgreSQL  password rather than reading from config.  --db-user TEXT The PostgreSQL DB database user to connect as.  --db-name TEXT The PostgreSQL DB database name to connect to.  --db-port INTEGER The PostgreSQL DB port to connect to.  --db-host TEXT The PostgreSQL DB host to connect to.  --statsd-prefix TEXT The environment prefix to prepend to all  StatsD metrics.  --statsd-port INTEGER The StatsD port to connect to on the  configured host.  --statsd-host TEXT The StatsD host to send metrics to.  --max-db-connections INTEGER The maximum DB connections to use concurrently  during this job.  --max-local-cpus INTEGER The maximum number of local CPUs to use  concurrently during this job.  --curr-date TEXT Sets current date in YYYYMMDD format for  testing. By default, uses system current date.  --no-full-lists If set, disable outputting full lists as CSV  for a performance improvement.  --no-cleanup If set, intermediate tables used to calculate  lists will not be deleted so that they can be  inspected.  --base INTEGER If set, will use this run ID as the base for  the delta CSV lists.  --help Show this message and exit. |

The dirbs-listgen command can be trivially scheduled using a crontab or Jenkins job to allow for daily list generation.

## Generating DIRBS reports – dirbs-report

Several reports can be generated using the dirbs-report command. The dirbs-report report\_type command generates all reports in DIRBS Core (see Table 3 -24).

Table 3-24 Report types

|  |  |
| --- | --- |
| Report commands (types) | Function |
| condition\_imei\_overlaps | Generates per-condition reports showing matched IMEIs seen on more than one MNO network |
| gsma\_not\_found | Generates report of all GSMA not found IMEIs as CSV |
| standard | Generates standard monthly per-operator and country-level reports as HTML, CSV and JSON |
| stolen\_violations | Generates per-MNO list of IMEIs seen on the network after they were reported stolen. |
| top\_duplicates | Generates report listing IMEIs seen with more than 5 IMSIs in a given month and year as CSV |

For help on all the options available to dirbs-report, run:

dirbs-report –-help

|  |
| --- |
| Usage: dirbs-report [OPTIONS] COMMAND [ARGS]...  DIRBS script to output reports (operator and country) for a given MONTH  and YEAR.  Options:  --version Show the version and exit.  -v, --verbose Print debug console output - file output is  unaffected.  --db-password-prompt If set, will prompt the user for a PostgreSQL  password rather than reading from config.  --db-user TEXT The PostgreSQL DB database user to connect as.  --db-name TEXT The PostgreSQL DB database name to connect to.  --db-port INTEGER The PostgreSQL DB port to connect to.  --db-host TEXT The PostgreSQL DB host to connect to.  --statsd-prefix TEXT The environment prefix to prepend to all StatsD  metrics.  --statsd-port INTEGER The StatsD port to connect to on the configured host.  --statsd-host TEXT The StatsD host to send metrics to.  --help Show this message and exit.  Commands:  condition\_imei\_overlaps Generate per-condition reports showing...  gsma\_not\_found Generate report of all GSMA not found IMEIs.  standard Generate standard monthly operator and...  stolen\_violations Generate per-MNO list of IMEIs seen on the...  top\_duplicates Generate report listing IMEIs seen with more... |

dirbs-report condition\_imei\_overlaps –-help

|  |
| --- |
| Usage: dirbs-report condition\_imei\_overlaps [OPTIONS] MONTH YEAR OUTPUT\_DIR  Generate per-condition reports showing matched IMEIs seen on more than one  MNO network.  Options:  --max-db-writers INTEGER The maximum write-intensive DB connections  to use concurrently during this job.  --max-db-connections INTEGER The maximum DB connections to use  concurrently during this job.  --max-local-cpus INTEGER The maximum number of local CPUs to use  concurrently during this job.  --debug-query-performance Enable this to print out more stats about  duration of queries during stats generation.  --disable-data-check Disable check to validate existence of data  for all configured operators in this  reporting month.  --disable-retention-check Disable check that stops reports being run  for months outside the retention period.  --force-refresh / --no-refresh Whether data in report should be refreshed  from latest data or from previously-  calculated data (default: --no-refresh).  --help Show this message and exit. |

dirbs-report gsma\_not\_found –-help

|  |
| --- |
| Usage: dirbs-report gsma\_not\_found [OPTIONS] MONTH YEAR OUTPUT\_DIR  Generate report of all GSMA not found IMEIs.  Options:  --max-db-writers INTEGER The maximum write-intensive DB connections  to use concurrently during this job.  --max-db-connections INTEGER The maximum DB connections to use  concurrently during this job.  --max-local-cpus INTEGER The maximum number of local CPUs to use  concurrently during this job.  --debug-query-performance Enable this to print out more stats about  duration of queries during stats generation.  --disable-data-check Disable check to validate existence of data  for all configured operators in this  reporting month.  --disable-retention-check Disable check that stops reports being run  for months outside the retention period.  --force-refresh / --no-refresh Whether data in report should be refreshed  from latest data or from previously-  calculated data (default: --no-refresh).  --help Show this message and exit. |

dirbs-report standard –-help

|  |
| --- |
| Usage: dirbs-report standard [OPTIONS] MONTH YEAR OUTPUT\_DIR  Generate standard monthly operator and country-level reports.  Options:  --max-db-writers INTEGER The maximum write-intensive DB connections  to use concurrently during this job.  --max-db-connections INTEGER The maximum DB connections to use  concurrently during this job.  --max-local-cpus INTEGER The maximum number of local CPUs to use  concurrently during this job.  --debug-query-performance Enable this to print out more stats about  duration of queries during stats generation.  --disable-data-check Disable check to validate existence of data  for all configured operators in this  reporting month.  --disable-retention-check Disable check that stops reports being run  for months outside the retention period.  --force-refresh / --no-refresh Whether data in report should be refreshed  from latest data or from previously-  calculated data (default: --no-refresh).  --help Show this message and exit. |

dirbs-report stolen\_violations –-help

|  |
| --- |
| Usage: dirbs-report stolen\_violations [OPTIONS] OUTPUT\_DIR  Generate per-MNO list of IMEIs seen on the network after they were  reported stolen.  Options:  --max-db-writers INTEGER The maximum write-intensive DB connections to  use concurrently during this job.  --max-db-connections INTEGER The maximum DB connections to use concurrently  during this job.  --max-local-cpus INTEGER The maximum number of local CPUs to use  concurrently during this job.  --newer-than TEXT Include violations newer than the date passed  in YYYYMMDD format.  --filter-by-conditions TEXT Specify a comma-separated list of condition  names if you wish to filter by those  conditions.  --help Show this message and exit. |

dirbs-report top\_duplicates –-help

|  |
| --- |
| Usage: dirbs-report top\_duplicates [OPTIONS] MONTH YEAR OUTPUT\_DIR  Generate report listing IMEIs seen with more than 5 IMSIs in a given month  and year.  Options:  --max-db-writers INTEGER The maximum write-intensive DB connections  to use concurrently during this job.  --max-db-connections INTEGER The maximum DB connections to use  concurrently during this job.  --max-local-cpus INTEGER The maximum number of local CPUs to use  concurrently during this job.  --debug-query-performance Enable this to print out more stats about  duration of queries during stats generation.  --disable-data-check Disable check to validate existence of data  for all configured operators in this  reporting month.  --disable-retention-check Disable check that stops reports being run  for months outside the retention period.  --force-refresh / --no-refresh Whether data in report should be refreshed  from latest data or from previously-  calculated data (default: --no-refresh).  --help Show this message and exit. |

output\_dir is the existing directory where HTML, JS, CSS, CSV, and JSON files will be output. dirbs-report automatically creates a timestamp-based subdirectory under this directory so there is no need for this directory to be empty.

### dirbs-report directory structure

Generated output from the dirbs-report command will be placed in the specified output\_dir.

The output\_dir will contain the HTML, JS, CSS, CSV, and JSON files, and based on the following directory naming convention:

‘report’\_’subcommand’\_’timestamp’\_’run\_id’\_’class\_run\_id’\_’data\_id’\_’month’\_’year’

where:

* subcommand is the dirbs-report subcommand
  + standard, gsma\_not\_found, top\_duplicates, condition\_imei\_overlaps, stolen\_violations
* timestamp is the run\_id\_start\_time in the job\_metadata table
  + Format is %Y%m%d\_%H%M%S, i.e., 20171102\_051731
* run\_id increments each time a report is run, i.e., ‘run\_id\_4
* class\_run\_id is the classification id of the last successful dirb-classify run,   
  i.e., ‘class\_id\_3’
* data\_id, i.e., data\_id\_1
* month, i.e., month\_7
* year, i.e., year\_2016

Sample listing of directory names for various subcommands

* standard

report\_standard\_20171102\_052206\_run\_id\_5\_class\_id\_3\_data\_id\_1\_month\_7\_year\_2016

* condition\_imei\_overlaps

This is the same name structure as gsma\_not\_found and top\_duplicate subcommands.  
data\_id is not used for these subcommands.

report\_condition\_imei\_overlaps\_20171102\_052800\_run\_id\_6\_class\_id\_3\_month\_11\_year\_2016

* stolen\_violation\_directory

Month, year and data\_id are not used for this subcommand

report\_stolen\_violations\_20171102\_051731\_run\_id\_4\_class\_id\_3

1. Visual reports depend on the JSON data, so it is not possible to publish just the HTML, CSS and JS files. Due to security restrictions imposed by the browser, HTML files generated by dirbs-report must be hosted by a webserver rather than opened locally from the filesystem. If you open the reports from the file system, you will receive an alert box stating that the JSON data could not be loaded.

## Accessing the API server

The API server provides information on the data catalog, job metadata, TAC, IMEI, MSISDN, and DIRBS code and schema version

Assuming that you have published the container's port 5000 to the host using the -p 5000:5000 option to docker run, you should be able to open a web browser on the host machine and access the API server on the:

* Data catalog API Version 1.0 (see Section 3.8.1)
* Data catalog API Version 2.0 (see Section 3.8.2)
* Job metadata API Version 1.0 (see Section 3.8.3)
* Job metadata API Version 2.0 (see Section 3.8.4)
* TAC API Version 1.0 (see Section 3.8.5)
* TAC API Version 2.0 (see Section 3.8.6)
* IMEI API Version 1.0 (see Section 3.8.7)
* IMEI API Version 2.0 (see Section 3.8.8)
* MSISDN API Version 1.0 (see Section 3.8.9)
* MSISDN API Version 2.0 (see Section 3.8.10)
* Version API Version 1.0 (see Section 3.8.11)
* Version API Version 2.0 (see Section 3.8.12)

### Data catalog API (Version 1.0)

Table 3-25 Data catalog API

|  |  |
| --- | --- |
| API endpoint | Description |
| /api/v1/catalog | Returns last 100 entries from the data\_catalog table sorted by last\_seen timestamp in descending order |
| **/api/v1/catalog?max\_results=1** | * Returns last 'x' entries from the data\_catalog table sorted by last\_seen timestamp in descending order   + 'x' is specified in max\_results parameter |
| **/api/v1/catalog?file\_type=gsma\_tac** | * Returns last 100 entries of file\_type 'x' from the data\_catalog table sorted by last\_seen timestamp in descending order   + 'x' is specified by file\_type parameter |
| **/api/v1/catalog?is\_valid\_zip=True** | * Returns last 100 entries with is\_valid\_zip status equal to 'x' from the data\_catalog table sorted by last\_seen timestamp in descending order   + 'x' is specified in is\_valid\_zip parameter |
| **/api/v1/catalog?modified\_since=20170825** | * Returns last 100 entries with modified\_time greater than equal to 'x' from the data\_catalog table sorted by last\_seen timestamp in descending order   + 'x' is specified in modified\_since parameter |
| **/api/v1/catalog?cataloged\_since=20170801** | * Returns last 100 entries with last\_seen greater than equal to 'x' from the data\_catalog table sorted by last\_seen timestamp in descending order   + 'x' is specified in cataloged\_since parameter |
| **/api/v1/catalog?modified\_since=20170825&is\_ valid\_zip=True** | * Returns last 100 entries with last\_seen greater than equal to 'x' and is\_valid\_zip equal to 'y' from the data\_catalog table sorted by last\_seen timestamp in descending order   + 'x' is specified in cataloged\_since and  'y' is\_valid\_zip parameters |

### Data catalog API (Version 2.0)

The Data Catalog API Version 2.0 is the extended version of API Version 1.0, it includes some extra features (having the previous features as well) which are described in the following:

* **Pagination support** has been introduced in it, now results can be paginated using the keys **limit** and **offset**

e.g api/v2/catalog?limit=10&offset=1

Now this will show total number of 10 results starting from the first. More details can be views in API Specification Guidelines.

* Result **sorting** support has been introduced in ascending or descending order, now results can be sorted based on the **file\_id** of the file in ascending or descending order.

e.g. api/v2/catalog?limit=10&offset=1&order=Ascending

api/v2/catalog?limit=10&offset=1&order=Descending

Now these will sort the results in ascending or descending orders based on the id of the file cataloged.

### Job metadata API (Version 1.0)

Table 3-26 Job metadata API

|  |  |
| --- | --- |
| API endpoint | Description |
| /api/v1/job\_metadata?max\_results=<n>(n defaults to 10) | The number of jobs to show in this list can be configured by the <max\_results> query parameter, which defaults to 10 and **must be a positive integer**. |
| /api/v1/job\_metadata?run\_id=<first\_run\_id>&run\_id=<second\_run\_id*>*(defaults to any run\_id) | Jobs can be filtered by a list of run\_ids using the <run\_id> query parameter, which defaults to any run\_id.  Each run\_id **must be a positive integer**. |
| /api/v1/job\_metadata?command=<first\_command\_name\_without\_quotes>& command=<second\_command\_name\_without\_quotes>(defaults to any command)  e.g. for import job :  /api/v1/job\_metadata?command=dirbs-import&command=dirbs-prune | Job command name can be specified using the <command> query parameter, which defaults to any command.  It is possible to specify more than one command name in the same query using the symbol "&" to concatenate the params.  **Each command name must refer to an existing command**, such as:  *"dirbs-import", "dirbs-classify", "dirbs-prune", "dirbs-listgen", "dirbs-catalog", "dirbs-report", "dirbs-db"*. |
| /api/v1/job\_metadata?subcommand=<subcommand\_name>(defaults to any sub\_command) | Jobs can be filtered by a list of job subcommands using the <subcommand> query parameter, which defaults to any subcommand. |
| /api/v1/job\_metadata?status=error&status=success(defaults to any status) | Jobs can be filtered by a list of job metadata using the <status> query parameter, which defaults to any status. Each job metadata**must be either  'running', 'success' or 'error'**. |
| /api/v1/job\_metadata?show\_details=True(defaults to True) | Extra details for the specific job can be retrieved in the extra\_metadata section in the JSON response by setting <show\_details> query parameter to True, which is the default value.  Show\_details **must have a boolean value: True, False, 0, 1.**If show\_details is set to False, extra\_metadata section will not be included in the JSON response. |
| /api/v1/job\_metadata?show\_details=True&status=error&status=success&max\_results=3&command=dirbs-import&command=dirbs-prune | Query parameters for jobs can be repeated to allow multiple values for the same param (if eligible). All query params are eligible for multiple values except max\_results and show\_details.  Multiple filters can be combined in the same query by adding query parameters separated by the symbol '&' (first query param must start with symbol '?'). |

### Job Metadata API (Version 2.0)

The Job Metadata API Version 2.0 is the extended version of API Version 1.0, it includes some extra features (having the previous features as well) which are described in the following:

* **Pagination support** has been introduced in it, now results can be paginated using the keys **limit** and **offset**

e.g api/v2/job\_metadata?limit=10&offset=1

Now this will show total number of 10 results starting from the first. More details can be views in API Specification Guidelines.

* Result **sorting** support has been introduced in ascending or descending order, now results can be sorted based on the **run\_id** of the job in ascending or descending order.

e.g. api/v2/job\_metadata?limit=10&offset=1&order=Ascending

api/v2/job\_metadata?limit=10&offset=1&order=Descending

Now these will sort the results in ascending or descending orders based on the run id of the job.

### TAC API (Version 1.0)

The TAC API returns relevant data from the GSMA TAC DB. The GSMA TAC fields for NFC, Bluetooth and WLAN are displayed as the raw content from the GSMA TAC DB.

http://localhost:5000/api/v1/tac/<tac\_num>

where <tac\_num> is the 8-digit TAC.

### TAC API (Version 2.0)

The API Version 2.0 supports both **GET** and **POST** methods in this version for the TAC API. The GET method returns the same relevant data from GSMA TAC DB as in the previous version however the TAC POST API (Batch TAC API) accepts 1000 TACs at the same time and returns results for them.

e.g. **curl -X POST --header "Content-Type: application/json" – header "Accept: application/json" -d "{**

**\"tacs\": [**

**\"12345678\",**

**\"56478377\"**

**]**

**}" "**[**http://localhost:5000/api/v2/tac**](http://localhost:5000/api/v2/tac)**"**

Now this will return results for both TACs from GSMA TAC Database. More details can be viewed in API Specifications Guidelines.

### IMEI API (Version 1.0)

The IMEI API returns all known information about the IMEI, as well as results of all 'conditions' evaluated as part of DIRBS Core.

The following realtime checks are also included information:

* Invalid IMEI
* GSMA not found
* Registration status
* IMEIs ever observed on the network

http://localhost:5000/api/v1/imei/<imei>?include\_seen\_with=<0,1,true,  
false>

where

include\_seen\_with determines whether or not the seen\_with field will be present in the response.

If the include\_seen\_with parameter is not set, it defaults to 0, meaning no seen\_with data will be calculated or sent.

### IMEI API (Version 2.0)

The IMEI API Version 2.0 has been divided into five different APIs which are as follow:

1. **IMEI API:** Returns information about an IMEI as the result of off conditions evaluated as a part of DIRBS Core. It also shows block date, real-time checks such as IMEI observation on network, its pairing status, if it is exempted device or if it is invalid. The IMEI API has two additional blocks in this version which are **registration\_status** which shows IMEI status from Registration List and **stolen\_status** which shows IMEI status from Stolen List.

It can be accessed as: api/v2/imei/<imei>

1. **IMEI Info API:** Returns information about an IMEI’s Brand Name, Device Type, Make, Model, Model Number, Radio Interfaces and current Status of it in the DIRBS-Core.

It can be accessed as: api/v2/imei/<imei>/info

1. **IMEI Pairings API:** Returns information core knows about the IMSI’s paired with the IMEI in the device pairing system. It supports pagination as well. It also supports results sorting in ascending or descending order based on MSISDN as key.

It can be accessed using: api/v2/imei/<imei>/pairings

It accepts **limit**, **offset** as pagination parameters and **Ascending** or **Descending** as sorting orders.

3. **IMEI Subscribers API:** Returns information Core knows about the IMSI- MSISDN pairs the IMEI has been seen with on the network. It supports pagination and sorting as well.

It can be accessed using: api/v2/imei/<imei>/subscribers

It accepts **limit**, **offset** as pagination parameters and **Ascending** or **Descending** as sorting orders.

4. **IMEI Batch API:** Returns information about IMEIs same as in single IMEI API described above. It accepts maximum 1000 IMEI’s at the same time and return results for each one collectively.

It can be accessed using POST: api/v2/imei-batch

### MSISDN API (Version 1.0)

The MSISDN API returns a list of IMEI, IMSI, GSMA Manufacturer, GSMA Model Name for the MSISDN specified:

http://localhost:5000/api/v1/msisdn/<msisdn>

### MSISDN API (Version 2.0)

The MSISDN API returns information about the MSISDN, such as IMEI, IMSI, GSMA Manufacturer, Model Name and Last Seen date for the specified MSISDN. It also returns Information from GSMA TAC Database as well as Device Registration System for the specified MSISDN.

It can be accessed using: api/v2/msisdn/<msisdn>

### Version API (Version 1.0)

This simple API returns the code, DB schema version, report schema version and version of the source code of DIRBS Core.

It can be accessed as: http://localhost:5000/api/v1/version

### Version API (Version 2.0)

This simple API returns the code, DB schema version, report schema version and version of the source code of DIRBS Core.

It can be accessed as: http://localhost:5000/api/v1/version

## Pruning old data

Table 3 -27 lists commands used to prune obsolete data from the DIRBS Core PostgreSQL.

Table 3-27 Prune commands

|  |  |
| --- | --- |
| Prune commands | Function |
| classification\_state | Prune obsolete classification\_state data. |
| triplets | Prune old seen\_triplets data. |

For help on all options available to dirbs-prune, run:

dirbs-prune --help

|  |
| --- |
| Usage: dirbs-prune [OPTIONS] COMMAND [ARGS]...  DIRBS script to prune obsolete data from the DIRBS Core PostgreSQL  database.  Options:  --version Show the version and exit.  -v, --verbose Print debug console output - file output is  unaffected.  --db-password-prompt If set, will prompt the user for a PostgreSQL  password rather than reading from config.  --db-user TEXT The PostgreSQL DB database user to connect as.  --db-name TEXT The PostgreSQL DB database name to connect to.  --db-port INTEGER The PostgreSQL DB port to connect to.  --db-host TEXT The PostgreSQL DB host to connect to.  --statsd-prefix TEXT The environment prefix to prepend to all StatsD  metrics.  --statsd-port INTEGER The StatsD port to connect to on the configured host.  --statsd-host TEXT The StatsD host to send metrics to.  --curr-date TEXT Sets current date in YYYYMMDD format for testing. By  default, uses system current date.  --help Show this message and exit.  Commands:  classification\_state Prune obsolete classification\_state data.  triplets Prune old seen\_triplets data. |

dirbs-prune classification\_state --help

|  |
| --- |
| Usage: dirbs-prune classification\_state [OPTIONS]  Prune obsolete classification\_state data.  Options:  --help Show this message and exit. |

dirbs-prune triplets --help

|  |
| --- |
| Usage: dirbs-prune triplets [OPTIONS]  Prune old seen\_triplets data.  Options:  --help Show this message and exit. |

# Understanding DIRBS Reports

## Standard reports

Standard monthly operator and country-level reports are generated in HTML, JSON and CSV formats. The formats and sections for the country and operator reports are the same. Operator reports are specific to their respective operators configured in the .dirbs.yml file. Country-level reports reflect all the IMEIs seen in the country.

The JSON file has a report schema version associated with any generated standard report and are explicit fields called “report\_schema\_version” and “software\_version”

Expect to see the version number incremented when:

* Fields are added, removed or renamed
* The method of calculation for a field is changed so that it cannot be compared to previous reports

The standard report has a --force-refresh / --no-refresh (default) CLI option:

* --no-refresh reports can be generated very quickly since no numbers are calculated
* --force-refresh tells dirbs-report to re-do stats generation if there is previous data available for the same month and year
* Standard report only looks for previous data with the same report\_schema\_version
  + If schema has changed, dirbs-report will always generate new data

Placeholder reports were created with no data for configured operators that have no data for the month. These reports are only created when the CLI option --disable-data-check is used.

Other CLI options for placeholder reports are:

* --max-db-connections <int>: Determines the number of parallel jobs to perform during stats generation (performance scales linearly with this number).
* --disable-data-check: By default, dirbs-report ensures that there is data available for all operators for the given month and year before generating a report. Disabling this allows a report to be generated even if data for one operator’s data is missing.
* --disable-retention-check: By default, dirbs-report will fail if there is an attempt to generate a report outside the retention period.
* --debug-query-performance: Provides more detail in the console output about query performance during the stats generation phase.

### Country report

#### HTML

The HTML country report covers:

* [Identifier counts](#HTMLIC)
* [Identifier trends](#HTMLIT)
* [Compliance breakdown](#HTMLCB)
* [IMEI compliance trends](#HTMLIMEICT)
* [Conditions breakdown](#HTMLconbreakdown)
* [Condition combinations](#HTMLCC)
* [Blacklist and blacklist violations](#HTMLblacklist)
* [Top models: counts](#HTMLTCcounts)
* [Top models: gross adds](#HTMLTCgrossadds)

1. Figures in this section show graphic representations of the same sections in the JSON report.

Figure 4 -4 shows the main page for country reports in HTML. Navigate to different sections of the report by clicking on the navigation pane on the left.

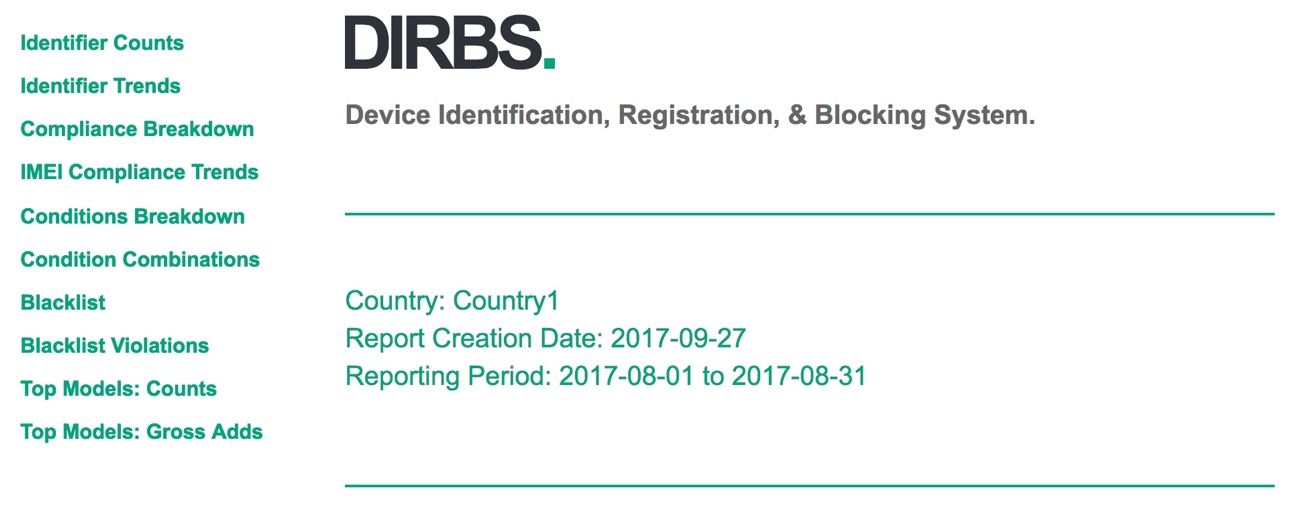


Figure 4-4 Country report main page – HTML

Identifier counts

Identifier counts show a distinct number of:

* IMEIs, MSISDNs, and IMSIs
* Combination pairs of IMEI-IMSI, IMEI-MSISDN, and IMSI-MSISDN
* Triplet combinations of IMEI-IMSI-MSISDN

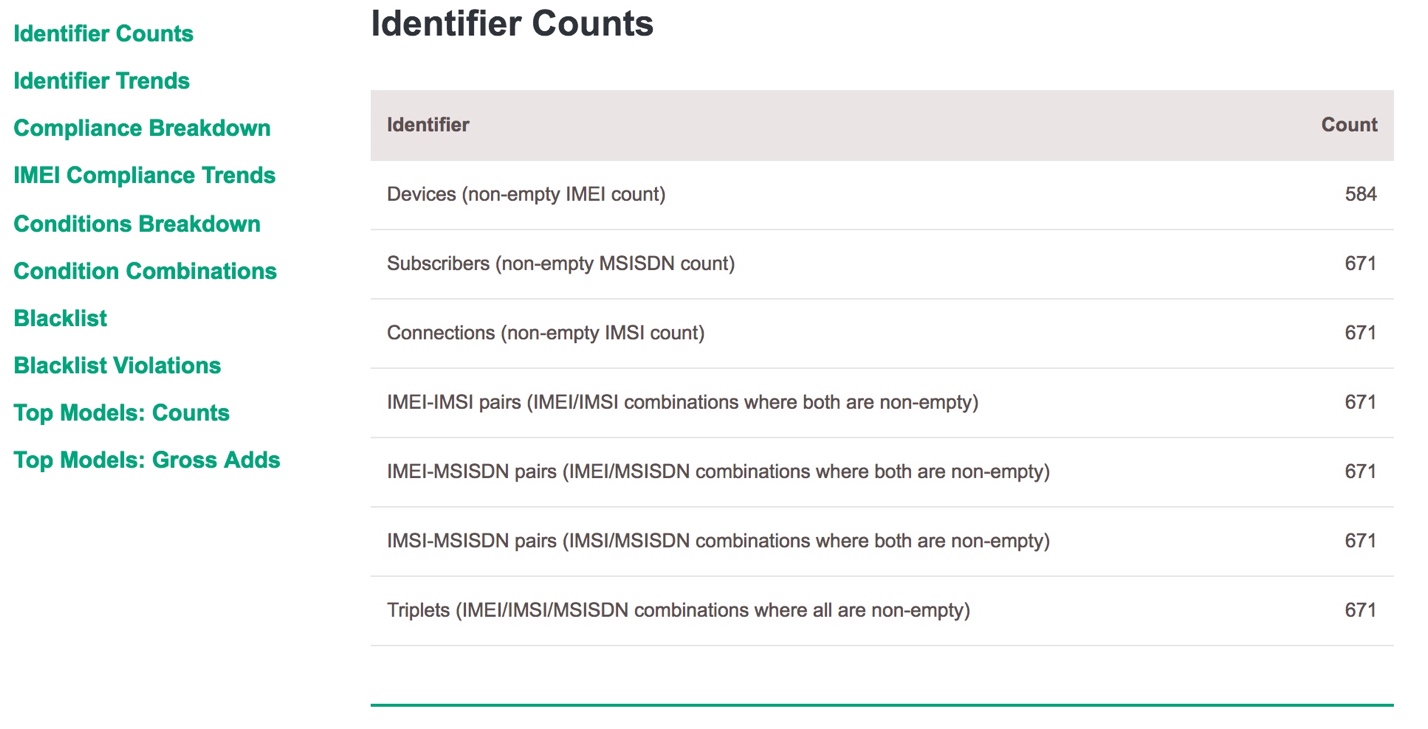


Figure 4-5 Identifier counts

Identifier trends

IMEIs, MSISDNs, and IMSIs counts for the months with data for the period specified in the configuration file.

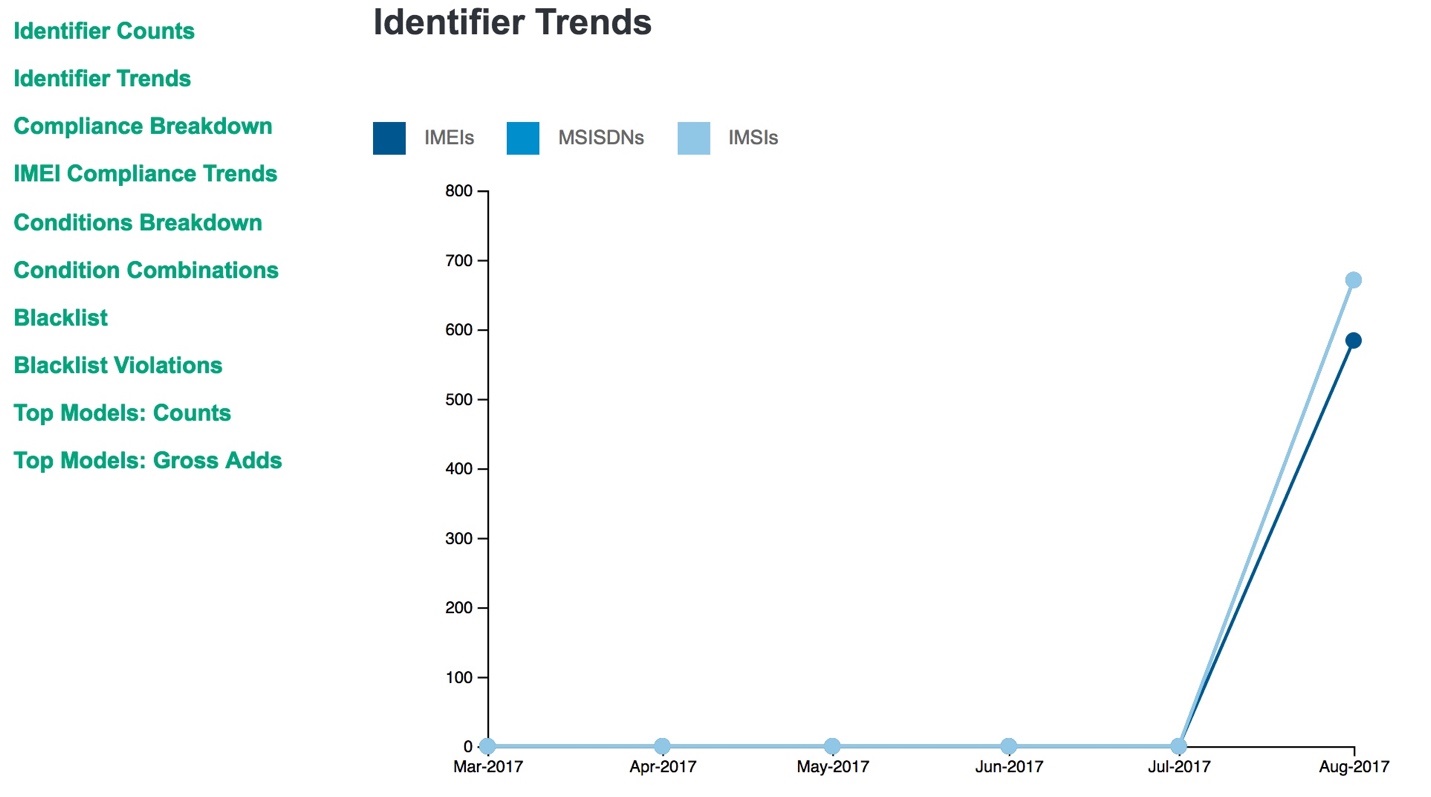


Figure 4-6 Identifier trends

Compliance breakdown

The compliance breakdown shows:

* Compliant IMEIs and triplets that do not meet any conditions or conditions that are non-blocking
* Non-compliant IMEIs and triplets that meet one or more blocking conditions

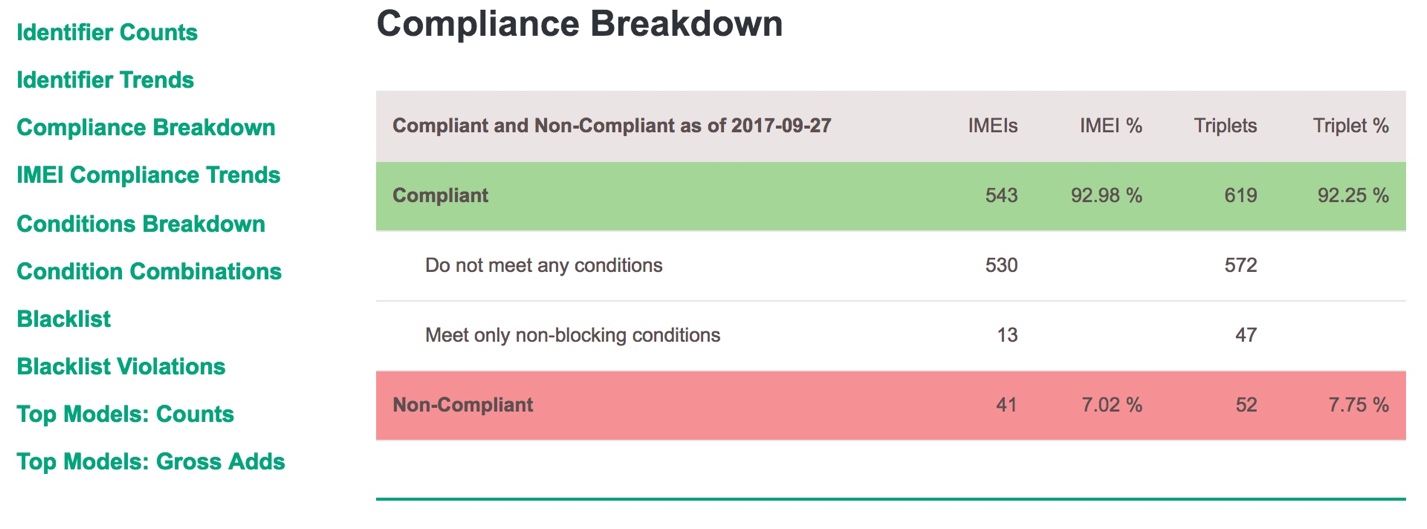


Figure 4-7 Compliance breakdown

IMEI compliance trends

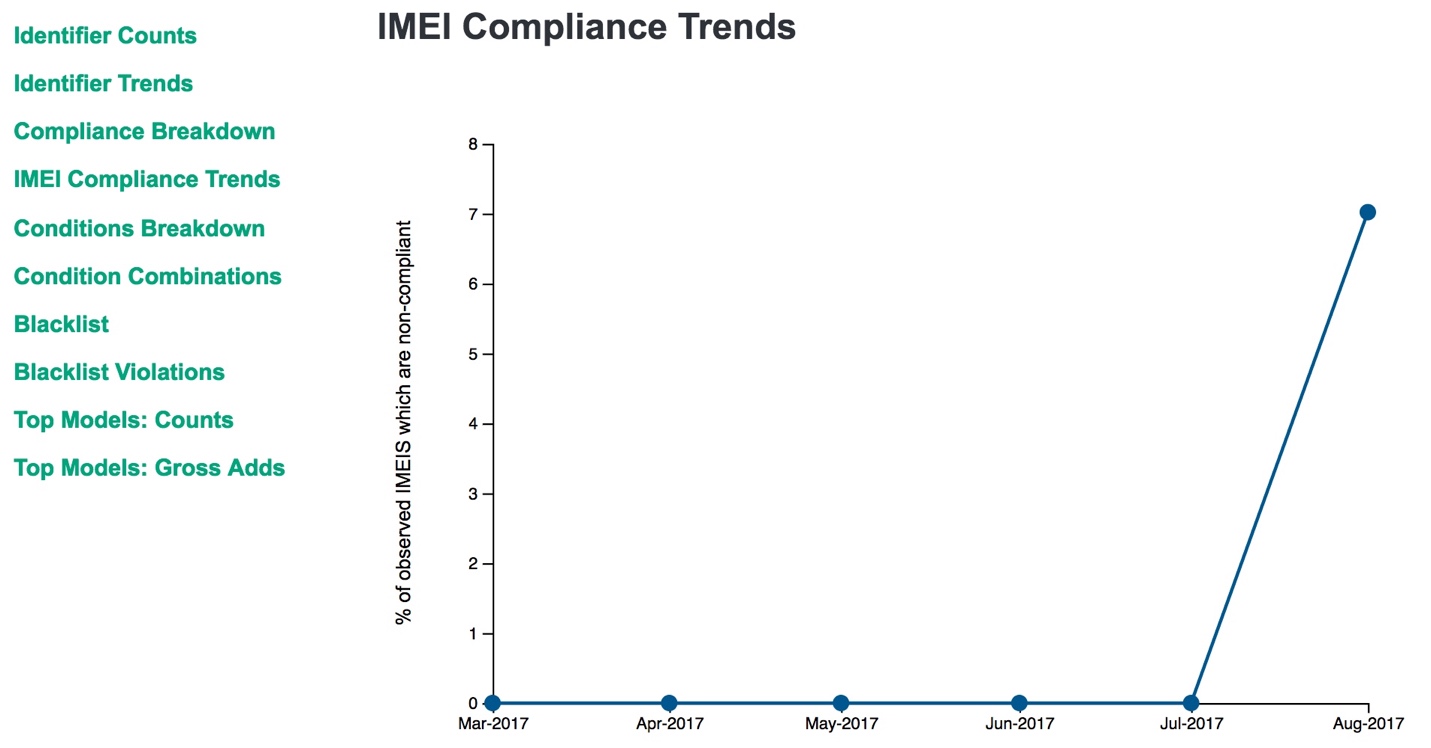


Figure 4-8 IMEI compliance trends

Conditions breakdown

Each condition is independent of each other. An IMEI can meet one or more conditions, and is counted on each condition it meets. The sum of the IMEIs for the conditions breakdown does not equal the number of IMEIs found on the compliance breakdown.



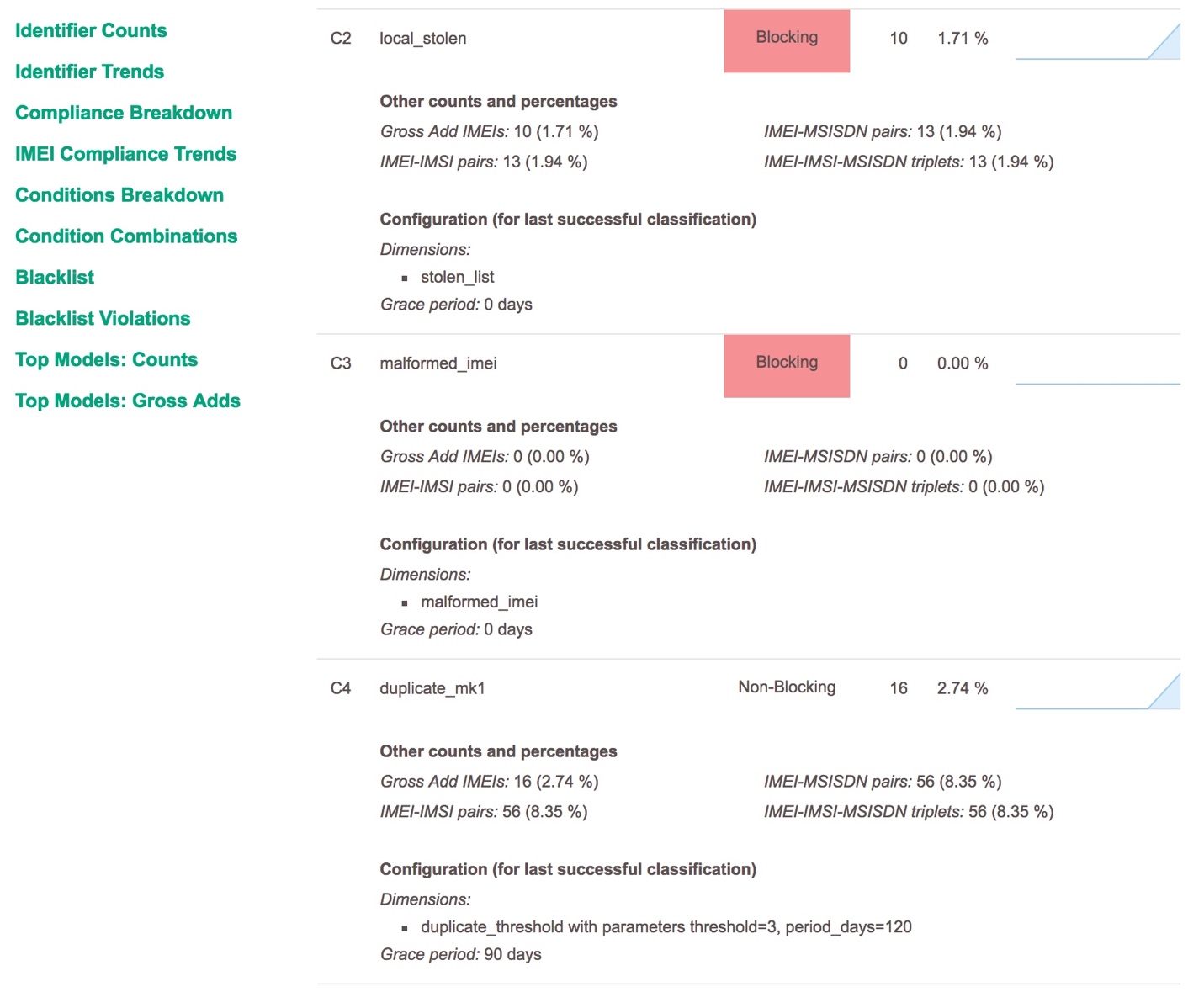


Figure 4-9 Conditions breakdown

Condition combinations

IMEIs, IMEI-IMSIs pairs, and triplets count for the conditions they meet including the combination of conditions.

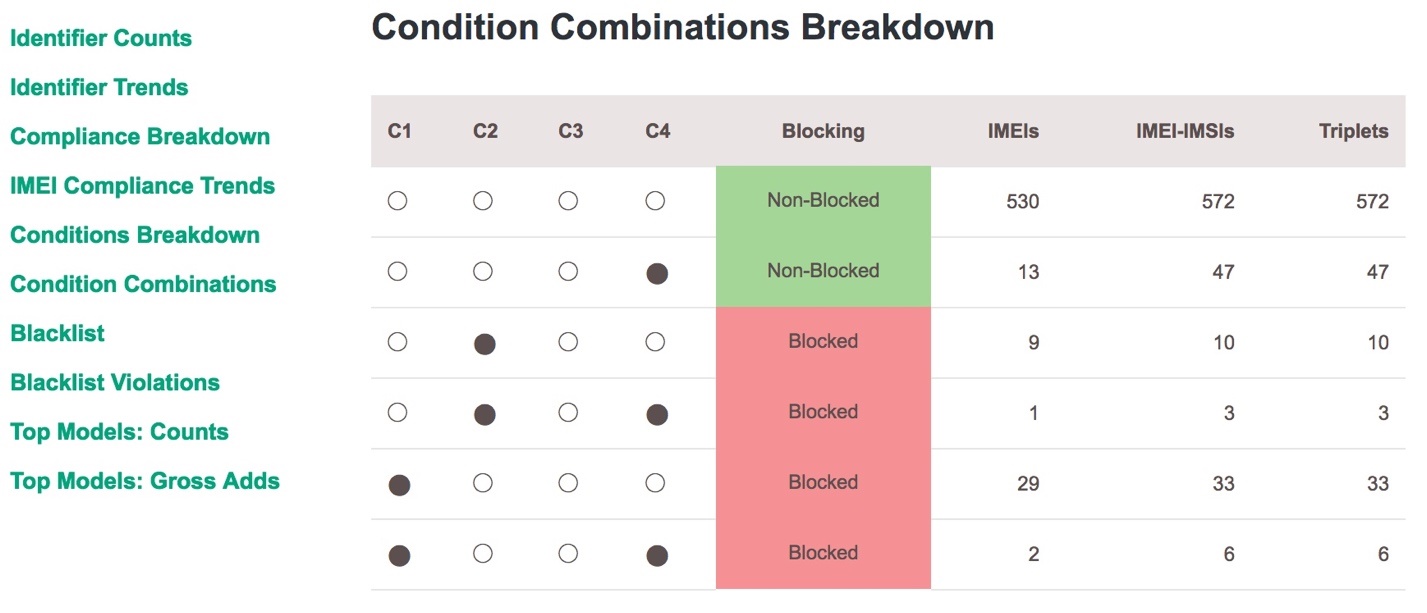


Figure 4-10 Condition combinations

Blacklist and blacklist violations

Blacklists and blacklist violations report the number of blacklisted IMEIs.

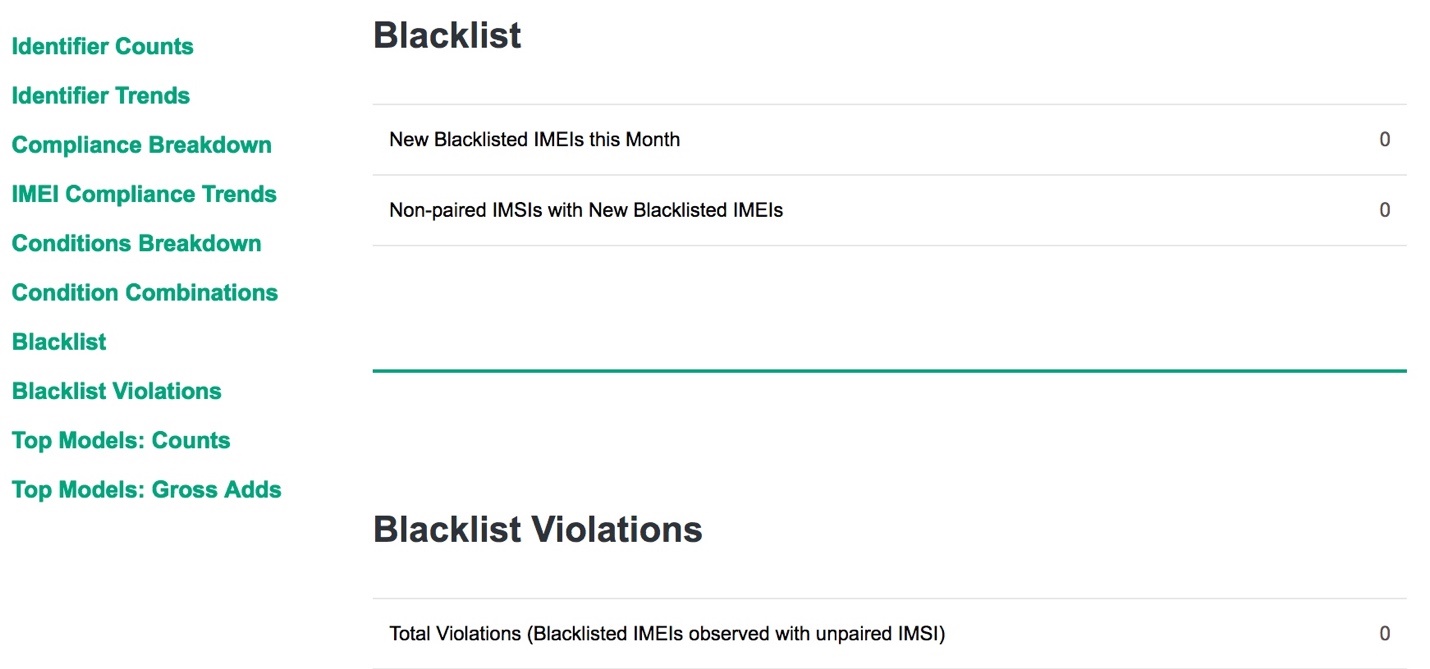


Figure 4-11 Blacklist and blacklist violations

Top models: counts

Top models show the top 10 models by IMEI counts.

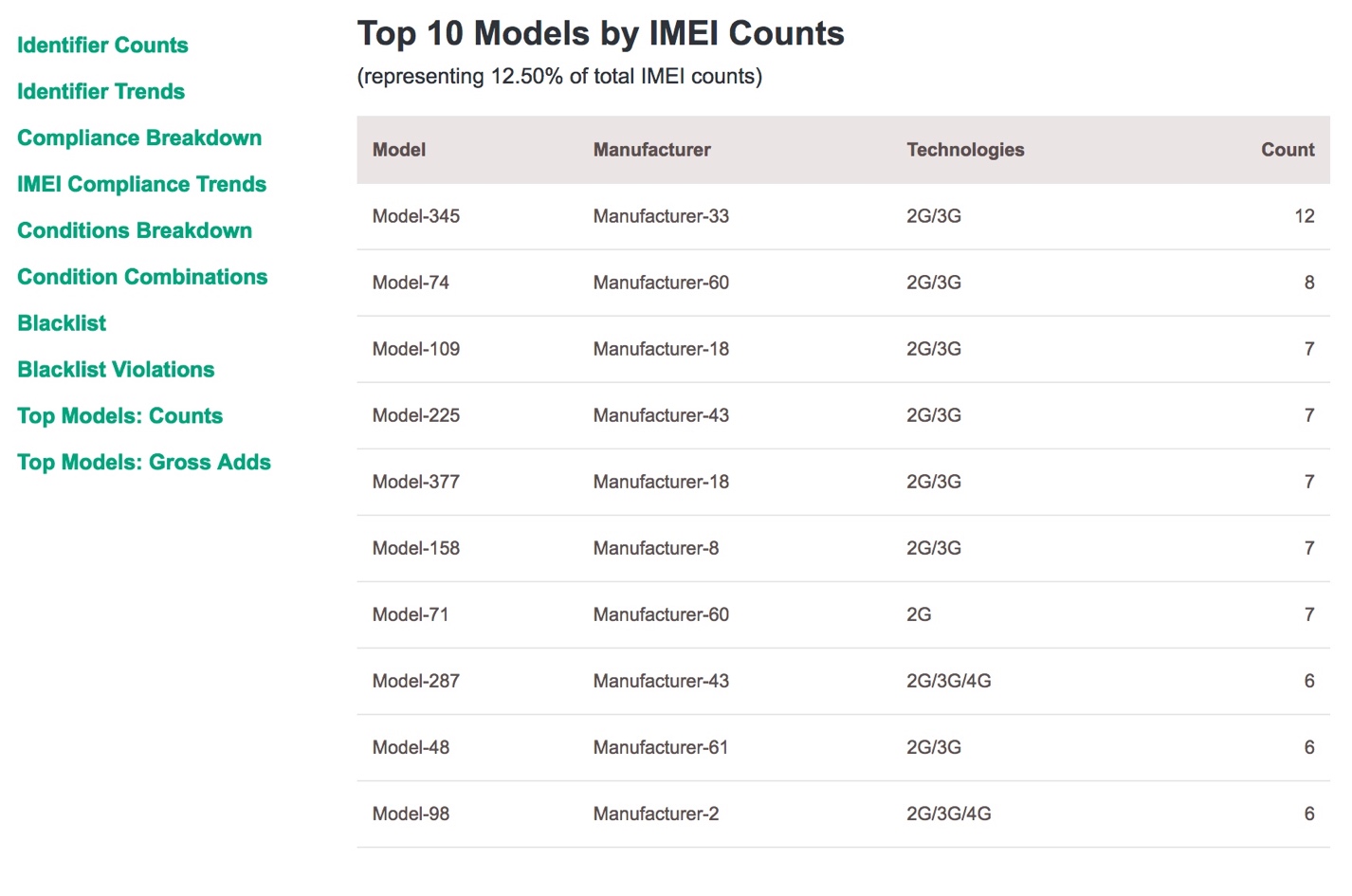


Figure 4-12 Top models: counts

Top models: gross adds

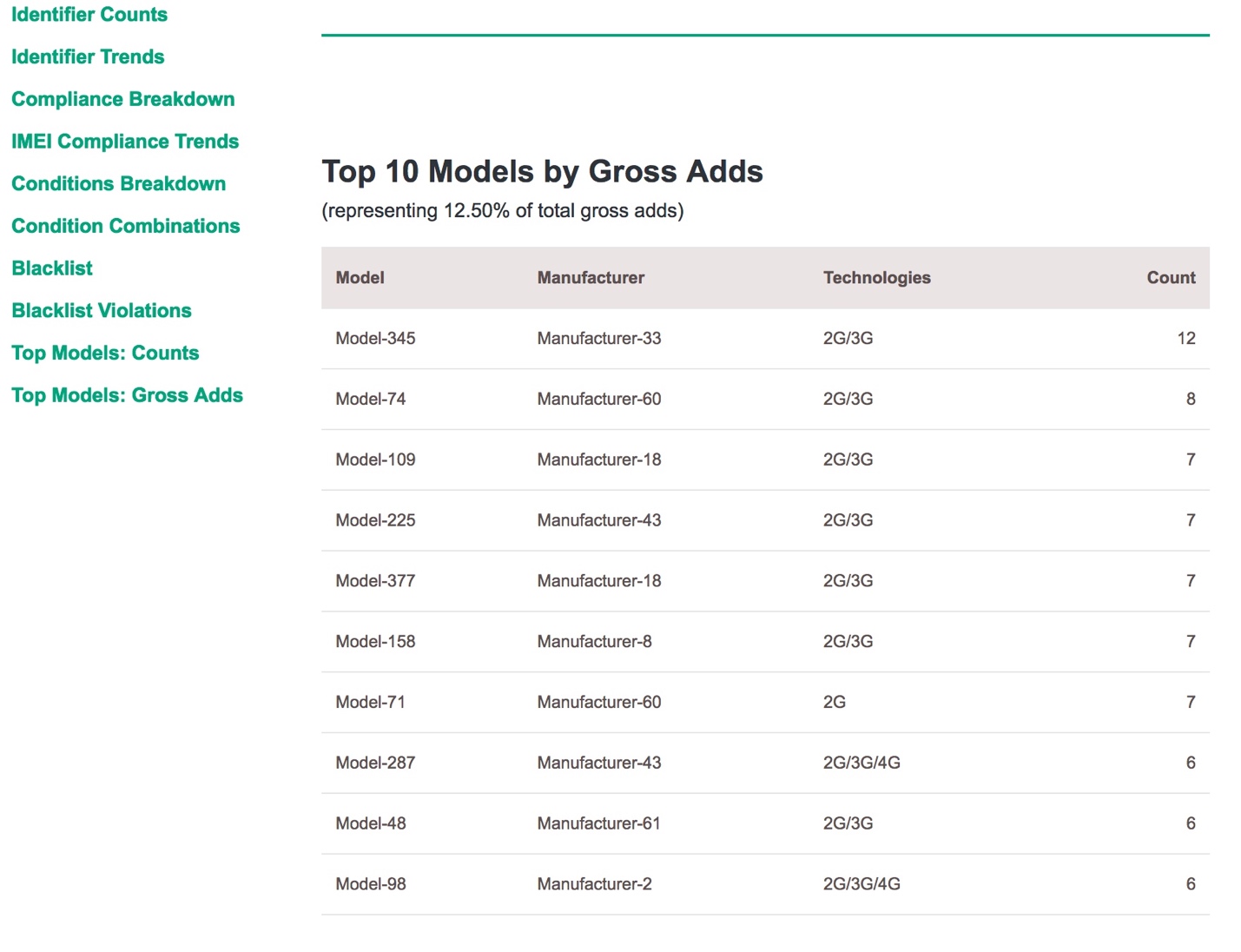


Figure 4-13 Top models: gross adds

#### JSON

The JSON country report covers:

* [Blacklist information](#JSONblacklist)
* [Classification conditions](#JSONCC)
* [Compliance breakdown](#JSONcombreakdown)
* [Condition combinations](#JSONconcombinations)
* [Conditions breakdown](#JSONconbreak)
* [Report name](#JSONreportname)
* [Historic blacklist adds](#JSONHblackadd)
* [Historic compliance breakdown](#JSONHcompbreakdown)
* [Historic conditions breakdown](#JSONHcondbreakdown)
* [Historic IMEI, IMSI, MSISDN, and triplet counts](#JSONHtriplet)
* [IMEI/IMSI and IMSI/IMEI overloading](#JSONoverloading)
* [Daily counts for IMEIs, IMSIs, and MSISDNs](#JSONdailycounts)
* [Top models](#JSONtopmodels)
* [Monthly counts](#JSONmonthly)

Blacklist information

Table 4-28 Blacklist information

|  |  |
| --- | --- |
| Field names | Description |
| blacklist\_adds | * imeis: Number of IMEIs seen on network during this month which were first blocked during the month * non\_paired\_imsis: Number of non-paired IMSIs seen on this network during this month associated with those IMEIs |
| blacklistviolations\_by\_age | Count of IMEIs seen after they were blacklisted, bucketed by the difference in days between when they were last seen during this month and the block date |

|  |
| --- |
| "blacklist\_adds": {  "imeis": 0,  "non\_paired\_imsis": 0  },  "blacklist\_violations\_by\_age": {  "1-2": 0,  "11-20": 0,  "21-30": 0,  "3-5": 0,  "31-90": 0,  "6-10": 0,  "90+": 0  }, |

Classification conditions

Classification number counts are always filtered by data appearing for that operator or for the whole country during the reporting month. For example, if there are 10 GSMA Not Found IMEIs, but only 6 were seen on an operator network during the month, the report will return 6.

The latest classification data is always used, not taken at the end of the reporting month. If a report was run in September for June, it would first take the IMEIs classified at the current date in September. To generate an IMEI count, it takes the subset of what appeared on the network during that month for that operator.

The following classification conditions were configured at the last dirbs-classify execution.

|  |
| --- |
| "classification\_conditions": [  {  "blocking": true,  "config": {  "blocking": true,  "dimensions": [  {  "invert": false,  "module": "gsma\_not\_found",  "parameters": {}  }  ],  "grace\_period\_days": 90,  "label": "gsma\_not\_found",  "max\_allowed\_matching\_ratio": 0.1,  "reason": "TAC not found in GSMA TAC database",  "sticky": false  },  "label": "gsma\_not\_found"  },  {  "blocking": true,  "config": {  "blocking": true,  "dimensions": [  {  "invert": false,  "module": "stolen\_list",  "parameters": {}  }  ],  "grace\_period\_days": 0,  "label": "local\_stolen",  "max\_allowed\_matching\_ratio": 0.1,  "reason": "IMEI found on local stolen list",  "sticky": false  },  "label": "local\_stolen"  },  {  "blocking": true,  "config": {  "blocking": true,  "dimensions": [  {  "invert": false,  "module": "malformed\_imei",  "parameters": {}  }  ],  "grace\_period\_days": 0,  "label": "malformed\_imei",  "max\_allowed\_matching\_ratio": 0.1,  "reason": "Invalid characters detected in IMEI",  "sticky": false  },  "label": "malformed\_imei"  },  {  "blocking": false,  "config": {  "blocking": false,  "dimensions": [  {  "invert": false,  "module": "duplicate\_threshold",  "parameters": {  "period\_days": 120,  "threshold": 3  }  }  ],  "grace\_period\_days": 90,  "label": "duplicate\_mk1",  "max\_allowed\_matching\_ratio": 0.1,  "reason": "Duplicate threshold exceeeded",  "sticky": false  },  "label": "duplicate\_mk1"  } |

Compliance breakdown

High level compliance data can be found in the compliance\_breakdown field. The breakdown lists stats on overall compliance across all configured conditions.

Table 4 -29 lists the properties of the configured conditions. There is also historical data for compliance breakdown in the historic\_compliance\_breakdown field.

Table 4-29 Compliance breakdown

|  |  |
| --- | --- |
| Field names | Description |
| num\_compliant\_imei\_imsis | Number of compliant IMEI-IMSIs (no NULLs) |
| num\_compliant\_imei\_msisdns | Number of compliant IMEI-MSISDNs (no NULLs) |
| num\_compliant\_imeis | Number of compliant IMEIs (no NULLs) |
| num\_compliant\_triplets | Number of compliant IMEI-IMSI-MSISDN triplets (no NULLs) |
| num\_noncompliant\_imei\_imsis\_blocking | IMEI-IMSIs (no NULLs) meeting 1+ blocking condition |
| num\_noncompliant\_imei\_imsis\_info\_only | As above, but meeting only non-blocking conditions |
| num\_noncompliant\_imei\_imsis | Sum of above 2 counts. Redundant |
| num\_noncompliant\_imei\_msisdns\_blocking | IMEI-MSISDNs (no NULLs) meeting 1+ blocking condition |
| num\_noncompliant\_imei\_msisdns\_info\_only | As above, but meeting only non-blocking conditions |
| num\_noncompliant\_imei\_msisdns | Sum of above 2 counts. Redundant |
| num\_noncompliant\_imeis\_blocking | IMEIs (no NULLs) meeting 1+ blocking condition |
| num\_noncompliant\_imeis\_info\_only | As above, but meeting only non-blocking conditions |
| num\_noncompliant\_imeis | Sum of above 2 counts. Redundant |
| num\_noncompliant\_triplets\_blocking | Triplets (no NULLs) meeting 1+ blocking condition |
| num\_noncompliant\_triplets\_info\_only | As above, but meeting only non-blocking conditions |
| num\_noncompliant\_triplets | Sum of above 2 counts. Redundant |

|  |
| --- |
| "compliance\_breakdown": {  "num\_compliant\_imei\_imsis": 572,  "num\_compliant\_imei\_msisdns": 572,  "num\_compliant\_imeis": 530, |

Condition combinations

Table 4 -30 lists the stats for every combination of conditions in the condition\_combination\_table field.

Table 4-30 Condition combinations

|  |  |
| --- | --- |
| Field names | Description |
| combination | * Describes combination of conditions for this entry in the list * If Cond A is True and the rest are False, it means Cond A only * If Cond A and Cond B are true and the rest are False, it means Cond A and Cond B only |
| compliance\_level | * 2 means compliant * 1 means non-compliant but informational * 0 means non-compliant and blocking * This is determined by the config of the conditions selected by combination |
| num\_imeis | Number of matching IMEIs (no NULLs) |
| num\_imei\_gross\_adds | Number of matching gross add IMEIs (no NULLs) |
| num\_imei\_imsis | Number of matching IMEI-IMSIs (no NULLs) |
| num\_imei\_msisdns | Number of matching IMEI-MSISDNs (no NULLs) |
| num\_subscriber\_triplets | Number of matching triplets (no NULLs) |

|  |
| --- |
| "condition\_combination\_table": [  {  "combination": {  "duplicate\_mk1": false,  "gsma\_not\_found": false,  "local\_stolen": false,  "malformed\_imei": false  },  "compliance\_level": 2,  "num\_imei\_gross\_adds": 530,  "num\_imei\_imsis": 572,  "num\_imei\_msisdns": 572,  "num\_imeis": 530,  "num\_subscriber\_triplets": 572  },  {  "combination": {  "duplicate\_mk1": true,  "gsma\_not\_found": false,  "local\_stolen": false,  "malformed\_imei": false  },  "compliance\_level": 1,  "num\_imei\_gross\_adds": 13,  "num\_imei\_imsis": 47,  "num\_imei\_msisdns": 47,  "num\_imeis": 13,  "num\_subscriber\_triplets": 47  }, |

Conditions breakdown

Overall stats about an individual condition can be found in the conditions\_breakdown field. There is an entry for each classification condition (matches label in classification\_conditions). There is also historical data for this in the historic\_conditions\_breakdown field

Table 4-31 Conditions breakdown

|  |  |
| --- | --- |
| Field names | Description |
| num\_imeis | Number of matching IMEIs (no NULLs) |
| num\_imei\_gross\_adds | Number of matching gross add IMEIs (no NULLs) |
| num\_imei\_imsis | Number of matching IMEI-IMSIs (no NULLs) |
| num\_imei\_msisdns | Number of matching IMEI-MSISDNs (no NULLs) |

|  |
| --- |
| "conditions\_breakdown": {  "duplicate\_mk1": {  "num\_imei\_gross\_adds": 16,  "num\_imei\_imsis": 56,  "num\_imei\_msisdns": 56,  "num\_imeis": 16,  "num\_triplets": 56  },  "gsma\_not\_found": {  "num\_imei\_gross\_adds": 31,  "num\_imei\_imsis": 39,  "num\_imei\_msisdns": 39,  "num\_imeis": 31,  "num\_triplets": 39  },  "local\_stolen": {  "num\_imei\_gross\_adds": 10,  "num\_imei\_imsis": 13,  "num\_imei\_msisdns": 13,  "num\_imeis": 10,  "num\_triplets": 13  },  "malformed\_imei": {  "num\_imei\_gross\_adds": 0,  "num\_imei\_imsis": 0,  "num\_imei\_msisdns": 0,  "num\_imeis": 0,  "num\_triplets": 0  }  }, |

Report name

|  |
| --- |
| "country\_name": "Country1",  "creation\_date": "2017-09-27",  "end\_date": "2017-08-31",  "has\_compliance\_data": true,  "has\_data": true, |

Historic blacklist adds

Historic statsfor the last five months of blacklist adds. This is used to generate drawing trends.

|  |
| --- |
| "historic\_blacklist\_adds": [  {  "imeis": 0,  "non\_paired\_imsis": 0  },  {  "imeis": 0,  "non\_paired\_imsis": 0  },  { |

Historic compliance breakdown

|  |
| --- |
| "historic\_compliance\_breakdown": [  {  "num\_compliant\_imei\_imsis": 572,  "num\_compliant\_imei\_msisdns": 572,  "num\_compliant\_imeis": 530,  "num\_compliant\_triplets": 572,  "num\_noncompliant\_imei\_imsis": 146,  "num\_noncompliant\_imei\_imsis\_blocking": 52,  "num\_noncompliant\_imei\_imsis\_info\_only": 47,  "num\_noncompliant\_imei\_msisdns": 52,  "num\_noncompliant\_imei\_msisdns\_blocking": 52,  "num\_noncompliant\_imei\_msisdns\_info\_only": 47,  "num\_noncompliant\_imeis": 54,  "num\_noncompliant\_imeis\_blocking": 41,  "num\_noncompliant\_imeis\_info\_only": 13,  "num\_noncompliant\_triplets": 99,  "num\_noncompliant\_triplets\_blocking": 52,  "num\_noncompliant\_triplets\_info\_only": 47  }  ], |

Historic conditions breakdown

|  |
| --- |
| "historic\_conditions\_breakdown": {  "duplicate\_mk1": [  {  "num\_imei\_gross\_adds": 0,  "num\_imei\_imsis": 0,  "num\_imei\_msisdns": 0,  "num\_imeis": 0,  "num\_triplets": 0  },  …  {  "num\_imei\_gross\_adds": 16,  "num\_imei\_imsis": 56,  "num\_imei\_msisdns": 56,  "num\_imeis": 16,  "num\_triplets": 56  }  ],  "gsma\_not\_found": [  {  "num\_imei\_gross\_adds": 0,  "num\_imei\_imsis": 0,  "num\_imei\_msisdns": 0,  "num\_imeis": 0,  "num\_triplets": 0  },  …  {  "num\_imei\_gross\_adds": 31,  "num\_imei\_imsis": 39,  "num\_imei\_msisdns": 39,  "num\_imeis": 31,  "num\_triplets": 39  }  ],  "local\_stolen": [  {  "num\_imei\_gross\_adds": 0,  "num\_imei\_imsis": 0,  "num\_imei\_msisdns": 0,  "num\_imeis": 0,  "num\_triplets": 0  },  …  {  "num\_imei\_gross\_adds": 10,  "num\_imei\_imsis": 13,  "num\_imei\_msisdns": 13,  "num\_imeis": 10,  "num\_triplets": 13  }  ],  "malformed\_imei": [  {  "num\_imei\_gross\_adds": 0,  "num\_imei\_imsis": 0,  "num\_imei\_msisdns": 0,  "num\_imeis": 0,  "num\_triplets": 0  },  …  "num\_imei\_gross\_adds": 0,  "num\_imei\_imsis": 0,  "num\_imei\_msisdns": 0,  "num\_imeis": 0,  "num\_triplets": 0  }  ]  }, |

Historic IMEI, IMSI, MSISDN and triplet counts

Table 4-32 Historic IMEI, IMSI, MSISDN and triplet counts

|  |  |
| --- | --- |
| Field names | Description |
| historic\_imei\_counts | Contains list of total\_imeis\_seen results for previous months for drawing trends |
| historic\_imsi\_counts | Contains list of total\_imsis\_seen results for previous months for drawing trends |
| historic\_msisdn\_counts | Contains list of total\_msisdns\_seen results for previous months for drawing trends |
| historic\_triplet\_counts | Contains list of total\_triplets\_seen results for previous months for drawing trends |

|  |
| --- |
| "historic\_imei\_counts": [  0,  0,  0,  0,  0,  584  ],  "historic\_imsi\_counts": [  0,  0,  0,  0,  0,  671  ],  "historic\_msisdn\_counts": [  0,  0,  0,  0,  0,  671  ],  "historic\_triplet\_counts": [  0,  0,  0,  0,  0,  671  ], |

IMEI/IMSI and IMSI/IMEI overloading

Table 4-33 IMEI/IMSI and IMSI/IMEI overloading

|  |  |
| --- | --- |
| Field names | Description |
| imsi\_imei\_overloading | Number of IMSIs seen with 1 IMEI, 2 IMEIs, 3 IMEIs, etc. |
| imei\_imsi\_overloading | Number of IMEIs seen with 1 IMSI, 2 IMSIs, 3 IMSIs, etc. (duplication) |

|  |
| --- |
| "imei\_imsi\_overloading": [  {  "num\_imeis": 521,  "seen\_with\_imsis": 1  },  {  "num\_imeis": 47,  "seen\_with\_imsis": 2  },  {  "num\_imeis": 13,  "seen\_with\_imsis": 3  },  {  "num\_imeis": 1,  "seen\_with\_imsis": 4  },  {  "num\_imeis": 1,  "seen\_with\_imsis": 5  },  {  "num\_imeis": 1,  "seen\_with\_imsis": 8  }  ], |

Daily counts for IMEIs, IMSIs and MSISDNs

Table 4-34 Daily counts for IMEIs, IMSIs and MSISDNs

|  |  |
| --- | --- |
| Field names | Description |
| imeis\_per\_day | Distinct IMEIs seen per day (no NULLs) |
| imsis\_per\_day | Distinct IMSIs seen per day (no NULLs) |
| msisdns\_per\_day | Distinct MSISDNs seen per day (no NULLs) |
| recs\_per\_day | Distinct triplets seen per day (no NULLs) |

|  |
| --- |
| "imeis\_per\_day": [  {  "count": 281,  "date": "2017-08-01"  },  …  {  "count": 266,  "date": "2017-08-31"  }  "imsi\_imei\_overloading": [  {  "num\_imsis": 671,  "seen\_with\_imeis": 1  }  …  "imsis\_per\_day": [  {  "count": 299,  "date": "2017-08-01"  },  …  {  "count": 281,  "date": "2017-08-31"  },  "msisdns\_per\_day": [  {  "count": 299,  "date": "2017-08-01"  },  …  {  "count": 250,  "date": "2017-08-31"  },  "recs\_per\_day": [  {  "count": 299,  "date": "2017-08-01"  },  …  {  "count": 250,  "date": "2017-08-31"  }, |

Report schema version and DIRBS core software version

|  |
| --- |
| "report\_schema\_version": 2,  "software\_version": "5.2.0",  "start\_date": "2017-08-01", |

Top models

Table 4-35 Top models

|  |  |
| --- | --- |
| Field names | Description |
| top\_models\_gross\_adds | List of top 10 models by ordered by IMEI gross adds. Each list item contains manufacturer, model, gross add IMEI count and tech generation (2G, etc.) |
| top\_models\_imei | List of top 10 models by ordered by raw IMEI count. Each list item contains manufacturer, model, gross add IMEI count and tech generation (2G, etc.) |
| top\_models\_gross\_adds\_count | Sum of IMEI gross add counts for top\_models\_gross\_adds. Used for percentage calculations. Technically redundant |
| top\_models\_imei\_count | Sum of IMEI gross add counts for top\_models\_imei. Used for percentage calculations. Technically redundant |

|  |
| --- |
| "top\_models\_gross\_adds": [  {  "count": 12,  "manufacturer": "Manufacturer-33",  "model": "Model-345",  "tech\_generations": "2G/3G"  },  {  "count": 8,  "manufacturer": "Manufacturer-60",  "model": "Model-74",  "tech\_generations": "2G/3G"  },  {  "count": 7,  "manufacturer": "Manufacturer-18",  "model": "Model-109",  "tech\_generations": "2G/3G"  },  …  "top\_models\_gross\_adds\_count": 73,  "top\_models\_imei": [  {  "count": 12,  "manufacturer": "Manufacturer-33",  "model": "Model-345",  "tech\_generations": "2G/3G"  },  {  "count": 8,  "manufacturer": "Manufacturer-60",  "model": "Model-74",  "tech\_generations": "2G/3G"  },  {  "count": 7,  "manufacturer": "Manufacturer-18",  "model": "Model-109",  "tech\_generations": "2G/3G"  },  …  "top\_models\_imei\_count": 73, |

Monthly counts

Table 4-36 Monthly counts

|  |  |
| --- | --- |
| Field names | Description |
| total\_imeis\_seen | Number of distinct IMEIs seen (no NULLs) |
| total\_imsis\_seen | Number of distinct IMSIs seen (no NULLs) |
| total\_msisdns\_seen | Number of distinct MSISDNs seen (no NULLs) |
| total\_imei\_imsis\_seen | Number of distinct IMEI-IMSI pairs (no NULLs) |
| total\_imei\_msisdns\_seen | Number of distinct IMEI-MSISDN pairs (no NULLs) |
| total\_imsi\_msisdns\_seen | Number of distinct IMSI-MSISDN pairs (no NULLs) |
| total\_gross\_adds | Number of IMEI gross adds |
| total\_records\_seen | Blind COUNT(\*) of all rows of data |
| total\_triplets\_seen | Number of distinct IMEI-IMSI-MSISDN triplets (no NULLS) |
| total\_null\_imei\_records | Rows of data containing a NULL IMEI |
| total\_null\_imsi\_records | Rows of data containing a NULL IMSI |
| total\_null\_msisdn\_records | Rows of data containing a NULL MSISDN |
| total\_invalid\_imei\_imsis | Distinct IMEI-IMSI pairs where IMEI or IMSI is NULL |
| total\_invalid\_imei\_msisdns | Distinct IMEI-IMSI pairs where IMEI or MSISDN is NULL |
| total\_invalid\_triplets | Distinct IMEI-IMSI-MSISDN triplets where any is NULL |
| total\_whitespace\_imsi\_records | Will always be zero in recent release (REMOVE) |
| total\_whitespace\_msisdn\_records | Will always be zero in recent release (REMOVE) |
| historic\_blacklist\_adds | Historic stats for the last five months for above for drawing trends |

|  |
| --- |
| "total\_blacklist\_violations": 0,  "total\_gross\_adds": 584,  "total\_imei\_imsis\_seen": 671,  "total\_imei\_msisdns\_seen": 671,  "total\_imeis\_seen": 584,  "total\_imsi\_msisdns\_seen": 671,  "total\_imsis\_seen": 671,  "total\_invalid\_imei\_imsis": 0,  "total\_invalid\_imei\_msisdns": 0,  "total\_invalid\_triplets": 0,  "total\_msisdns\_seen": 671,  "total\_null\_imsis": 0,  "total\_null\_msisdns": 0,  "total\_records\_seen": 671,  "total\_triplets\_seen": 671,  "total\_whitespace\_imsis": 0,  "total\_whitespace\_msisdns": 0 |

#### CSV

Country1\_8\_2017.csv

Country1\_8\_2017.csv shows conditions met per TAC and the additional data in the header.

|  |
| --- |
| TAC,gsma\_not\_found,local\_stolen,malformed\_imei,duplicate\_mk1,IMEI count,IMEI gross adds count,IMEI-IMSI count,IMEI-MSISDN count,Subscriber triplet count,Compliance Level  35929705,False,False,False,False,2,2,2,2,2,2  35347306,False,False,False,False,1,1,1,1,1,2  35544905,False,False,False,False,1,1,1,1,1,2  35295707,False,False,False,False,1,1,1,1,1,2  35305902,False,False,False,False,1,1,1,1,1,2  35211906,False,False,False,False,1,1,1,1,1,2  35627206,False,False,False,False,1,1,1,1,1,2  35730805,False,False,False,False,1,1,1,1,1,2  … |

Country1\_8\_2017\_condition\_counts.csv

Country1\_8\_2017\_condition\_counts.csv shows all configured conditions and additional data in the header.

|  |
| --- |
| gsma\_not\_found,local\_stolen,malformed\_imei,duplicate\_mk1,IMEI count,IMEI gross adds count,IMEI-IMSI count,IMEI-MSISDN count,Subscriber triplet count,Compliance Level  False,False,False,False,530,530,572,572,572,2  False,False,False,True,13,13,47,47,47,1  False,True,False,False,9,9,10,10,10,0  False,True,False,True,1,1,3,3,3,0  True,False,False,False,29,29,33,33,33,0  True,False,False,True,2,2,6,6,6,0 |

### Operator reports

HTML and JSON operator reports are identical.

#### CSV

Country1\_operator1\_8\_2017.csv shows conditions met per TAC and additional data in the header.

Country1\_operator1\_8\_2017.csv

TAC,gsma\_not\_found,local\_stolen,malformed\_imei,duplicate\_mk1,IMEI count,IMEI gross adds count,IMEI-IMSI count,IMEI-MSISDN count,Subscriber triplet count,Compliance Level

99000435,False,False,False,False,1,1,1,1,1,2

01140800,False,False,False,False,1,1,1,1,1,2

86809701,False,False,False,True,1,1,2,2,2,1

86809701,False,False,False,False,1,1,1,1,1,2

Country1\_operator1\_8\_2017\_condition\_counts.csv

Country1\_operator1\_8\_2017\_condition\_counts.csv shows all configured conditions and additional data in the header.

gsma\_not\_found,local\_stolen,malformed\_imei,duplicate\_mk1,IMEI count,IMEI gross adds count,IMEI-IMSI count,IMEI-MSISDN count,Subscriber triplet count,Compliance Level

False,False,False,False,488,488,488,488,488,2

True,False,False,False,25,25,25,25,25,0

False,True,False,True,2,2,5,5,5,0

False,False,False,True,55,55,131,131,131,1

## Condition IMEI overlaps reports

Condition IMEI overlaps reports generate per-condition reports showing matched IMEIs seen on more than one MNO network.

**Country1\_8\_2017\_condition\_imei\_overlap\_duplicate\_mk1.csv**

IMEI,Operators

01170100000001,operator1|operator2

01206400000001,operator1|operator2

01219000000001,operator1|operator2

01223745000001,operator1|operator2

**Country1\_8\_2017\_condition\_imei\_overlap\_gsma\_not\_found.csv**

IMEI,Operators

01134900000001,operator1|operator2

01223745000001,operator1|operator2

01349800000001,operator1|operator2

01392300000001,operator1|operator2

**Country1\_8\_2017\_condition\_imei\_overlap\_local\_stolen.csv**

IMEI,Operators

01368900000001,operator1|operator2

01388500000001,operator1|operator2

01453800000001,operator1|operator2

35236005000001,operator1|operator2

**Country1\_8\_2017\_condition\_imei\_overlap\_malformed\_imei.csv**

IMEI,Operators

0113AA00000001,operator1|operator2

0122AA45000001,operator1|operator2

0136AA00000001,operator1|operator2

0138AA00000001,operator1|operator2

## GSMA not found reports

Country report with list of IMEIs seen on the network that are not found in the GSMA TAC:

**Country1\_8\_2017\_gsma\_not\_found.csv**

IMEI

01134900000001

01134900000001

01223745000001

01223745000001

## Stolen violations reports

Stolen violations reports generate a per-MNO list of IMEIs seen on the network after they were reported stolen.

**stolen\_violations\_operator1.csv**

imei\_norm,last\_seen,reporting\_date

35236005000001,20170831,20170809

35930705000001,20170831,20170809

35819806000002,20170829,20170809

35570805000002,20170831,20170809

**stolen\_violations\_operator2.csv**

imei\_norm,last\_seen,reporting\_date

35819806000002,20170829,20170809

35793806000001,20170830,20170809

01388500000001,20170829,20170809

01453800000001,20170831,20170809

## Top duplicates reports

Country reports of all IMEIs seen with more than five IMSIs:

Country1\_8\_2017\_duplicates.csv

IMEI,IMSI count

01206400000001,16

35177105000001,10

35840304000001,8

01381500000001,6

# Understanding DIRBS Lists

The dirbs-listgen command creates .zip files containing both the full lists and all the delta lists in CSV format for blacklists, notifications, and exceptions. ZIP files are named as shown below, where both date\_string and operator\_id are variables based on the list generation timestamp and the operator id:

* <date string>\_blacklist.zip (same for every MNO)
* <date\_string>\_notifications\_<operator\_id>.zip
* <date\_string>\_exceptions\_<operator\_id>.zip

Full lists contain all the entries that are on the respective list, while the delta list only contains changes between the list-generation runs

1. Running listgen with no explicit curr-date parameter bases the end of its lookback window off the most recent operator data date rather than the current date.

## Blacklist

The <date\_string> blacklist.zip file will contain the full blacklist and the delta blacklists .CSVs as listed in the sample filenames below. The same blacklists are distributed to all operators,

* 20180217\_000302\_blacklist.csv
* 20180217\_000302\_blacklist\_delta\_-1\_42\_blocked.csv
* 20180217\_000302\_blacklist\_delta\_-1\_42\_changed.csv

20180217\_000302\_blacklist\_delta\_-1\_42\_unblocked.csv

### Full blacklist

The full blacklist file will contain the following information:

* Lists IMEIs that have met a blocking condition and where the current date has exceeded the block date.
* A CSV file containing the complete blacklist is distributed to all MNOs.
* One row per IMEI containing these fields:
  + IMEI
  + Block date for IMEI (earliest block date for all the blocking classification conditions that the IMEI meets)
  + List of reasons for this condition (one reason for each condition resulting in the IMEI being blocked, pipe-separated)

20180217\_000302\_blacklist.csv

imei,block\_date,reasons

31111106045110,20160503,TAC not found in GSMA TAC database

41111101365980,20160503,TAC not found in GSMA TAC database

12640904324427,20171016,IMEI found on local stolen list

12909602872723,20171016,IMEI found on local stolen list

### Delta blacklist

Delta blacklists will also be included in the .zip file.

Each file contains the difference between the results of previous list generation run ID for each event type. The file format is:

<date\_string>\_blacklist\_delta\_-<Previous\_RunID>\_<Current\_RunID>\_<event\_type>.csv

The following are sample delta blacklist file names:

* 20180217\_000302\_blacklist\_delta\_-1\_42\_blocked.csv
* 20180217\_000302\_blacklist\_delta\_-1\_42\_changed.csv
* 20180217\_000302\_blacklist\_delta\_-1\_42\_unblocked.csv

The delta blacklist file contains the same fields as the full list.

20180217\_000302\_blacklist\_delta\_-1\_42\_blocked.csv

imei,block\_date,reasons

31111106045110,20160503,TAC not found in GSMA TAC database

41111101365980,20160503,TAC not found in GSMA TAC database

12640904324427,20171016,IMEI found on local stolen list

12909602872723,20171016,IMEI found on local stolen list

Table 5-37 Blacklist event types

|  |  |
| --- | --- |
| Event | Example scenarios |
| blocked | * Grace period for an IMEI has expired on a previously-met condition * IMEI meets a blocking condition for the first time and grace period was 0 * IMEI was on the golden list, was meeting a blocking condition and then golden list entry was removed |
| unblocked | * IMEI was previously blocked but no longer meets any blocking condition (non-sticky blocking condition) * IMEI was added to the golden list |
| changed | * IMEI previously was blacklisted, but reasons or block date changed * Perhaps stolen and GSMA Not Found and then the TAC got allocated in GSMA so that the new reasons are just stolen |

## Notifications lists

The <date\_string>\_notifications\_<operator\_id>.zip file will be generated for each operator and contain the full notification and the delta notification CSVs, as listed in the following sample file names:

* 20180217\_000302\_notifications\_operator1.csv
* 20180217\_000302\_notifications\_operator1\_delta\_-1\_42\_blacklisted.csv
* 20180217\_000302\_notifications\_operator1\_delta\_-1\_42\_changed.csv
* 20180217\_000302\_notifications\_operator1\_delta\_-1\_42\_new.csv
* 20180217\_000302\_notifications\_operator1\_delta\_-1\_42\_no\_longer\_seen.csv
* 20180217\_000302\_notifications\_operator1\_delta\_-1\_42\_resolved.csv

### Full notification list

The full notification list file will contain the following information:

* Lists IMEIs that have met a blocking condition where the current date is still within the grace period for the condition. Does not include any IMEI already on the blacklist.
* For each IMEI, subscriber triplets are generated based on imported operator data. There is one row in the list for each triplet.
* Determines the home network for each triplet based on IMSI and configured MCC/MNC pairs for each configured operator.
* If a triplet does not match any MCC/MNC pairing for a configured operator (roamers, etc.), we notify all operators whose data they have been seen in.
* Each operator gets a different list containing their subscribers and any fallback triplets seen on their network.
* These fields are included in each row:
  + IMEI
  + IMSI
  + MSISDN (if available in country)
  + Block date for IMEI (earliest block date for all the blocking classification conditions that the IMEI meets)
  + List of reasons for this condition (one reason for each condition met by the IMEI, pipe-separated)
  + Amnesty granted field (set to either True or False)
    - Specifies if IMEI is eligble for amnesty

20180217\_000302\_notifications\_operator1.csv

imei,imsi,msisdn,block\_date,reasons, amnesty\_granted

38674133009747,11101536296900,22300001929746,20161206,IMEI not found on local registration list,blacklisted,false

### Delta notification lists

Delta notification lists will also be included in the .zip file.

Each file will contain the difference between the results of previous list generation run ID for each event type. The file format is:

<date\_string>\_notification\_<operator\_id>\_delta\_-<Previous\_RunID>\_<Current\_RunID>\_<event\_type>.csv

The following are sample delta notification list file names:

* 20180217\_000302\_notifications\_operator1\_delta\_-1\_42\_blacklisted.csv
* 20180217\_000302\_notifications\_operator1\_delta\_-1\_42\_changed.csv
* 20180217\_000302\_notifications\_operator1\_delta\_-1\_42\_new.csv
* 20180217\_000302\_notifications\_operator1\_delta\_-1\_42\_no\_longer\_seen.csv
* 20180217\_000302\_notifications\_operator1\_delta\_-1\_42\_resolved.csv

The delta notification list will contain the same fields as the full list.

20171208\_235247\_notifications\_operator1\_delta\_36\_41.csv

imei,imsi,msisdn,block\_date,reasons,amnesty\_granted

38674133009747,11101536296900,22300001929746,20161206,IMEI not found on local registration list,blacklisted,false

Table 5-38 Notification list event types

|  |  |
| --- | --- |
| Event | Example scenarios |
| new | * IMEI has met a blocking condition for the first time, and there is a non-zero grace period * A new subscriber triplet has been seen with an IMEI meeting a blocking condition and in grace period (changed SIM) * Pairing has been removed for a subscriber using an IMEI meeting a blocking condition and in grace period * IMSI did not have an identifiable home network (via MCC-MNC) and gfipldg was seen for the first time on a network in the lookback window |
| resolved | * IMEI no longer meets a blocking condition and was in grace period previously * Pairing added for a subscriber using an IMEI meeting a blocking condition and in grace period * IMEI added to golden list and was in grace period previously * Triplet no longer seen during lookback window so no longer needs to be notified |
| blacklisted | * IMEI met a blocking condition and was in grace period, but now grace period has expired * IMEI met a new blocking condition that had 0 grace period |
| changed | IMEI is in grace period, but reasons or block date changed (blocking condition added or removed since last list generation) |
| no\_longer\_seen | The triplet was removed from the notifications list, is not paired or blacklisted but the IMEI is still being notified |

## Exceptions lists

The <date\_string>\_exceptions\_<operator\_id>.zip file will be generated for each operator and contain the full exception and the delta exceptions CSVs, as listed in the following sample filenames:

* 20180217\_000302\_exceptions\_operator1.csv
* 20180217\_000302\_exceptions\_operator1\_delta\_-1\_42\_added.csv
* 20180217\_000302\_exceptions\_operator1\_delta\_-1\_42\_removed.csv

### Full exceptions list

A full exceptions list file contains the following information:

* Each operator gets a copy of the pairing list, split into per-operator exception lists based again on their IMSI and the configured MCC/MNC pairs for the configured operators.
* If a pairing's IMSI matches any MCC/MNC pairing for a configured operator (roamers, etc.), the pairing is placed on each operator's exception lists which that IMEI/IMSI combination has been seen.
* These fields are included in each row:
  + IMEI
  + IMSI

20180217\_000302\_exceptions\_operator1.csv

imei,imsi

811111013136464,111038001111111

311111060451100,111035111111111

411111013659808,310035111111111

### Delta exceptions list

Delta exceptions lists will also be included in the .zip file.

Each file contains the difference between the results of previous list generation run ID for each event type. The file format is:

<date\_string>\_exceptions\_<operator\_id>\_delta\_-<Previous\_RunID>\_<Current\_RunID>\_<event\_type>.csv

The following are sample delta notification list file names:

* 20180217\_000302\_exceptions\_operator1\_delta\_-1\_42\_added.csv
* 20180217\_000302\_exceptions\_operator1\_delta\_-1\_42\_removed.csv

20180217\_000302\_exceptions\_operator1\_delta\_1\_42.csv

imei,imsi,change\_type

64220299727231,111041012987198,added

Table 5-39 Exceptions list change types

|  |  |
| --- | --- |
| Event | Example scenarios |
| added | * IMEI-IMSI pair has been added to the pairing list since last run * IMSI did not have an identifiable home network (via MCC-MNC) and was seen for the first time on a network * Config setting restrict\_exceptions\_list\_to\_blacklisted\_imeis is True and the IMEI associated with this pairing just got blacklisted |
| removed | * IMEI-IMSI pair has been removed from the pairing list since last run * IMSI did not have an identifiable home network (via MCC-MNC) and was no longer seen on a network * Config setting restrict\_exceptions\_list\_to\_blacklisted\_imeis is True and the IMEI associated with this pairing was unblocked |

# Frequently Asked Questions

## How does duplicate averaging work?



Figure 7-14 Duplicate averaging

## Reported error during dirbs-classify or dirbs-import

The following error message is the result of a connection timeout between the DIRBS Core and the PostgreSQL server:

2017-11-24 13:03:49,826 - dirbs.exception - ERROR - DIRBS encountered an uncaught software exception

...

psycopg2.DatabaseError: SSL SYSCALL error: Connection timed out

...

psycopg2.OperationalError: SSL SYSCALL error: EOF detected

...

During handling of the above exception, another exception occurred:

...

psycopg2.DatabaseError: SSL SYSCALL error: Connection timed out

Build step 'Execute shell' marked build as failure

Finished: FAILURE

The timeout can be caused and fixed by either or both of the following:

* PostgreSQL server requires tuning. Logs from the server must be analyzed during the tuning process. Check the following on the PostgreSQL server:
  + tcp\_keepalives\_count
  + tcp\_keepalives\_idle
  + tcp\_keepalives\_interval
* Network device configuration, i.e., firewalls:
  + Increase TCP timeout to greater than 1800

## Reported error during dirbs-import

The following error message occurred while importing operator data and is the result of insufficient disk space on the PostgreSQL server:

File "/usr/lib/python3.5/concurrent/futures/\_base.py", line 357, in \_\_get\_result

    raise self.\_exception

File "/usr/lib/python3.5/concurrent/futures/thread.py", line 55, in run

    result = self.fn(\*self.args, \*\*self.kwargs)

File "/home/dirbs/dirbs-venv/lib/python3.5/site-packages/dirbs/importer/abstract\_importer.py", line 338, in \_upload\_file\_to\_staging\_table

    cursor.copy\_expert(sql=self.\_upload\_batch\_to\_staging\_table\_query(), file=f)

psycopg2.OperationalError: could not extend file "base/24702/25222.8": wrote only 4096 of 8192 bytes at block 1162731

**HINT: Check free disk space.**

CONTEXT:  COPY staging\_operator\_import\_5, line 320617

This issue can be resolved by adding additional disk space to your PostgreSQL Server.

## Understanding gsma\_not\_found Reporting Body Index delay configuration

The dirbs.yml file enables the configuration of the Reporting Body Index (RBI) delays to be used when classifying the gsma\_not\_found condition. For syntax and default values, see Appendix B.

Due to delays by the reporting body, there can be a lag between the TAC allocation date and the GSMA TAC DB. New IMEIs may be seen on the network before the TAC is included in the GSMA TAC DB and can be erroneously reported as gsma\_not found and potentially prematurely blocked.

The RBI delay enables the configuration of a delay in days on a per RBI basis (see Figure 6 -15). An IMEI that contains an RBI listed in Appendix B will not be classified as gsma\_not\_found until the RBI delay period has elapsed.



Figure 6-15 RBI delay

The default values configured in the DIRBS Core have been selected based on the analysis of historical data.

1. For all other RBIs that are not listed in Appendix B or configured in the dirbs.yml, the RBI delay is 0. Any found IMEI whose tag does not include a legitimate RBI will be immediately classified as gsma\_not\_found.

## Duplicate and conflicting rows in non-operator imports

This section provides the rationale behind the "conflicting rows" check that has been implemented in DIRBS 7.0.0 and for which there is no option to disable. This includes:

* The difference between duplicate and conflicting rows, and
* Why DIRBS Core cannot safely import files containing conflicting data for the same records.

### Key and metadata columns

Every non-operator import in DIRBS Core (stolen list, pairing list, registration list, and golden list) has columns that fall into two categories:

* **Key columns** uniquely identify the device or pairing. This is often the normalized IMEI for many imports.
* **Metadata columns** contain metadata associated with the device or pairing identified by the key columns. This might be make, model, status, reporting date, etc.

Table 6 -40 summarizes the columns for each type of non-operator import.

Table 6-40 Key and metadata columns

|  |  |  |
| --- | --- | --- |
| Import type | Key columns | Metadata columns in 7.0.0 |
| Stolen list | Normalized IMEI | Reporting date |
| Pairing list | Normalized IMEI, IMSI | None |
| Registration list | Normalized IMEI | None |
| Golden list | Normalized IMEI or Hashed Normalized IMEI | None |

#### Normalized IMEI – imei\_norm

Unfortunately, there is no single definition of an IMEI. There are at least four variants of the same standards-compliant IMEI:

* 14-digit IMEI (no check digit or software version)
* 15-digit IMEI (with Luhn check digit calculated and appended)
* 15-digit IMEI (with 0 transmitted as last digit, as sent over the air)
* 16-digit IMEI (with 2 digit software version appended to 14 digit number)

DIRBS must normalize IMEIs using some well-defined process to ensure that variations listed above map to the same IMEI during classification and list generation.

The algorithm used by DIRBS Core is:

* Trim leading and trailing whitespace to produce TRIMMED\_IMEI. If TRIMMED\_IMEI is an empty string, convert it to NULL.
  + If TRIMMED\_IMEI starts with 14 digit characters (0-9), use those 14 digits as the normalized IMEI.
  + Else, return the uppercase version of TRIMMED\_IMEI as the normalized IMEI.

### Problems

#### Duplicate keys in the file

After normalization of IMEIs, there might be duplicate keys in an input file.

Table 6 -41 and Table 6 -42 show an example file of a stolen list before and after normalization, respectively.

Table 6-41 Stolen list

|  |  |
| --- | --- |
| imei | reporting\_date |
| 123456789012345 | 2017-01-01 |
| 1234567890123463 | 2017-01-01 |

Table 6-42 Stolen list after normalization

|  |  |
| --- | --- |
| imei\_norm | reporting\_date |
| 12345678901234 | 2017-01-01 |
| 12345678901234 | 2017-01-01 |

The imei\_norm column in Table 6 -42 is now duplicated in the file. There may have also been IMEIs in the original file that were duplicated even before normalization.

In general, there are two possible scenarios:

* There are duplicate keys, but all the metadata columns agree , i.e., they all contain the same value. This is the case in Table 6 -41 and Table 6 -42, where the reporting\_date column has the same value for both rows with the duplicate imei\_norm. This is called a duplicate row, because it is an exact duplicate of the other row. **These kinds of duplicates can be safely ignored by the importer.**
* There are duplicate keys, but the metadata columns do not agree, i.e., they contain different vaues. These are called conflicting rows because the rows contain conflicting data for the same key. **These kinds of rows cannot be safely imported and will fail the conflicting rows validation check.**

#### Prior resolution of conflicting rows required

Table 6 -43 provides an example of two conflicting rows. These rows cannot be safely imported because DIRBS Core does not know how to resolve the conflict.

Table 6-43 Conflicting rows

|  |  |
| --- | --- |
| imei\_norm | reporting\_date |
| 12345678901234 | 2017-01-01 |
| 12345678901234 | 2018-03-02 |

If the only metadata column is reporting\_date, it might be possible to simply take the minimum of the reporting\_date and store that, but this may be an incorrect assumption and does not also hold for metadata that are not dates.

Table 6 -44 is a hypothetical example of conflicting rows, containing extra columns proposed in a future DIRBS Core release for the registration list.

Table 6-44 Future DIRBS Core registration list

|  |  |  |  |
| --- | --- | --- | --- |
| imei\_norm | make | model | status |
| 12345678901234 | Samsung | Galaxy | whitelist |
| 12345678901234 | Apple | iPhone | pending\_approval |

If the registration\_list importer were to import the above rows, how would it know which row to import? DIRBS Core does not have enough information to make a decision.

Why not import both rows?

Importing both rows causes the following issues:

* The key columns (imei\_norm) are currently used by the database table as a primary key, meaning that they must be unique. This is only a technical limitation only. The primary key can be changed at the expense of performance and optimal query plans if there is a very good reason to do so.
* If the same IMEI is on the list twice, what make/model should be returned? Is the status of that IMEI that it is registered and therefore whitelisted, or that it is still pending approval from the regulator? Systems like the DVS do not know what answer to return for an IMEI if there is conflicting information for the same IMEI.
* Importing both rows can mess up reporting. If we choose the wrong reporting date, the blacklist violations report might say that there was a stolen list violation that is a false positive.

Delta imports

Delta imports pose an additional, related problem if the same IMEI is in the delta file after normalization with different change types.

For example, after normalizing an IMEI might be 'added' and 'removed' in the same delta file   
(see Table 6 -45).

Table 6-45 Normalized delta file

|  |  |  |
| --- | --- | --- |
| imei | reporting\_date | change\_type |
| 12345678901234 | 2017-01-01 | add |
| 123456789012345 | 2017-01-01 | remove |

This is another conflict that DIRBS Core cannot resolve. This special check for delta files is called the 'multiple\_changes\_check' and only performs if the importer is in delta mode. If a delta file import is failing this check, it is a simlar issue to the conflicting rows check above.

### Options for resolving a conflicting row problem

There are three options for resolving a conflicting row problem:

* Ensure that subsystems that record IMEIs use the same normalization rules as DIRBS Core.
  + This way there are never any duplicates detected by DIRBS Core that could potentially have conflicting data. If the same IMEI is reported stolen twice, for example, the IMEI should already be blocked and a second report should not be lodged. This would be required for new subsystems using replication to directly replicate into DIRBS Core database tables.
* Do not import with any metadata columns.
  + The subsystems in some deployments may not use extra metadata columns. The only use of metadata columns is in the stolen list import, which uses this information for the stolen\_violations report. This might not be a viable option.
* Pre-process data during export for DIRBS.
  + If there is a business need to store duplicate data in the subsystems, these can simply be filtered out during export of data for DIRBS Core using a conflict resolution process based on business rules.

## DIRBS Amnesty feature

The DIRBS Amnesty (grandfathering) feature enables DIRBS operators to grant a grace period for IMEIs that have been seen on the network that match certain blocking and amnesty eligible conditions.

When an IMEI is eligible for amnesty, it will not be blacklisted or blocked from the network for the duration of the amnesty period and enables IMEIs to continue working on the operator network. Once the amnesty period expires, IMEIs are classified based on the rules and conditions configured in .dirbs.yml.

The DIRBS system maintains the list of amnesty eligible IMEIs internally in the DIRBS database.

DIRBS Amnesty contains three phases:

* **Amnesty evaluation:** A pre-configured period of time that determines IMEI amnesty eligibility based on conditions configured in the .yml file. Amnesty eligible IMEIs are determined by running the dirbs-classify command.
* **Amnesty:** A pre-configured period of time that defines when amnesty is in effect. Amnesty eligible IMEIs will be on the notification list with a block date that is the amnesty end date (see Section 1.2.3). IMEIs are no longer evaluated for amnesty eligibility.
* **Post-amnesty:** Amnesty period has expired. The system no longer checks whether an IMEI is eligible for amnesty. Normal system classification and notifications are in effect.

### Enabling and configuring amnesty in .dirbs.yml

To enable the amnesty feature and configure amnesty evaluation and period end dates:

# Definition of settings to be used for amnesty feature. Amnesty feature enables native grandfathering support within

# DIRBS Core. A list of whitelisted IMEIs is managed within Core transparent to EIRs during the amnesty period.

# The amnesty list is mutable during the amnesty evaluation period, immutable during the amnesty period.

# During the amnesty evaluation period, the amnesty\_list table is overwritten each time dirbs-classify is run.

amnesty:

# Boolean value to indicate whether to enable this feature or not.

amnesty\_enabled: True

# End date of the amnesty evaluation period & start of the amnesty period.

evaluation\_period\_end\_date: 20180131

# End of amnesty period. Must be greater than the evaluation period end date.

amnesty\_period\_end\_date: 20180417

Conditions that determine amnesty eligibility are configured in .dirbs.yml:

- label: simple\_dimension

dimensions:

- module: gsma\_not\_found

grace\_period\_days: 0

blocking: True

amnesty\_eligible: True

reason: Violated simple dimension

1. The blocking parameter must be set to True for the amnesty eligible parameter to take effect.

### Eligibility and notifications

During the amnesty evaluation period, IMEIs will be classified using the configured amnesty eligible conditions. When running dirbs-classify and the IMEI meets the amnesty eligible condition, it will be classified as amnesty eligible and stored as such in the DIRBS classification state table.

Running dirbs-listgen during this period will not generate notifications for amnesty eligible IMEIs.

Running dibs-listgen during the amnesty\_period will generate notifications for amnesty eligible IMEIs in the format described in Section 5.2.

#### Modifications

Evaluation period

The evaluation period can be extended or reduced by modifying the evaluation\_period\_end\_date and re-running dirbs-classify.

Amnesty period

The amnesty\_period can be extended or reduced by modifying the amnesty\_period\_end\_date and re-running dirbs-classify.

#### Disabling

Amnesty can be disabled prior to the amnesty evaluation end date by toggling the amnesty\_enabled parameter in the .dirbs.yml to False.

Changing this value to False after the evaluation period or during the amnesty period does not disable the feature as amnesty eligible IMEIs have already been classified and stored in the database.

During the amnesty period, the feature can be disabled by accelerating the amnesty period end date to the current date and re-running dirbs-classify. IMEIs will then be classified based on the configured conditions and notified/blocked accordingly.

### Stolen, paired, and golden IMEI interaction

Amnesty eligible IMEIs added to the golden list will not be added to the notification list.

Amnesty eligible IMEIs added to the pairing list will exclude a subset of triplets from the notification list.

Example

Triplets seen on the network:

* IMEI-1 - IMSI-1 - MSISDN-1
* IMEI-1 - IMSI-2 - MSISDN-2
* IMEI-1 - IMSI-3 - MSISDN-3

If IMEI-1 is amnesty eligible and also on the golden list, then none of the above triplets are on the notification list.

If IMEI-1 is amnesty eligible and IMEI-1-IMSI-1 is on the pairing list, then IMEI-1-IMSI-2-MSISDN-2 and IMEI-1-IMSI-3-MSISDN-3 triplets will be on the notification list.

The behavior for amnesty eligible IMEIs that are added to the stolen list must be specifically configured in the conditions section of the .dirbs.yml.

It is recommended to treat stolen (not duplicated) IMEIs, and stolen and duplicated IMEIs differently:

* For an amnesty eligible IMEI that was stolen and not duplicated, the IMEI should be blocked and blacklisted. This should be configured as a compound dimension in the conditions section.

- label: compound\_dimension1

dimensions:

- module: stolen\_list

- module: duplicate\_threshold

parameters:

threshold: 2

period\_days: 30

invert: True

grace\_period\_days: 0

blocking: True

reason: Violated compound dimension stolen & not duplicate

max\_allowed\_matching\_ratio: 0.1

* For an amnesty eligible IMEI that was stolen and duplicated, it is recommended to not block the IMEI so as not to impact other IMEIs that are amnesty eligible.

It is recommended that the Operator terminate the subscription associated with the stolen device.

- label: compound\_dimension2

dimensions:

- module: stolen\_list

- module: duplicate\_threshold

parameters:

threshold: 2

period\_days: 30

invert: False

grace\_period\_days: 0

blocking: False

amnesty\_eligible: Trus

reason: Violated compound dimension of stolen & duplicate

max\_allowed\_matching\_ratio: 0.1

#### Post-amnesty

After the amnesty period, DIRBS Core will classify IMEIs based on conditions configured in .dirbs.yml. It is recommended that the conditions are reviewed and reconfigured as required.

DIRBS Configuration File Sample: YML

Sample annotated config for DIRBS Core configuration

# (C) 2016-2017 Qualcomm Technologies, Inc. All rights reserved.

#

# PostgreSQL settings used to build connection string

postgresql:

# Database name (an empty database on the first run). Overridden by environment

# variable DIRBS\_DB\_DATABASE if set.

database: dirbs

# Host that the PostgreSQL server runs on. Overridden by environment

# variable DIRBS\_DB\_HOST if set.

host: localhost

# PostgreSQL port if not running on standard port of 5432. Overridden by environment

# variable DIRBS\_DB\_PORT if set.

port: 5432

# Database role/user that DIRBS will connect to PostgreSQL as. Overridden by environment

# variable DIRBS\_DB\_USER if set.

user: dirbs

# Password used to connect to the database.

#

# There are a number of ways to set the password, with each option having pros and cons

# dependent on the level of security required vs. ability to automate

# - Firstly, the password can be defined here in clear text. This file would

# then have its permissions set appropriately to restrict access to non-admin

# users

# - If the setting is not defined in this config file, the user's .pgpass file

# will be read from their home directory. Note that this file will only be

# read if its permissions are set appropriately (must only be readable by the user)

# - If the DIRBS\_DB\_PASSWORD environment variable is set, this value will

# overwrite any value configured in here or in .pgpass

# - Finally, the --db-password-prompt command-line option can be used to prompt

# the user for a password when a command is run.

#

# Uncomment the below line to set the password explicitly in this config file

#

# password: <change me>

# Definitions of regional settings used by DIRBS core for reporting and

# for input validation.

region:

# Name is used for the country level report

name: Country1

# Whether or not MSISDN data is present and should be imported for this region

import\_msisdn\_data: True

# Whether or not RAT data is present and should be imported for this region

import\_rat\_data: True

# country\_codes are used to validate MSISDNs during operator data import

country\_codes:

- "22"

# exempted\_device\_types contains a list of GSMA device types that do not require

# registration in this country. Specifiying a list of device types here will mean

# that the not\_in\_registration\_list classification dimension will ignore IMEIs

# whose TACs correspond to the listed device types. They will also be ignored in

# the IMEI API's realtime registration check. The expected syntax for this is:

#

# exempted\_device\_types:

# - Module

# - Tablet

exempted\_device\_types: []

# operators map operator IDs to a more human-friendly display string for

# reporting purposes

operators:

- id: operator1

name: First Operator

# mcc\_mnc values are used to:

# - validate IMSIs during operator data import

# - work out which operators notifications about an offending subscriber

# should be sent to

# - work out which operators excepted IMEI-IMSI pairings should be sent to

mcc\_mnc\_pairs:

- mcc: "111"

mnc: "01"

- id: operator2

name: Second Operator

mcc\_mnc\_pairs:

- mcc: "111"

mnc: "02"

- id: operator3

name: Third Operator

mcc\_mnc\_pairs:

- mcc: "111"

mnc: "03"

- id: operator4

name: Fourth Operator

mcc\_mnc\_pairs:

- mcc: "111"

mnc: "04"

# Definitions of configuration variables related to pruning of subscriber data

# after a specified retention window

data\_retention:

# The number of months from the start of the current months that DIRBS core

# will retain data about a triplet seen in its DB. After this time, the triplet

# will be erased from the seen\_triplet table. The IMEI will continue to be stored

# after this date as it is needed for continued list generation, etc.

# All references to IMSI and MSISDN will be pruned after this date.

months\_retention: 6

# Definitions of configuration variables used by DIRBS Core in the list generation process.

list\_generation:

# The number of days that DIRBS core will look back through data from current date to determine IMSIs/MSISDNs

# which were associated with the notifiable IMEIs.

lookback\_days: 180

# If true, the exception list will contain only those IMEI-IMSI pairs where the IMEI is on the blacklist.

# By default, all IMEI-IMSI pairs part of the pairing list are output to the exception list.

restrict\_exceptions\_list\_to\_blacklisted\_imeis: false

# If true, generate a check digit for IMEIs during list generation.

# Check digit will only be added to "valid IMEIs"

generate\_check\_digit: false

# If true, output only "valid" IMEIs.

# Valid IMEIs start with 14 digits as they will have 15 digits if the check digit append has been enabled

output\_invalid\_imeis: true

# Definitions of configuration variables used by DIRBS Core in the report generation process.

report\_generation:

# This setting is used by blacklist violations and stolen list violations reports to give the MNO

# some processing time (in days) before an IMEI appearing on the network is considered a violation.

blacklist\_violations\_grace\_period\_days: 2

# Definitions of configuration variables used by DIRBS Core to determine how many workers to use to parallelise

multiprocessing:

# The maximum number of local processing blade workers to use to achieve DIRBS Core tasks. This is particularly

# useful for pre-validation of large operator import jobs where we can run multiple instances of the pre-validator

# in parallel on different parts of the file. The default is to use half of the available CPUs in the

# system will be used.

# max\_local\_cpus: 10

# The maximum number of database connections to use to parallelise DIRBS Core tasks. PostgreSQL 9.6 has support

# for parellelising tasks internally - this setting does not affect parellelisation for a single connection.

# Where PostgreSQL is unable to parallelise a single query by itself, we use this number of workers to issue

# multiple queries at once on different connections. Generally this scales very well - it is safe to set this

# reasonably high. It should probably be set to roughly the number of disks in your RAID array in case there are

# I/O intensive DB operations going on. If using SSD, can be set to a higher value.

max\_db\_connections: 4

# Definition of ratio limits for the various checks on operator data.

operator\_threshold:

# The proportion of the entries in the data that are allowed to have a NULL IMEI

null\_imei\_threshold: 0.05

# The proportion of the entries in the data that are allowed to have a NULL IMSI

null\_imsi\_threshold: 0.05

# The proportion of the entries in the data that are allowed to have a NULL MSISDN (ignored if MSISDN disabled)

null\_msisdn\_threshold: 0.05

# The proportion of the entries in the data that are allowed to have a NULL RAT (ignored if RAT disabled)

null\_rat\_threshold: 0.05

# The proportion of the entries in the data that are allowed to have any column equal to NULL

# This only includes columns enabled in the import (MSISDN and RAT may be excluded)

null\_threshold: 0.05

# The proportion of the non-NULL IMEIs in the data that are allowed to not start with 14 digits

unclean\_imei\_threshold: 0.05

# The proportion of the non-NULL IMSIs in the data that are allowed to not be 14-15 digits

unclean\_imsi\_threshold: 0.05

# The proportion of entries in the data that are allowed to have either a unclean IMEI or an unclean IMSI

unclean\_threshold: 0.05

# The proportion of the non-NULL IMSIs in the data that are allowed to have a MCC that does not match the

# configured region

out\_of\_region\_imsi\_threshold: 0.1

# The proportion of the non-NULL MSISDNs in the data that are allowed to have a CC that does not match the

# configured region. Ignored ir MSISDN disabled

out\_of\_region\_msisdn\_threshold: 0.1

# The combined proportion of entries in the data that are allowed to have either a CC (IMSI) or MCC (MSISDN)

# that does not match the configured region. Ignored if MSISDN if disabled, as this would then be the same as the

# out of region IMSI check.

out\_of\_region\_threshold: 0.1

# The proportion of the entries in the data that are allowed to have an IMSI not starting with one of the MCC-MNC

# prefixes associated with the operator the data is being imported for

non\_home\_network\_threshold: 0.2

# The minimum valid ratio of average daily IMEI count against historical daily IMEI count for a data dump to be

# considered valid.

historic\_imei\_threshold: 0.9

# The minimum valid ratio of average daily IMSI count against historical daily IMSI count for a data dump to be

# considered valid.

historic\_imsi\_threshold: 0.9

# The minimum valid ratio of average daily MSISDN count against historical daily MSISDN count for a data dump to be

# considered valid. Ignored if MSISDN if disabled

historic\_msisdn\_threshold: 0.9

# Each of the following importers specifies 2 historic\_thresholds which can be used to validate new import

# row count against previously imported data for the same importer.

# - import\_size\_variation\_absolute: The most an import can decrease in absolute row count before

# it is rejected as invalid. By setting this variable to -1, this check will be disabled.

# - import\_size\_variation\_percent: The most an import can decrease in percentage row count before

# it is rejected as invalid. 0.75 indicates a new import must be at least 75% of the previous

# import's row count or it will be rejected. Therefore, setting this variables to 0 will disable

# this check.

gsma\_threshold:

import\_size\_variation\_absolute: 100

import\_size\_variation\_percent: 0

pairing\_list\_threshold:

import\_size\_variation\_absolute: 1000

import\_size\_variation\_percent: 0.95

stolen\_list\_threshold:

import\_size\_variation\_absolute: -1

import\_size\_variation\_percent: 0.75

registration\_list\_threshold:

import\_size\_variation\_absolute: -1

import\_size\_variation\_percent: 0.75

golden\_list\_threshold:

import\_size\_variation\_absolute: -1

import\_size\_variation\_percent: 0.75

# Definition of conditions used by the DIRBS system. There are zero or more

# conditions used to drive the classification. A system with zero conditions

# does no classification at all

conditions:

# Each condition specifies the following properties

# label: A name for the condition. This is the id/key for the condition. If

# this is changed, all previous classifications will be reset. Likewise,

# if you change the dimensions but keep the condition label the same,

# existing classifications for that condition will be retained.

# dimensions: A list of dimensions whose intersection forms the IMEI set

# result for the condition. Each of these can take parameters that

# are particular for the dimension being used. Additionally, they all

# accept an 'invert' property, which basically NOTs the result of the

# dimension by taking the all-time observed IMEIs list and subtracting the

# set of IMEIs returned by this dimension

# grace\_period\_days: The integer number of days that an IMEI failing

# this condition will remain on the notification list before moving

# to the black list.

# blocking: A boolean stating whether this condition contributes to

# list generation or is simply informational. Information conditions

# can be used to try out new modules or to tweak parameters.

# reason: A string sent to the operators describing why the IMEI is

# to be blacklisted.

# max\_allowed\_matching\_ratio: The maximum percentage of all-time seen IMEIs

# this condition is allowed to match. This is a safety check implemented

# to catch a missing GSMA TAC DB, registration list, etc.

#

# The following are just sample conditions designed to show the features of

# DIRBS Core and just an example of simple/compound conditions. They are not

# supposed to represent suggestions for real business rules. Please consult

# the release documentation for available dimensions and their parameters.

- label: simple\_dimension

dimensions:

- module: gsma\_not\_found

grace\_period\_days: 30

blocking: true

reason: Violated simple dimension

max\_allowed\_matching\_ratio: 0.1

- label: compound\_dimension

dimensions:

- module: stolen\_list

- module: duplicate\_daily\_avg

parameters:

threshold: 3.1

period\_days: 30

min\_seen\_days: 5

invert: True

grace\_period\_days: 0

blocking: true

reason: Violated compound dimension

max\_allowed\_matching\_ratio: 0.1

# Definition of settings to be used for logging output of DIRBS system.

logging:

# Logging level determines the verbosity of logs. This is also set to 'debug' by the -v CLI option

level: info

# Format string can be configured here

format: '%(asctime)s - %(name)s - %(levelname)s - %(message)s'

# Set this to true if you want to see logging message for StatsD

show\_statsd\_messages: False

# Set this to true if you want to see Werkzeug internal log messages from TAC/IMEI APIs

show\_werkzeug\_messages: False

# Set this to true if you want to see SQL messages from DIRBS (most are debug level)

show\_sql\_messages: False

# If log\_directory is set to a value that is not "null", DIRBS will log to a file as well as to the console. The

# log files will all be generated in the directory specified by this setting. This directory needs to exist and

# be writable

log\_directory: /var/log/dirbs

# Uncomment and set this value if you want to prefix all log files created on this host with a prefix

# to distinguish them from other host

file\_prefix: null

# Set the number of bytes before a logfile is rotated. If this or file\_rotation\_backup\_count is zero, rotation

# is disabled

file\_rotation\_max\_bytes: 100000000

# Sets the numbwe old logs to keep

file\_rotation\_backup\_count: 100

# Definition of settings to be used for forwarding application-defined metrics to a StatsD server

# for aggregation

statsd:

# The hostname for the StatsD server. Overridden by environment

# variable DIRBS\_STATSD\_HOST if set.

#

# Uncomment this and set to a real StatsD hostname to enable collection of metrics

# hostname = statsd.local

#

# The UDP port that the StatsD server is listening on for metrics. Overridden by environment

# variable DIRBS\_STATSD\_PORT

port: 8125

# The prefix to be used for for all metrics collected from this instance. This is useful

# if you have multiple hosts or environments sending data to the same StatsD server and you

# with to differentiate them, Overridden by the environment variable DIRBS\_ENV if set.

#

# Uncomment this and set to a prefix string to enable prefixing of StatsD metrics

# prefix =

# Definition of settings to be used during data cataloging process.

catalog:

# The prospector harvests all files in the path adding them to the data catalog.

# Each prospector specifies the following properties:

# file\_type: Type of files contained within the specified paths.

# It should match the keyword specified during dirbs-import (eg. operator, gsma\_tac etc.)

# paths: Directories and/or files to be harvested. Sub-directories within the listed path are not

# traversed automatically; they should be listed separately if files within them need to be cataloged.

# Multiple paths can be defined for each file type and the path used should be absolute and globally unique.

# schema\_filename: Schema file to be used for data pre-validation (if enabled).

# Multiple prospectors can be defined for the same file\_type if files exist across multiple schema versions.

prospectors:

- file\_type: operator

paths:

- /path/to/operator\_data/directory

schema\_filename: OperatorImportSchema\_v2.csvs

- file\_type: operator

paths:

- /path/to/operator\_data/directory/operator\_data\_file

schema\_filename: OperatorImportSchema.csvs

- file\_type: gsma\_tac

paths:

- /path/to/gsma\_tac/directory

schema\_filename: GSMASchema.csvs

# Set this to true if pre-validation should be performed on the data files.

# Note: Enabling this can slow down the process if there are a lot of uncataloged files.

perform\_prevalidation: False

Sample Conditions: YML

Table B-46 YML sample configuration

|  |  |  |
| --- | --- | --- |
| Condition module name | Sample configuration | Default config |
| duplicate\_daily\_avg | - label: duplicate\_daily\_avg  dimensions:  - module: duplicate\_daily\_avg  parameters:  threshold: 2.0  period\_days: 30  min\_seen\_days: 2  grace\_period\_days: 0  blocking: true  sticky: false  reason: Duplicate daily avg detected | N/A |
| duplicate\_threshold | - label: duplicate\_mk1  dimensions:  - module: duplicate\_threshold  parameters:  threshold: 5  period\_days: 120  grace\_period\_days: 90  blocking: false  reason: Duplicate threshold exceeeded  max\_allowed\_matching\_ratio: 0.1 | N/A |
| gsma\_not\_found  Note: Do not use this condition if there is a live DRS enforcing GSAM not found. | - label: gsma\_not\_found  dimensions:  - module: gsma\_not\_found  parameters:  ***ignore\_rbi\_delays: False***  ***per\_rbi\_delays:***  ***"00": 0***  ***"01": 0***  grace\_period\_days: 0  blocking: true  reason: TAC not found in GSMA TAC database  max\_allowed\_matching\_ratio: 0.1  ***Note: “bold italics” indicates optional parameters*** | ‘RBI’:’Delay(days)’  '00': 32,  '01': 40,  '35': 20,  '86': 19,  '91': 20,  '99': 69  See [GSM Association Non Confidential Official Document TS.06 (DG06) IMEI Allocation and Approval Guidelines](https://www.gsma.com/newsroom/terminal-steering-group/ts-06-imei-allocation-and-approval-process/) for additional information on RBIs |
| malformed\_imei | - label: malformed\_imei  dimensions:  - module: malformed\_imei  grace\_period\_days: 0  blocking: true  reason: Invalid characters detected in IMEI  max\_allowed\_matching\_ratio: 0.1 | N/A |
| not\_on\_registration\_list | - label: not\_on\_registration\_list  dimensions:  - module: not\_on\_registration\_list  grace\_period\_days: 0  blocking: true  reason: IMEI not found on local registration list  max\_allowed\_matching\_ratio: 1.0 | N/A |
| stolen\_list | - label: local\_stolen  dimensions:  - module: stolen\_list  grace\_period\_days: 0  blocking: true  max\_allowed\_matching\_ratio: 0.1  reason: IMEI found on local stolen list | N/A |
| used\_by\_dirbs\_subscriber | - label: used\_by\_local\_non\_dirbs\_roamer  dimensions:  - module: used\_by\_dirbs\_subscriber  parameters:  lookback\_days: 2  grace\_period\_days: 0  reason: IMEI found for local non DIRBS roamer | N/A |
| used\_by\_international\_roamer | - label: used\_by\_local\_non\_dirbs\_roamer  dimensions:  - module: used\_by\_international\_roamer  parameters:  lookback\_days: 2  grace\_period\_days: 0  reason: IMEI found for local non DIRBS roamer | N/A |
| used\_by\_local\_non\_dirbs\_roamer | - label: used\_by\_local\_non\_dirbs\_roamer  dimensions:  - module: used\_by\_local\_non\_dirbs\_roamer  parameters:  lookback\_days: 2  grace\_period\_days: 0  reason: IMEI found for local non DIRBS roamer | N/A |