

Title	Synthetic Mobility Data
Creator / Responsible Partner	TID, IBM
Dataset Identifier	02_SyntheticMobilityDataTID_I-BiDaaS_WP2_D2.1_v0.1.csv.gz
DOI	10.5281/zenodo.4274458 ¹
Dataset Description	This is a synthetic data stream based on real-time data, formed by so-called cell network events. These events are picked up by the antennas that are closer to the mobile phone thus providing an approximate location of the device. Every transaction of a mobile phone generates one of those events. A transaction can be, for instance, placing or receiving a call, sending or receiving an SMS, asking for a specific URL in your mobile phone browser, or sending a text message or a data transaction from/to any mobile phone app. There are also some synchronization events like, for instance, turning your mobile phone on or off, or when switching between location area networks (relatively big geographical areas comprising several cell towers).
Work Package/Deliverable	WP2/D2.1
Source	TID and IBM determined the rule set for the data fabrication using IBM's TDF. Through an iterative analysis, the rules were improved in order to obtain the fabricated dataset used for testing the MVP.
Processing	All the fields containing sensitive information were encrypted using the following encryption formats: <ul style="list-style-type: none"> • Format preserving encryption for categorical fields. Order preserving encryption for numerical fields.
Repository	Zenodo
Language	English
Code list	Each network event consists of 8 fields, which are described in detail below: <ul style="list-style-type: none"> • Date: An 8-digit date with the following format: YYYYMMDD. • EventID: A categorical variable with seven possible values: <ul style="list-style-type: none"> ○ rau: 2% ○ l_service_request: 64% ○ l_tau: 11% ○ l_detach: 5% ○ l_handover: 14% ○ l_attach: 3% ○ l_dedicated_bearer_deactivate: 1% • EventResult: A binary categorical variable (and their distribution): <ul style="list-style-type: none"> ○ Success: 96%

¹ <https://doi.org/10.5281/zenodo.4274458>

	<ul style="list-style-type: none"> ○ Reject: 2% ○ Abort: 2% • UserID: A hash consisting of 44 digits. • SectorID: This variable has the following format of five consecutive digits followed by dash and additional five consecutive digits: XXXXX-XXXXX (X=digit) • Mobile Technology: Categorical variable with six possible values: <ul style="list-style-type: none"> ○ wcdma ○ lte ○ gsm ○ 2G ○ 3G ○ 4G • DeviceModel: many different models • IMEI (IMEI): 15-digit sequences of numbers
Type	Categorical, Numerical
Format	CSV
Expected Size	1.87GB
Keywords	mobility, network events, antennas
Version	V0.1
Date of Repository Submission	17/11/2020
Necessary software	N/A
Rights	TID
Access Information	Open