UJIIndoorLoc Data Set

Abstract.

The UJIIndoorLoc is a Multi-Building Multi-Floor indoor localization database to test Indoor Positioning System that rely on WLAN/WiFi fingerprint.

Source.

Donors/Contact
Joaquín Torres-Sospedra: jtorres +@+ uji.es
Raul Montoliu: montoliu +@+ uji.es
UJI - Institute of New Imaging Technologies, Universitat Jaume I, Avda. Vicente Sos Baynat S/N, 12071, Castellón, Spain

Creators

Joaquín Torres-Sospedra, Raul Montoliu, Adolfo Martínez-Usó, Tomar J. Arnau, Joan P. Avariento, Mauri Benedito-Bordonau, Joaquín Huerta, Yasmina Andreu, óscar Belmonte, Vicent Castelló, Irene Garcia-Martí, Diego Gargallo, Carlos Gonzalez, Nadal Francisco, Josep López, Ruben Martínez, Roberto Mediero, Javier Ortells, Nacho Piqueras, Ianisse Quizán, David Rambla, Luis E. Rodríguez, Eva Salvador Balaguer, Ana Sanchís, Carlos Serra, and Sergi Trilles.

Data set information.

Many real world applications need to know the localization of a user in the world to provide their services. Therefore, automatic user localization has been a hot research topic in the last years. Automatic user localization consists of estimating the position of the user (latitude, longitude and altitude) by using an electronic device, usually a mobile phone. Outdoor localization problem can be solved very accurately thanks to the inclusion of GPS sensors into the mobile devices. However, indoor localization is still an open problem mainly due to the loss of GPS signal in indoor environments. Although, there are some indoor positioning technologies and methodologies, this database is focused on WLAN fingerprint-based ones (also know as WiFi Fingerprinting).

Although there are many papers in the literature trying to solve the indoor localization problem using a WLAN fingerprint-based method, there still exists one important drawback in this field which is the lack of a common database for comparison purposes. So, UJIIndoorLoc database is presented to overcome this gap. We expect that the proposed database will become the reference database to compare different indoor localization methodologies based on WiFi fingerprinting.

The UJIIndoorLoc database covers three buildings of Universitat Jaume I with 4 or more floors and almost 110.000m^2. It can be used for classification, e.g. actual building and floor identification, or regression, e.g. actual longitude and latitude estimation. It was created in 2013 by means of more than 20 different users and 25 Android devices. The database consists of 19937 training/reference records (trainingData.csv file) and 1111 validation/test records (validationData.csv file).

The 529 attributes contain the WiFi fingerprint, the coordinates where it was taken, and other useful information.

Each WiFi fingerprint can be characterized by the detected Wireless Access Points (WAPs) and the corresponding Received Signal Strength Intensity (RSSI). The intensity values are represented as negative integer values ranging -104dBm (extremely poor signal) to 0dbM. The positive value 100 is used to denote when a WAP was not detected. During the database creation, 520 different WAPs were detected. Thus, the WiFi fingerprint is composed by 520 intensity values.

Then the coordinates (latitude, longitude, floor) and Building ID are provided as the attributes to be predicted.

Additional information has been provided.

The particular space (offices, labs, etc.) and the relative position (inside/outside the space) where the capture was taken have been recorded. Outside means that the capture was taken in front of the door of the space.

Information about who (user), how (android device & version) and when (timestamp) WiFi capture was taken is also recorded.

Files.

training.csv   → Data set used for training the model.

validation.csv   → Data set used for validation of the model.

test.csv   → Data set used to test the accuracy of the model.

Attibute information.

Attribute 001 (WAP001): Intensity value for WAP001. Negative integer values from -104 to 0 and +100. Positive value 100 used if WAP001 was not detected.
....
Attribute 520 (WAP520): Intensity value for WAP520. Negative integer values from -104 to 0 and +100. Positive Vvalue 100 used if WAP520 was not detected.
Attribute 521 (Longitude): Longitude. Negative real values from -7695.9387549299299000 to -7299.786516730871000
Attribute 522 (Latitude): Latitude. Positive real values from 4864745.7450159714 to 4865017.3646842018.
Attribute 523 (Floor): Altitude in floors inside the building. Integer values from 0 to 4.

Attribute 524 (BuildingID): ID to identify the building. Measures were taken in three different buildings. Categorical integer values from 0 to 2.

Attribute 525 (SpaceID): Internal ID number to identify the Space (office, corridor, classroom) where the capture was taken. Categorical integer values.

Attribute 526 (RelativePosition): Relative position with respect to the Space (1 - Inside, 2 - Outside in Front of the door). Categorical integer values.

Attribute 527 (UserID): User identifier (see below). Categorical integer values.

Attribute 528 (PhoneID): Android device identifier (see below). Categorical integer values.

Attribute 529 (Timestamp): UNIX Time when the capture was taken. Integer value.

UserID Anonymized user Height (cm)
----
0 USER0000 (Validation User) N/A
1 USER0001 170
2 USER0002 176
3 USER0003 172
4 USER0004 174
5 USER0005 184
6 USER0006 180
7 USER0007 160
8 USER0008 176
9 USER0009 177
10 USER0010 186
11 USER0011 176
12 USER0012 158
13 USER0013 174
14 USER0014 173
15 USER0015 174
16 USER0016 171
17 USER0017 166
18 USER0018 162

PhoneID Android Device Android Ver. UserID
----
0 Celkon A27 4.0.4(6577) 0
1 GT-I8160 2.3.6 8
2 GT-I8160 4.1.2 0
3 GT-I9100 4.0.4 5
4 GT-I9300 4.1.2 0
5 GT-I9505 4.2.2 0
6 GT-S5360 2.3.6 7
7 GT-S6500 2.3.6 14
8 Galaxy Nexus 4.2.2 10
9 Galaxy Nexus 4.3 0
10 HTC Desire HD 2.3.5 18
11 HTC One 4.1.2 15
12 HTC One 4.2.2 0
13 HTC Wildfire S 2.3.5 0,11
14 LT22i 4.0.4 0,1,9,16
15 LT22i 4.1.2 0
16 LT26i 4.0.4 3
17 M1005D 4.0.4 13
18 MT11i 2.3.4 4
19 Nexus 4 4.2.2 6
20 Nexus 4 4.3 0

21 Nexus S 4.1.2 0
22 Orange Monte Carlo 2.3.5 17
23 Transformer TF101 4.0.3 2
24 bq Curie 4.1.1 12

Results file.

The test file provides the intensity value of the signal for 2589 observations. If you want to upload a results file to the platform in order to get a score for your model accuracy, this file must follow these guidelines:

- format: csv (comma separated values)

- columns: 4 named columns (LONGITUDE, LATITUDE, FLOOR and BUILDINGID)

- rows: 1st row for names and 2589 rows for results.

Relevant papers.

Joaquín Torres-Sospedra, Raúl Montoliu, Adolfo Martínez-Usó, Tomar J. Arnau, Joan P. Avariento, Mauri Benedito-Bordonau, Joaquín Huerta. UJIIndoorLoc: A New Multi-building and Multi-floor Database for WLAN Fingerprint-based Indoor Localization Problems. In Proceedings of the Fifth International Conference on Indoor Positioning and Indoor Navigation, 2014.
Available at: [Web Link]

Citation request.

Joaquín Torres-Sospedra, Raúl Montoliu, Adolfo Martínez-Usó, Tomar J. Arnau, Joan P. Avariento, Mauri Benedito-Bordonau, Joaquín Huerta. UJIIndoorLoc: A New Multi-building and Multi-floor Database for WLAN Fingerprint-based Indoor Localization Problems. In Proceedings of the Fifth International Conference on Indoor Positioning and Indoor Navigation, 2014.
Available at: [Web Link]