Solutions for signal mapping campaigns of Wi-Fi networks

Mendoza-Silva, Germán M.; Rodríguez-Pupo, Luis E.; Torres-Sospedra, Joaquín; Huerta-Guijarro, Joaquín

The boom of smart mobile devices with several types of sensors has enabled applications that engage people in collecting information about their surroundings so that they can contribute to citizen science projects. In this paper, we address a set of software solutions that aim to enable the general public to participate in WiFi signal samples collection campaigns. We expect these solutions to be appealing for researchers working in WiFi-based indoor positioning due to the widespread presence of WiFi antennas, the popularity of smartphone able to connect to those antennas, and because it is usually required to create a WiFi fingerprint database, which is a very time-consuming activity. The solutions set addresses three step in the WiFi signal samples database creation process: The campaign planning, the WiFi signal collection and the database construction and sharing. By the end of the process, the research community is provided with sets of geo-located points whose attributes include the signal intensities of the detected WiFi access points. The solutions set that we described in this paper can be extended to include campaigns focused on measuring other physical phenomena by using other sensors found in mobile devices.

KEYWORDS

WiFi mapping, Citizen Science, WiFi Networks, Web, Android Apps, Data Infrastructures.

INTRODUCTION

People are nowadays more aware of the importance of location, as a result in part of the usage of applications related to location-based services (LBS) and the increasing number of objects around us able to communicate and to identify and locate themselves (Internet of Things). In general, and as stated by Kerski (2015), people now have more geography-related skills and a willingness to contribute with data to science. Therefore, they are more capable of being involved in tasks in which map/orientation/location skills are required. One these tasks is the WiFi signal strengths collection for localization purposes.

The indoor positioning has become a hot research topic [2]. Reasons behind it include (i) a huge growth in the location-based services market, (ii) a greater awareness about its potentials and (iii) there is not yet an established, general solution for indoor localization. The widespread presence of WiFi antennas and the pervasive usage of smartphones with WiFi signal strength measurement capabilities have driven the research community interest towards WiFi-based positioning, especially on the WiFi fingerprinting technique [3]. The first step in WiFi fingerprinting is to build a radio map (or training database). Then, this database is used to train an automatic learning algorithm. After the training is completed, the algorithm is able to predict the locations where new signal strengths have been measured [4]. The WiFi fingerprint database construction is a very time-consuming activity, usually involving an ad-hoc software and several people who follow specific instructions. However, these costs can be reduced by using the right tools and methodologies, e.g., citizen science or crowdsourcing approaches.

Citizen science is not a new concept, and several project groups have been harnessing its power along several years [5]. It can provide a valuable help to scientists (mainly) regarding data collection and analysis, which are time and thus economically costly labors. In recent years, citizen science has experienced a boom [6], with two factors decisively influencing these raise: (i) new technologies, particularly improved sensors, better communication options and smart mobile devices [7, 8] and the new roles and areas opened for citizen scientists [6]. With these changes, citizen science has moved to allowing people anywhere and from a wide range of backgrounds to remotely (by online means) contribute to it. According to [9], online citizen science should address a technological pillar, by providing a computer system to facilitate and handle the contributions, and a motivational pillar, by attracting and retaining volunteers. Key aspects of citizen science are the quantity and quality of the contributor results [9], and usually, the contributors are provided with some training or detailed instructions to perform their tasks with high quality.

The crowdsourcing approach has also been used in recent years [10–14]. The crowdsourced WiFi data may have, or not, a location tag. The location tag can be assigned by the user, and as it happens with the citizen science approach, there are factors affecting the quality and quantity. Because the data collection task for WiFi-based localization is relatively simple, the main challenge here is the quality assurance. Despite several research papers have addressed this topic, and no final and dominant result have been obtained. Based on the previous facts, in the work presented in this paper we have mainly focused on a citizen science approach, and in particular on the technological pillar. We provide a set of solutions that facilitates the participation of the general public in projects created by professional scientists. With the term “general public” we refer to people without previous LBS background or training in WiFi signal strength collection. The process of taking WiFi sample is not error-free even when performed by trained operators.

WIFI SIGNAL MEASUREMENT FOR INDOOR LOCALIZATION

People involved in campaigns of WiFi signal collection belong to two main roles: Planners and operators. The planners are generally researchers who work on WiFi-based localization techniques research or development. The operators can be researchers, trained workers or the general public. The operators perform the actual WiFi signal collection by following the instructions given by planners. The main goal of the solutions described in this paper is to provide tools for easing the above mentioned top-down approach.

The researchers working on WiFi localization usually choose specific areas to carry out their experimentation: a lab, a mall, a university building, etc. For these areas, they carefully select the specific points where the sampling is going to be performed: Too many points and the process could require too much effort; while if too few points are mapped, certain methods cannot be properly tested and the localization error could dramatically increase. Apart from its location, other aspects of a recorded fingerprint could be also important. These aspects include, e.g., its order (number) in the whole set of fingerprints, the facing direction of the recorder device (smartphone) and the operator body position regarding the device. The above issues require a planning step that is provided in the first stage of our solutions: **The planning tools**.

Once the instructions given by the researchers are ready, the operators can follow the instructions to perform the WiFi signal samples collection. The collection process is usually made using smartphones. The collected data should reflect the planning made by the main researchers, and should clearly state which user have taken samples, what campaign the data belong to, what time each sample was collected at, etc. These attributes are important because a user can be participating in several collection campaigns and one campaign can include data collected by several users. Once a campaign is complete, the users should provide the main researchers with the results by sharing the campaign’s collected data, along with possible annotations describing errors or suggestions notes. All these facts are taken into account in the second stage of our solutions: **The collection tools**.

When a researcher (campaign planner) has all the data collected by all users participating in a campaign, their data should be revised, i.e., the planner has to review all the data looking for possible inconsistencies, and taking into account the user notes. The planner can refuse specific points or the whole data provided by an operator. Once the data is reviewed, the planner creates a database with it. The planner then decides whether to share it or not with the rest of the research community. The database construction, reviewing and sharing are addressed in the third stage of our solutions: **The cleaning and sharing tools**.

The following sections present a general description of our solutions kit and how it addresses each of the above stages.

THE SOLUTIONS KIT

The kit is a set of collaborative web and android applications. The concepts of user and campaign link these applications. The users, who can be both planners and operators, need to register into a web application. The planners define new campaigns using the planning tools provided in this web application. When defining a campaign, a user chooses what other users can participate as operators. Once the campaign is completely defined, the allowed users carry out the planned collection using the collection tools, which are a mobile Android application. The collection process is guided by the rules established in the campaign definition stage. When the WiFi signal collection is complete, the operator uses the mobile application to upload the gathered data into the web application. Several users can contribute to the same campaign. To ensure the quality of the data, the planner can review the data, delete inconsistencies and make corrections. When the planner considers the data is satisfactory, she/he can make it publicly available.

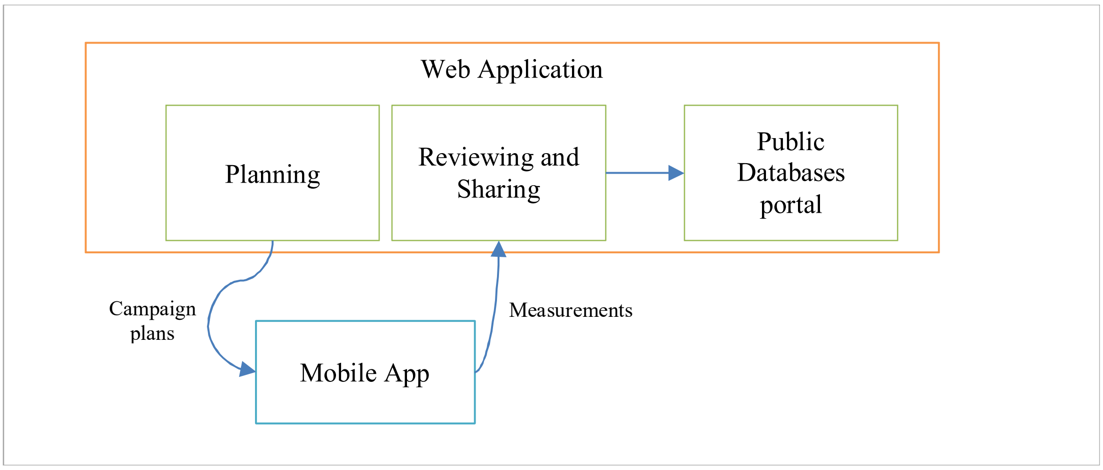


Figure 1. Relation between the tools composing the solutions kit.

THE PLANNING TOOLS

In our solutions, a campaign plan is a set of instructions that the operators need to follow in order to collect quality data. These rules are expressed in a model that is shown in Figure 2. A campaign of WiFi signal collection can be performed by the users for whom the planner has granted permissions. Also, a campaign could only be performed in certain validity period of time.

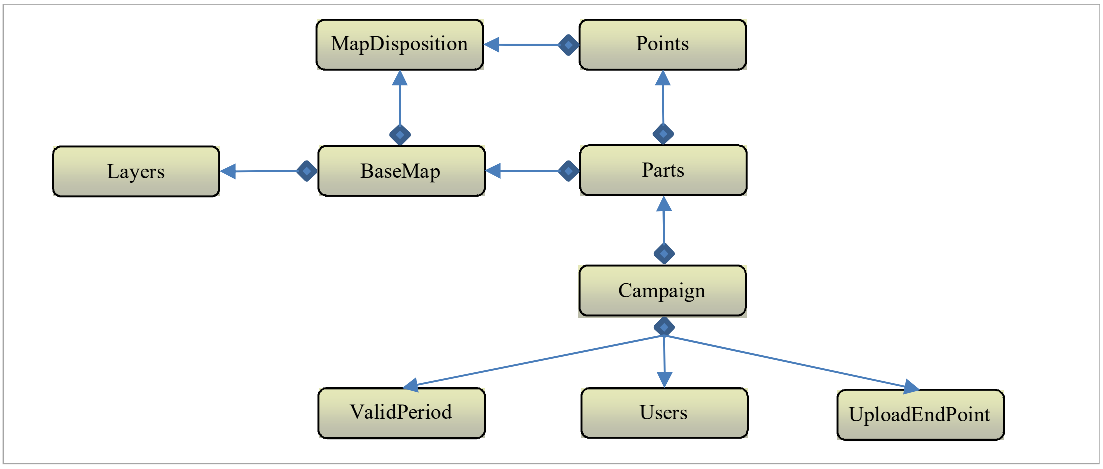


Figure 2. Simplified campaign definition model. Each box represent a model entity that store some data and the arrows represent containment.

Each campaign is composed of several parts, which contain the collection instructions. There are two kinds of instructions: textual instructions and point instructions. The textual instructions are descriptions of what a campaign part is intended for, how to proceed at each moment, and what to pay attention to. The point descriptions establish the locations where the planner want the operators to take WiFi signal measurements. Those locations will be later presented on a map to the operator, along with descriptions about how to collect the WiFi signals, e.g., what direction the operator should face. To improve the collection user experience of the operators, each campaign part can be configured with a base map, which is composed of layers. These layers can be described by an online data resource and a layer descriptor. The data resource can be a WMS service or an Esri runtime geodatabase[1]. With the layer descriptors the planner can achieve a proper layer visualization. It is also possible to provide instructions on how the base map is presented, i.e., orientation and zoom, at each point of the collection process.

```
{
  "campaigns":[
    {...
    "parts":[
      {...
        "samplesPerPoint":6,
        "baseMap":{
          ...
          "resourceUrl":"http://sampleserver//resource.geodatabase",
          "descriptor":{
            "levels":[
              {"index": 0,"levelId": "0","layers": [{"layerId": 1},{"layerId": 2}]},
              {"index": 1,"levelId": "1","layers": [{"layerId": 3},{"layerId": 4}]}
            ]
          }
        },
        "points":[
          {... "x":-7745.5319,"y":4865056.6125,"building":"CD1302BL","number":1,"floor":3},
          {... "x":-7750.4608,"y":4865059.4976,"building":"CD1302BL","number":2,"floor":3},
          {... "x":-7755.4019,"y":4865062.2650,"building":"CD1302BL","number":3,"floor":3},
          {... "x":-7746.7009,"y":4865054.5380,"building":"CD1302BL","number":4,"floor":3}
        ]
      }
    ]
  }]
}
```

Figure 3. Base map and point specification in a campaign definition text. The JSON text has been simplified by omitting some properties (which is indicated by the ellipsis "…" symbol).

The planner creates the model described above as a JSON[2] text and supply it to the web application. Figure 3 shows a simplified version of what the above text can look like. The example shown in this figure highlights the definition of the base map (and the identification of its layers) and the points that form a campaign part. The most laborious part in this text construction is the point location specification. This step can be done with a part of our web application (presented in Figure 4) that allows them to put points on a map, keeping the record of the point order (number). Once the points are located on the map, their locations and orders can be exported to later incorporate them into the campaign's description.

[1] Details can be found on "Creating ArcGIS runtime content", http://desktop.arcgis.com/en/arcmap/10.3/map/working-with-arcmap/creating-arcgis-runtime-content.htm, visited on 12/09/16.
[2] Details can be found on http://www.json.org/, visited on 12/09/16.

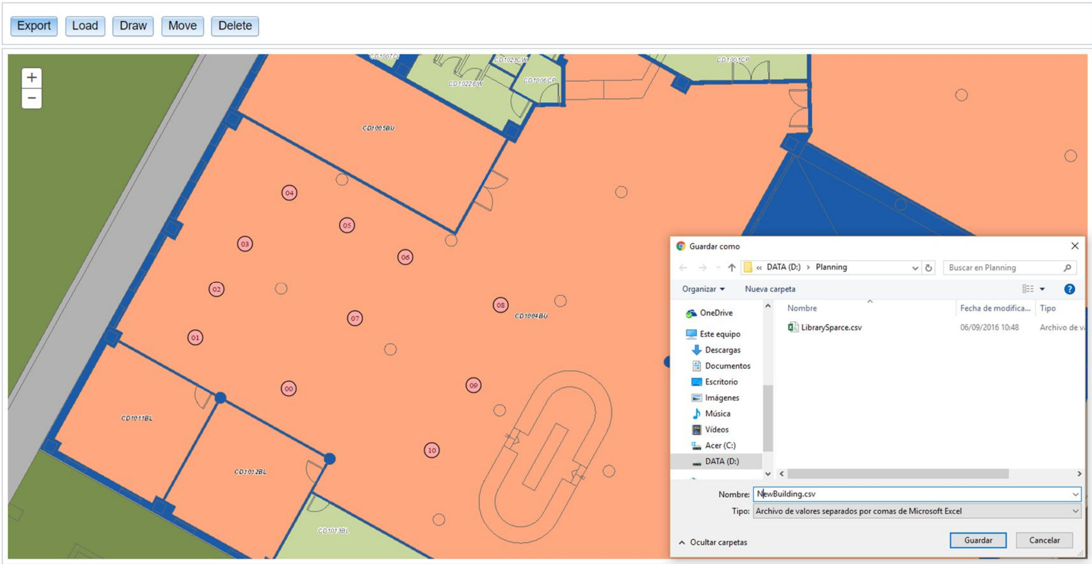


Figure 4. Point plan creation.

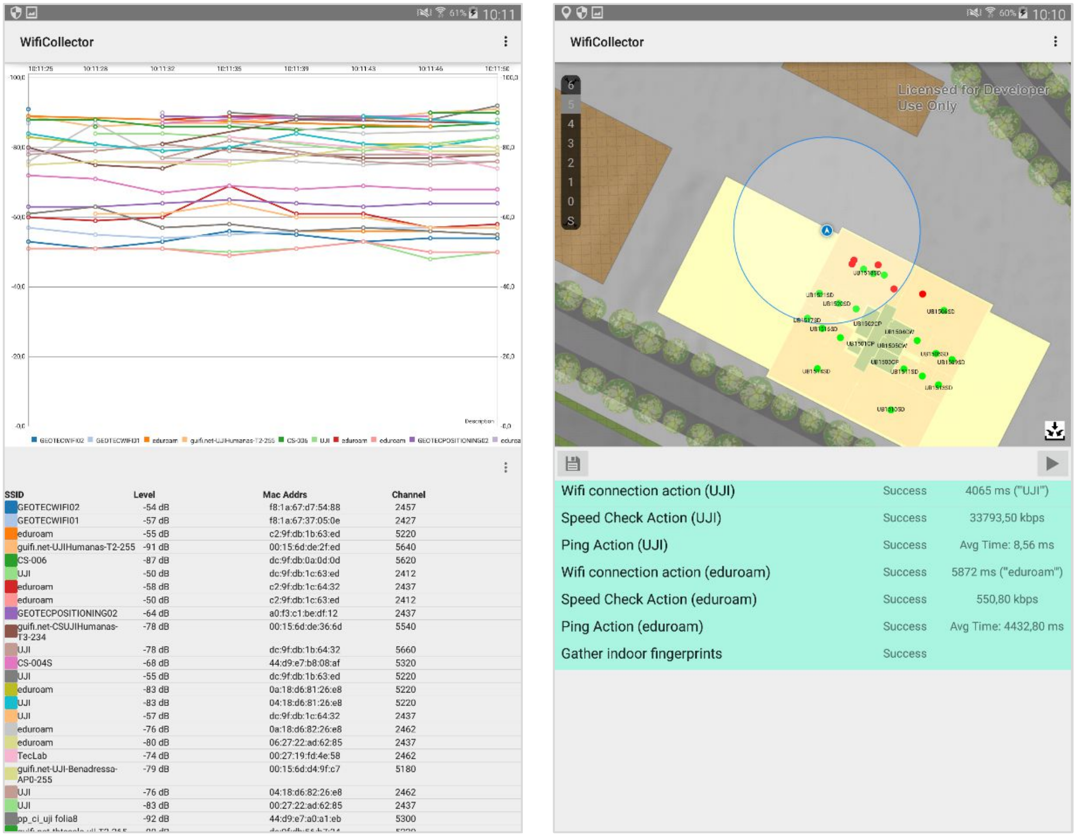
THE COLLECTION TOOLS

The collection tools are grouped into a smartphone (Android) application called WiFiCollector. This application uses offline maps and building information layers, if available, for including spatial and context data to the measurements. Figure 5 presents two screenshots of this application. The application can be used in two ways: A freestyle mode and a campaign-based mode. With the first mode, the operator has total freedom to choose where the samples are been taken. The operator defines the measurements' locations by tapping on a map. With the second mode, the application guides the operator through a set of instructions (points with descriptions) established in the planning stage. In both cases, the application (i) shows in real time the behavior of WiFi signals and (ii) save every recorded measurement in local storage (SQLite database[3]). The fact the data is kept locally make this tool appropriate for scenarios where the WiFi connection is poor or not available.

When working in the campaign-based mode, the WiFiCollector first presents the operator with the option to select the campaign and campaign part she/he wants work on, and later presents the related campaign instructions. Then, on the user commands, points are presented one by one as the collection process advances. For each point, the application sets the proper base map settings (layer, orientation, and zoom) and presents the point description (including its number and floor). When an operator collects all the points in a campaign, the campaign's data can be exported to the web application along with the operator notes or directly as a local database or plain text that can be later

[3] https://www.sqlite.org/, visited on 12/09/16.

shared by, e.g., email.



(a) (b)

Figure 5. Usage of WiFiCollector: (a) presents the real time monitoring and (b) WiFi measurements collection process.

THE CLEANING AND SHARING TOOLS

The tools addressed in this section make the data collected by the operators available to the planners. The data reviewing process is campaign-centric, i.e., for each campaign, the planner can explore the contribution of each user one by one. The planner can delete the data from a specific (i) measurement point, (ii) WiFi antenna or (iii) a user. Once each individual contribution is cleaned, they can be combined into one database. This database has a tabular form in the way described in [2]. Using this form, the databases provides geo-located WiFi signal measurements, i.e., each row corresponds to a set of signal intensity measurements associated with a location. The location data is composed of geographical coordinates, and floor and building when available. The planner decides if the database she/he has created will be publicly available. For public databases, the planner has to supply a name, a description, and keywords. The databases can be later downloaded, explored and visualized using another tool in the web application. This tool (presented in Figure 6) allows to explore several databases and do some basic operations, like filtering, printing, and organization.

Figure 6. Tool for visualization. The measurement points are clustered to avoid over-crowding.

CONCLUSIONS

This paper has presented a set of solutions that eases the process of collecting WiFi signal measurements for indoor localization research. The solutions address the three stages of the process of constructing a WiFi sample dataset: Planning, collection, and assembly. Besides, the solutions provide a platform for sharing the resulting WiFi databases. We expect the solutions described in this paper to be appealing for researchers working in WiFi-based indoor positioning. Reasons behind this expectation include that the WiFi database construction is a very time-consuming process and if the individuals are given with the right instructions, then their contribution can have higher quality and thus be more valuable to researchers.

REFERENCIAS

[1] Kerski, J.J.: Geo-awareness, Geo-enablement, Geotechnologies, Citizen Science, and Storytelling: Geography on the World Stage. Geogr. Compass. 9, 14–26 (2015).

[2] Torres-Sospedra, J., Montoliu, R., Martínez-Usó, A., Avariento, J.P., Arnau, T.J., Benedito-Bordonau, M., Huerta, J.: UJIIndoorLoc: A new multi-building and multi-floor database for WLAN fingerprint-based indoor localization problems. In: Indoor Positioning and Indoor Navigation (IPIN), 2014 International Conference on. pp. 261–270 (2014).

[3] Torres-Sospedra, J., Montoliu, R., Trilles, S., Belmonte, Ó., Huerta, J.: Comprehensive analysis of distance and similarity measures for Wi-Fi fingerprinting indoor positioning systems. Expert Syst. Appl. 42, 9263–9278 (2015).

[4] Marques, N., Meneses, F., Moreira, A.: Combining similarity functions and majority rules for multi-building, multi-floor, WiFi positioning. In: Indoor Positioning and Indoor Navigation (IPIN), 2012 International Conference on. pp. 1–9 (2012).

[5] Silvertown, J.: A new dawn for citizen science. Trends Ecol. Evol. 24, 467–471 (2009).

[6] Bonney, R., Cooper, C., Ballard, H.: The Theory and Practice of Citizen Science: Launching a New Journal. Citiz. Sci. Theory Pract. 1, (2016).

[7] Newman, G., Wiggins, A., Crall, A., Graham, E., Newman, S., Crowston, K.: The future of citizen science: emerging technologies and shifting paradigms. Front. Ecol. Environ. 10, 298–304 (2012).

[8] O'Grady, M.J., Muldoon, C., Carr, D., Wan, J., Kroon, B., O'Hare, G.M.P.: Intelligent Sensing for Citizen Science. Mob. Networks Appl. 21, 375–385 (2016).

[9] Nov, O., Arazy, O., Anderson, D.: Scientists@ Home: what drives the quantity and quality of online citizen science participation? PLoS One. 9, e90375 (2014).

[10] Park, J., Charrow, B., Curtis, D., Battat, J., Minkov, E., Hicks, J., Teller, S., Ledlie, J.: Growing an organic indoor location system. In: Proceedings of the 8th international conference on Mobile systems, applications, and services. pp. 271–284 (2010).

[11] Bolliger, P.: Redpin-adaptive, zero-configuration indoor localization through user collaboration. In: Proceedings of the first ACM international workshop on Mobile entity localization and tracking in GPS-less environments. pp. 55–60 (2008).

[12] Moreira, A., Meneses, F.: Where@UM - Dependable organic radio maps. In: Indoor Positioning and Indoor Navigation (IPIN), 2015 International Conference on. pp. 1–9 (2015).

[13] Pérez-Penichet, C., Moreira, A.: Analyzing the quality of crowd sensed WiFi data. In: Pervasive Computing and Communications Workshops (PERCOM Workshops), 2014 IEEE International Conference on. pp. 272–277 (2014).

[14] Zhuang, Y., Syed, Z., Li, Y., El-Sheimy, N.: Evaluation of Two WiFi Positioning Systems Based on Autonomous Crowdsourcing of Handheld Devices for Indoor Navigation. IEEE Trans. Mob. Comput. 15, 1982–1995 (2016).

AUTHORS

Germán M. Mendoza-Silva
gmendoza@uji
University Jaume I
Institute of New Imaging Technologies

Luis E. Rodríguez-Pupo
luis.rodriguez@uji.es
University Jaume I
Institute of New Imaging Technologies

Joaquín Torres-Sospedra
jtorres@uji.es
University Jaume I
Institute of New Imaging Technologies

Joaquín Huerta-Guijarro
huerta@uji.es
University Jaume I
Institute of New Imaging Technologies