Indoor positioning using ambient radio signals: Data acquisition platform for a long-term study

Andrei Popleteev
University of Luxembourg, Luxembourg
E-mail: andrei.popleteev@uni.lu

***Abstract*—This paper presents an ongoing long-term study exploring indoor positioning systems based on ambient radio signals (such as FM, TV and GSM). We introduce an open-source platform designed to facilitate data acquisition in indoor localization experiments. The platform is currently employed for the creation of a public dataset of georeferenced ambient radio signal samples. The paper discusses the system design as well as the challenges and current lessons of the year-long experiment.**

***Index Terms*—Indoor localization, ambient radio, signals of opportunity, infrastructure-free positioning, fingerprinting, FM radio, TV, GSM, Wi-Fi**

I. Introduction

In contrast to the global satellite-based navigation systems, indoor positioning solutions are highly local and are bounded to buildings with dedicated infrastructure (such as Wi-Fi networks or ultra-wideband transceivers). Ambient radio signals—such as those from commercial radio stations or cellular networks—represent a promising opportunity for globally available indoor positioning; while not specifically designed for localization purposes, these radio signals are transmitted with high power and are intended for indoor reception in populated areas worldwide.

Indoor positioning systems based on ambient radio signals (also known as infrastructure-free systems) have been previously demonstrated by several research groups [1]–[3]. However, there are several open research questions that separate the first successful proof-of-the-concept studies from a realistic global indoor positioning system. In particular, the limited scope of the pioneering studies provides little insight into the actual performance of infrastructure-free localization systems across different seasons of the year, different weather and environment conditions.

This paper presents an approach to a long-term study of indoor positioning based on ambient radio sources. In this context, we introduce an open-source data acquisition platform (DAQ) created to facilitate data collection for indoor localization experiments. As a key part of an ongoing study, the DAQ is employed for building a georeferenced dataset of multi-band radio samples in several indoor testbeds, for the duration of one year. By collecting raw radio-frequency (RF) signal samples from a software-defined radio (SDR) receiver, this approach separates data acquisition from the extraction of location-dependent signal features, thus offering unprecedented flexibility for the evaluation of classic and novel localization methods (potentially including those to be devised).

The following sections present our approach, introduce the DAQ system design and implementation, discuss the challenges of the long-term study, and conclude with a summary of the lessons learned so far.

II. Related work

Indoor localization methods based on ambient radio signals (also known as signals of opportunity) leverage signals broadcast by commercial FM radio and TV stations, as well as those of cellular networks. While not specifically designed for positioning, the wide availability of these signals makes them a promising source for indoor positioning systems.

Due to the difficult signal propagation conditions indoors (with multiple signal reflections and attenuation by walls, ceilings and other objects), timing-based localization approaches are typically not feasible there, as even minor movements of the receiver lead to large positioning errors [4]. Therefore the majority of infrastructure-free indoor positioning studies focus on the fingerprinting approach, which employs an empirical signal distribution map (collected during system calibration phase) and machine learning algorithms to locate the user.

In one of such works, Varshavsky et al. [2] used wide GSM fingerprints with 35 GSM channels; leave-one-out evaluation approach demonstrated a 4 m median accuracy on a single-day dataset. However, the authors did not consider weather dependence and left longer-term experiments for future work.

TV-based localization systems, in turn, tried to leverage the high bandwidth of the TV signals and focused on time-based approaches—with limited success. For instance, Eggert [5] employed time difference of arrival (TDOA) method on analogue TV signals and reported indoor positioning errors of up to 300 m. These results were improved by the Rosum system, which managed to achieve 23 m mean accuracy with digital TV (DTV) synchronization signals [6], [7]. To the best of our knowledge, the only fingerprinting study involving DTV signals was performed only outdoors and reported median accuracy of 130 m [8].

FM radio differs from the other ambient radio broadcasts by its relatively long 3-m wavelength, which theoretically makes FM reception less sensitive to weather and terrain conditions [9]. Following the early outdoor experiments by Krumm et al. [10], Popleteev et al. explored several aspects of FM-based indoor positioning [1], [9], [11]. Further studies demonstrated 89% room recognition rate [12] with ambient FM radio signals, their suitability for context sensing [13] and meter-level indoor localization accuracies [14], [15].

While the initial experiments demonstrated promising results, they were based on small-scale datasets collected over few days and thus provide little insight into the real-world performance of the proposed methods. Before reaching the target indoor environment, ambient radio signals propagate mainly outdoors—a process which is subject to multiple interfering factors such as weather conditions (temperature, humidity, precipitation) and environment dynamics (car traffic, moving tree foliage). The impact of these factors on the localization performance of different radio signals is currently unknown.

Another limitation of the pioneering studies is their focus on a single signal property, namely the received signal strength (RSS). However, Chen et al. [12] suggested that consideration of some FM-specific signal characteristics can improve localization accuracy. Moreover, there are other signal features that could be taken from Wi-Fi based positioning systems; for instance, channel state information (CSI) is known to significantly improve Wi-Fi localization accuracy [16]. Unfortunately, advanced signal features are typically not available from dedicated FM, TV or cellular receivers.

Finally, due to the lack of a public dataset of ambient radio samples, the researchers have to invest considerable time into collecting their own data using specialized hardware, instead of focusing on algorithms and methods. This paper presents our ongoing work on building such a dataset, which would provide a common benchmark for different signal features and localization methods, and thus facilitate further research of indoor positioning based on ambient radio signals.

III. Our approach

The proposed approach is focused on the research challenges identified in the previous section, namely: 1) studying the impact of weather conditions and environment dynamics on localization performance, 2) exploring advanced signal features potentially suitable for positioning, and 3) creating a common reference dataset of georeferenced ambient radio signals.

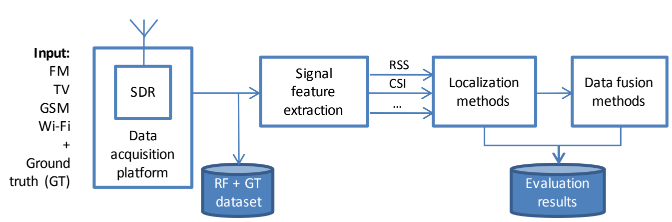

Fig. 1. Overview of the study approach.

An overview of our approach is presented in Fig. 1. In this paper, we focus on the the first phase of the study: the acquisition of a comprehensive long-duration dataset of radio samples at several predefined test locations (see Fig. 1). Raw radio samples are provided by a software-defined receiver which is controlled by the data acquisition platform (DAQ). The resulting dataset contains raw RF samples of multiple radio types (FM, digital TV, GSM-900/1800 and Wi-Fi), along with the ground truth coordinates of the test locations and several metadata parameters.

A. *Software-defined radio*

In this study, ambient radio signals are the variety of signals broadcast by radio transmitters external to the target indoor environment and independent from it. In particular, we focus on FM radio and TV broadcasts, as well as GSM network signals. Due to the considerable differences in frequency bands, channel parameters and modulation types, each radio technology would typically require a dedicated receiver with a predetermined set of measured signal characteristics (such as signal strength).

Software-defined radio (SDR) [17] is free from these limitations. Using a relatively simple RF front end, SDR can tune to a wide range of frequencies and digitize raw radio signals; further signal processing, such as filtering and demodulation, is performed by a computer. Low-level RF samples and the wide frequency range provide unprecedented flexibility for data processing, which can be done offline. In particular, this approach allows post-factum extraction of signal features which were not initially foreseen, whereas the features provided by the traditional specialized receivers would be limited to a predefined set.

B. *Metadata*

In contrast to the conventional indoor localization systems where all the infrastructure is deployed inside or near the target building, propagation of the ambient signals occurs predominantly outdoors. As a result, a number of environmental factors can influence the signal, in particular, weather conditions (rain or wet ground) and more generally the season of the year (foliage or snow) [18], [19]. Movement of people and objects inside the building can also affect propagation and consequently received signal features in a given test point. In order to evaluate the impact of these factors on the localization performance, every data collection session is associated with several meta-parameters, such as timestamp, population dynamics (empty building, working hours, dense crowd), and current weather conditions (automatically acquired from a webservice).

C. *Testbeds*

The duration of the study and the controlled parameters dictate a number of requirements for the selection of the testbed environments. First of all, several testbeds are required to ensure diverse reception conditions. All of them should be accessible in different weather (thus, no rainy terrace) and with different population dynamics. In turn, well-controlled laboratory testbeds would likely be small-scale and might not provide insight into real-world localization performance. Ideally, it should also be possible to deploy ground-truth markers for the test points (which we found impractical in our case). Another limitation is imposed by the data acquisition equipment which might attract unwanted attention and raise security concerns in public places such as airports, shopping malls or train stations. University buildings satisfied all the conditions, as they are available at all times and are spatially distributed across different campuses; moreover, university staff and students are used to seeing research equipment and would not be alarmed by it.

Unfortunately, the considerable effort associated with multi-building multi-floor data collection puts a feasibility limit on the number of test points and time interval between experimental sessions. Also, while signal distribution properties can be influenced by the human presence [20] and in particular by the facing direction of the user [21], consideration of multiple facing directions would proportionally increase the effort required; it is therefore left out of the scope of the study in preference to more frequent sampling. In order to reduce the effect of human body on the measurements, data collection at every test point is consistently performed with the same orientation.

D. Further processing

In the next phase, when the dataset collection is complete, the dataset will be processed to extract location-dependent signal features (such as RSS, CSI or others). As the considered radio technologies vary widely in terms of channel structure, modulation, and content, we plan to leverage GNU Radio [22] toolkit for signal processing.

Once the location-dependent signal features are extracted, these data along with the ground truth coordinates and environmental meta-parameters will be used to evaluate the performance of various localization and data fusion methods. This will provide understanding of real-world performance of indoor positioning systems based on ambient radio signals, as well as their sensitivity to weather conditions and other environmental factors.

IV. Data acquisition platform

Data acquisition platform (DAQ) has been designed to facilitate creation of indoor signal fingerprinting datasets. This section introduces the DAQ as an open-source tool for the indoor positioning community[1].

A. System design

The DAQ consists of three main components (see Fig. 2): an SDR receiver; the DAQ core, which controls the workflow of the system; and the user interface. In order to ensure system modularity and compatibility with different hardware configurations, the user interface is implemented as a dynamic web page which communicates with the DAQ core via HTTP and WebSocket [23] protocols. This also allows separation of the tasks between different devices, for example running the resource-intensive SDR streaming on a heavy powerful laptop and controlling the process from a lightweight tablet.

In order to accommodate SDR receivers with different programming interfaces, SDR control is delegated to a dedicated hardware-specific command-line utility, which can be adapted for different receivers. Once the user requests acquisition of an RF data sample, the DAQ launches the SDR control utility, providing it with a list of tuning parameters, sampling duration, and a directory for storing the acquired data. Overall, the system is rather flexible and SDR receiver can even be replaced with other location-sensing hardware, such as magnetic or light sensors.

[1]Available online: http://popleteev.com/projects/indoors

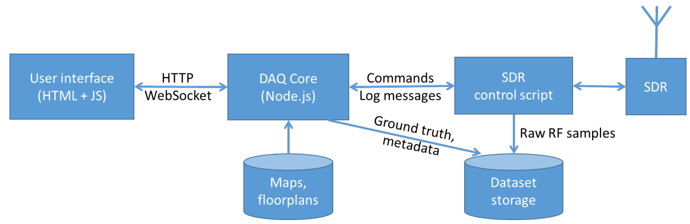

Fig. 2. DAQ system architecture.

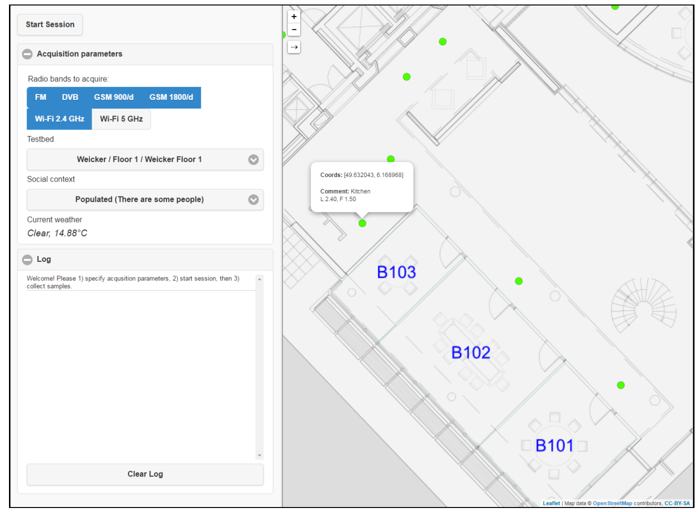

Fig. 3. DAQ user interface.

B. User interface

The user interface of the DAQ features two key parts: a control panel and a floorplan with predefined test points (Fig. 3).

At the start of the experimental session, the operator selects radio bands of interest and the testbed where the session takes place, and specifies the environment's social dynamics — that is, whether the session takes place in an empty building (e.g. offices on weekends), populated building (e.g. offices during work days), or in a dense crowd (e.g. student canteen during lunch). Current weather conditions are acquired from a web service [24] and include such parameters as temperature, humidity, precipitation intensity, pressure, cloud density, wind speed and direction and even ozone levels.

In the course of the experimental session, the operator visits the predefined test locations, selects the current location on the floorplan, and starts RF sample acquisition. Depending on the selected radio bands of interest, sampling takes from 15 s to 2 minutes per point. The user can monitor the sampling progress via log messages from the DAQ core.

Fig. 4. DAQ hardware prototype.

V. LESSONS LEARNED SO FAR

In its current implementation, the DAQ runs on a MacBook Pro laptop with an Intel Core i7 processor, 16 GiB of memory and a fast solid-state storage. Along with a USRP B210 SDR receiver by Ettus Research [25], the laptop is installed on a lightweight mobile cart (see Fig. 4).

1) Mobile cart: The first prototype of the system was considerably more portable, as both SDR and the laptop were concealed in a backpack and controlled remotely from a handheld smartphone. Such a setup was more suitable for experiments in public places such as shopping malls. Unfortunately, due to the lack of proper ventilation inside the backpack, the laptop used to overheat within few minutes. Moreover, it was impossible to ensure consistent antenna height and ground truth position when the measurements were performed by different people.

Installing the DAQ on a mobile cart resolved these issues, and the operator was relieved from carrying the weight of the equipment. The new setup also resulted in an unexpected bonus of increased internal publicity: during the measurements, students and researchers are attracted by the apparatus and are eager to learn about the project.

2) Floorplans: An important part of the DAQ interface is the floorplan of the test environments; this map facilitates definition of test points and make it easier to monitor the progress of the experimental session. While acquiring a floorplan image is not always a trivial task, we have identified and successfully tried the following approaches:

- *Official floorplans* can be requested from a building manager; this works well for research institutions, such as university buildings.
- *Unofficial floorplans* can be derived from the photos of the fire evacuation plans available in buildings like shopping malls, train stations and hotels.
- *Custom floorplans* can be constructed manually, by directly measuring testbed dimensions; this approach is best suited for small environments such as private apartments.

A prepared floorplan image is then aligned with the satellite map of the building surroundings using QGIS georeferencer tool [26] and integrated into the DAQ's interactive map (powered by Leaflet library [27]).

3) Ground truth: Initially, we defined test points that were aligned with easily recognizable features of the environment, such as doors, building pillars, electric sockets, etc. Test locations were marked on the floorplan and supplied with a textual description (for example, "in front of C123 office door"). In our testbeds, the distance between test points varies from about 3 m (office floors) to about 20 m (underground parking); these values are the result of the inevitable trade-off between testbed coverage and feasibility constraints.

On one hand, such a landmark-based ground truth (GT) approach is well-suited for a long-term study, because it does not require any markers which could be removed by the cleaning personnel. On the other hand, however, such GT definitions are open for misinterpretation even by the same person a few weeks later. Moreover, landmark-based locations proved to be ambiguous also over short periods of time [28]. The issue became even more severe in the underground garage with locations defined as "between parking lots 123 and 321"—where parking lots are 3 m wide and 5 m apart.

In order to ensure accurate and consistent ground truth during the long experiment, the mobile cart was equipped with laser rangefinders and the test points were redefined in terms of distances to the surrounding walls. After than, mobile cart's position is always established and verified using the rangefinders. This step has helped to reduce ground truth uncertainty from about 0.3 m [28] to less than 5 cm.

4) Weather dynamics: One of the implicit assumptions of the system design was that the weather conditions remain stable during the experimental session. While most of the times the assumption is true, it does not always hold. In particular, spring and autumn weather in Luxembourg can be rather volatile and easily change from 'clear sky' to 'drizzle' (and back) within one hour. Therefore, for some studies it might be beneficial to record weather information on a per-point basis instead of the current per-session one.

5) Session schedule: Rather unsurprisingly, adhering to a strict measurement schedule proved to be unfeasible. A variety of factors, such as testbed reconstructions, travels, and major life events will interfere with the experimental sessions, even if those are planned by weeks, let alone days.

VI. SUMMARY

The paper presented our approach to studying long-term performance of indoor positioning systems based on ambient radio signals. We also introduced an open-source data acquisition platform designed to facilitate signal data collection in this and similar studies. The project is currently in progress and a preprocessed dataset is scheduled for release in 2017.

VII. Acknowledgments

This work was supported by the National Research Fund of Luxembourg (INDOORS project, grant number C14/IS/8311593).

References

[1] A. Popleteev, “Indoor positioning using FM radio signals,” Ph.D. dissertation, University of Trento, 2011.
[2] A. Varshavsky, E. de Lara, J. Hightower, A. LaMarca, and V. Otsason, “GSM indoor localization,” *Pervasive and Mobile Computing*, vol. 3, no. 6, pp. 698–720, December 2007.
[3] S.-H. Fang, C.-H. Wang, S.-M. Chiou, and P. Lin, “Calibration-free approaches for robust wi-fi positioning against device diversity: A performance comparison,” in *Vehicular Technology Conference (VTC Spring), 2012 IEEE 75th*. IEEE, 2012, pp. 1–5.
[4] R. M. Faragher, “Effects of multipath interference on radio positioning systems,” Ph.D. dissertation, University of Cambridge, 2007.
[5] R. J. Eggert, “Evaluating the navigation potential of the national television system committee broadcast signal,” Ph.D. dissertation, Air Force Institute of Technology, OH, USA, 2004.
[6] M. Rabinowitz and J. J. Spilker Jr, “A new positioning system using television synchronization signals,” *IEEE Transactions on Broadcasting*, vol. 51, no. 1, pp. 51–61, 2005.
[7] D. Rubin and T. Young, “Femtocells: Bringing reliable location and timing indoors,” *InsideGNSS*, vol. Fall, pp. 40–46, 2008.
[8] S. Fang and T. Lin, “Cooperative multi-radio localization in heterogeneous wireless networks,” *IEEE Transactions on Wireless Communications*, vol. 9, no. 5, pp. 1547–1551, 2010.
[9] A. Popleteev, V. Osmani, and O. Mayora, “Investigation of indoor localization with ambient FM radio stations,” in *Proceedings of PerCom-2012*. IEEE, 2012, pp. 171–179, (acceptance rate 11
[10] J. Krumm, G. Cermak, and E. Horvitz, “RightSPOT: A novel sense of location for a smart personal object,” in *Proceedings of the 5th International Conference on Ubiquitous Computing (Ubicomp 2003)*, ser. Lecture Notes in Computer Science, A. K. Dey, A. Schmidt, and J. F. McCarthy, Eds., vol. 2864. Seattle, WA, USA: Springer Berlin / Heidelberg, October 12-15 2003, pp. 36–43.
[11] A. Popleteev, “Device-free indoor localization using ambient radio signals,” in *Adjunct Proceedings of UbiComp-2013*. ACM, 2013, pp. 549–552.
[12] Y. Chen, D. Lymberopoulos, J. Liu, and B. Priyantha, “FM-based indoor localization,” in *Proceedings of MobiSys 2012*, 2012.
[13] S. Shi, S. Sigg, W. Zhao, and Y. Ji, “Monitoring attention using ambient FM radio signals,” *IEEE Pervasive Computing*, vol. 13, no. 1, pp. 30–36, 2014.
[14] V. Moghtadaiee, A. Dempster, and S. Lim, “Indoor localization using FM radio signals: A fingerprinting approach,” in *Proceedings of IPIN-2011*, 2011.
[15] S. Yoon, K. Lee, and I. Rhee, “FM-based indoor localization via automatic fingerprint DB construction and matching,” in *Proceeding of the 11th annual international conference on Mobile systems, applications, and services (MobySys-2013)*. ACM, 2013, pp. 207–220. [Online]. Available: http://dl.acm.org/citation.cfm?id=2464445
[16] Z. Yang, Z. Zhou, and L. Yunhao, “From RSSI to CSI: Indoor localization via channel response,” *ACM Computing Surveys*, vol. 46, no. 2, p. 25, Nov. 2013.
[17] J. Mitola, “The software radio architecture,” *IEEE Communications Magazine*, vol. 33, no. 5, pp. 26–38, 1995.
[18] International Telecommunication Union, “Recommendation ITU-R P.840-4: Attenuation due to clouds and fog,” Online: http://www.itu.int/rec/R-REC-P.840-4-200910-I/en, Oct. 2009, retrieved 2011-01-18.
[19] ——, “Recommendation ITU-R P.833-6: Attenuation in vegetation,” Online: http://www.itu.int/rec/R-REC-P.833-6-200702-I/en, Feb. 2007, retrieved 2011-01-18.
[20] B. Graham, C. Tachtatzis, F. Di Franco, M. Bykowski, D. Tracey, N. Timmons, and J. Morrison, “Analysis of the effect of human presence on a wireless sensor network,” *International Journal of Ambient Computing and Intelligence (IJACI)*, vol. 3, no. 1, pp. 1–13, 2011.
[21] P. Bahl and V. N. Padmanabhan, “RADAR: An in-building RF-based user location and tracking system,” in *Proceedings of 9th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2000)*, vol. 2, 26-30 March 2000, pp. 775–784.
[22] “Gnu radio toolkit,” Online: http://gnuradio.org.
[23] Internet Engineering Task Force, “RFC 6455 - The WebSocket protocol,” Online: https://tools.ietf.org/html/rfc6455, dec 2011.
[24] “The Dark Sky Forecast API,” Online: htts://developer.forecast.io.
[25] Ettus Research, “USRP B-210 USB software defined radio (SDR),” Online: https://www.ettus.com/product/details/UB210-KIT.
[26] QGIS Development Team, *QGIS Geographic Information System*, Open Source Geospatial Foundation, 2016. [Online]. Available: http://qgis.osgeo.org
[27] V. Agafonkin, “Leaflet - a JavaScript library for interactive maps,” Online: http://leafletjs.com.
[28] A. Popleteev, “HIPS: Human-based indoor positioning system,” in *Proceedings of the 7th International Conference on Indoor Positioning and Indoor Navigation (IPIN-2016)*, 4-7 2016.