This is the pre-print version of the published paper:

Joaquín Torres-Sospedra, Adriano Moreira, Stefan Knauth, Rafael Berkvens, Raul Montoliu, Oscar Belmonte, Sergio Trilles, Maria João Nicolau, Filipe Meneses, António Costa, Athanasios Koukofikis, Maarten Weyn and Herbert Peremans. *A Realistic Evaluation of Indoor Positioning Systems Based on Wi-Fi Fingerprinting: The 2015 EvAAL-ETRI Competition*. Article in *Journal of Ambient Intelligence and Smart Environments* Vol.9(2):263-279, February 2017.

The final publication is available at IOS Press through:
`http://dx.doi.org/10.3233/AIS-170421`

A Realistic Evaluation of Indoor Positioning Systems Based on Wi-Fi Fingerprinting: The 2015 EvAAL-ETRI Competition

Joaquín Torres-Sospedra [a,*,**], Adriano Moreira [b], Stefan Knauth [c], Rafael Berkvens [d], Raul Montoliu [a], Oscar Belmonte [a], Sergio Trilles [a], Maria João Nicolau [b], Filipe Meneses [b,e], António Costa [b], Athanasios Koukofikis [c], Maarten Weyn [d] and Herbert Peremans [f]

[a] *Institute of New Imaging Technologies, Universitat Jaume I, Castellón, Spain*
E-mail: { jtorres , montoliu , belmonte , strilles }@uji.es
[b] *Algoritmi Research Centre, University of Minho, Portugal*
E-mail: { adriano.moreira , joao.nicolau , antonio.costa }@algoritmi.uminho.pt
[c] *Faculty for Computer Science and Mathematics, HFT Stuttgart - Stuttgart University of Applied Sciences,*
Schellingstr. 24, 70174 Stuttgart, Germany
E-mail: { stefan.knauth , athanasios.koukofikis }@hft-stuttgart.de
[d] *IDLab, Faculty of Applied Engineering, University of Antwerp - imec, Groenenborgerlaan 171, 2020 Antwerpen,*
Belgium
E-mail: { rafael.berkvens , maarten.weyn }@uantwerpen.be
[e] *Centro de Computação Gráfica (CCG), Portugal*
E-mail: filipe.meneses@ccg.pt
[f] *ENM, Faculty of Applied Economics, University of Antwerp, Prinsstraat 13, 2000 Antwerpen, Belgium*
E-mail: herbert.peremans@uantwerpen.be

Abstract. This paper presents results from comparing different Wi-Fi fingerprinting algorithms on the same private dataset. The algorithms where realized by independent teams in the frame of the off-site track of the EvAAL-ETRI Indoor Localization Competition which was part of the Sixth International Conference on Indoor Positioning and Indoor Navigation (IPIN 2015). Competitors designed and validated their algorithms against the publicly available UJIIndoorLoc database which contains a huge reference- and validation data set. All competing systems were evaluated using the mean error in positioning, with penalties, using a private test dataset. The authors believe that this is the first work in which Wi-Fi fingerprinting algorithm results delivered by several independent and competing teams are fairly compared under the same evaluation conditions. The analysis also comprises a combined approach: Results indicate that the competing systems where complementary, since an ensemble that combines three competing methods reported the overall best results.

Keywords: Wi-Fi Fingerprinting, Off-site competition, Indoor positioning

---

*Corresponding author. E-mail: jtorres@uji.es adriano.moreira@algoritmi.uminho.pt stefan.knauth@hft-stuttgart.de rafael.berkvens@uantwerpen.be

**Joaquín Torres-Sospedra, Adriano Moreira, Stefan Knauth and Rafael Berkvens contributed equally in the elaboration of this work as first authors

*Corresponding author. E-mail: jtorres@uji.es adriano.moreira@algoritmi.uminho.pt stefan.knauth@hft-stuttgart.de rafael.berkvens@uantwerpen.be

**Joaquín Torres-Sospedra, Adriano Moreira, Stefan Knauth and Rafael Berkvens contributed equally in the elaboration of this work as first authors

1. Introduction

Localization, with an expected 4.4 billion market in 2019 [48], is one of the main pillars for indoor services. Most of the newest applications need to know the user's location to customize their services [19,54,62], monitor people [9], or track Internet-of-Things objects [42], among others. Moreover, the location can also be used for detecting the user's activities and to provide services based on them. Accurate positioning is also considered as a fundamental enabling technology for future 5G mobile networks [59,16]

Although the Global Navigation Satellite Systems (GPS, GLONASS, Galileo or Beidou) support location, they cannot operate satisfactory in indoor scenarios due to numerous factors. Many different technologies and Indoor Positioning Systems (IPS) have been proposed to deal with location indoors. In fact, a spectacular growth of studies about indoor location has been witnessed since RADAR [3] was proposed in 2000. The ubiquity of Wi-Fi Internet connectivity and smartphones make Wi-Fi-based positioning very popular [3,41,17,22,60,47, 49,11,39,42], despite the diversity of the technologies available for indoor positioning: Radio-Frequency Identification (RFID) [33,51,9], Bluetooth [20,44], ZigBee [50], Ultrasound [30,26,55], Magnetic field variations[14,27], LED light [38,34], Ultra Wide Band [25,23], and Hybrid solutions [32], among others. The IPSs based on Wi-Fi fingerprinting techniques are preferred to those based on the propagation model, angle of arrival, time of arrival and time difference of arrival because they do not require any very specialized hardware, line-of-sight to the emitter, or knowledge about the precise location of radio emitters to operate [71,70].

Wi-Fi fingerprinting is based on the Received Signal Strength Indicator (RSSI) associated to each of the Wireless Access Points (WAPs) that are available and on comparisons to a reference database (or radio-map). This reference database contains a set of previously recorded fingerprints at well-known reference points. The location of a device is commonly determined by computing the distance or the similarity between a fingerprint collected by the device and the fingerprints contained in the reference database [49]. Wi-Fi fingerprinting is a complex subject which can profit from well-established Expert System techniques by imple-

menting advanced machine learning techniques (*Bayesian Inference* [72], *Neural Networks* [37,10], *Decision Trees* [69], and *Random Forest* [9], among others).

However, there still exists one important drawback which is the lack of a common framework for evaluation and comparison purposes. Appropriate, comparable and reproducible ways of evaluating IPSs is crucial, from research and commercial points of view. Otherwise the scientific advantages on the state-of-art proposals may remain unclear since data, experimental setup and evaluation metrics may be different.

Driven by this objective, the private databases for Indoor Positioning based on Wi-Fi[1] [63] and Magnetic field[2] [65] were published. Moreover, the off-site track of the EvAAL-ETRI Indoor Location Competition[3] at the Sixth International Conference on Indoor Positioning and Indoor Navigation (IPIN 2015) [4] was organized to promote a meaningful comparison of indoor location algorithms and existing working IPSs. The EvAAL-ETRI competition aimed at establishing benchmarks and evaluation metrics for comparing Ambient Assisted Living (AAL) solutions.

Furthermore, it is well-known that the Wi-Fi signal is affected by some factors, including the presence of human bodies [24]. In the previous on-site competitions (see Section 2.4), the participants had to profile the competition environment by themselves the day before competition. So, a simultaneous profiling, where two or more competing teams are profiling the same scenario at the same time, may occur. This simultaneous profiling might have different side-effects on the Wi-Fi signals because the interferences caused by people attending the conferences and competitions might not equally affect all competition teams. For instance, team A employed only one person to profile the scenario and nobody else was present during its profiling, but team B profiled the scenario when it was crowded. If the actor who holds the device is not surrounded by anybody during the evaluation, Team A may have higher chances to win because the people distribution is very similar to the one used in the profiling. So the team, which has a people distribution similar to the one in the evaluation, may have a higher chance to win the competition. Moreover, the results presented in these

---

[1]https://archive.ics.uci.edu/ml/datasets/UJIIndoorLoc
[2]https://archive.ics.uci.edu/ml/datasets/UJIIndoorLoc-Mag
[3]http://evaal.aaloa.org
[4]http://www.ucalgary.ca/ipin2015/

on-site competitions also depends on external factors such as the device used for profiling & positioning or the strategy used for profiling. In summary, the competitors relied on Wi-Fi fingerprinting generated radio maps with different data. Thus, the previous on-line competitions were not only evaluating the accuracy of the proposed IPSs, but also the mapping strategies and the "luck" during profiling. In the EvAAL-ETRI off-site competition, these external factors were not considered, since competitors had the same reference dataset to generate the radio map. So, we consider that the comparison we present in this paper is more representative about the IPS accuracy itself.

A description of the main contributions of this paper are:

– We describe the evaluation criteria used to test and compare different Indoor Positioning Systems (IPSs) based on Wi-Fi fingerprinting.
– A comparative study of the competing systems under equal conditions. Competitors had the same data, a public database, for fine-tuning and training and they had the same time to provide their estimations on a private unlabeled test set. It is worth mentioning that the competitors had no "inside information" about the private test set and they were not able to over-tune their systems for the competition.
– A simple ensemble is introduced, which combines the strengths of the competing systems in a single method. Combining diverse alternatives improves the accuracy of the global system, as for regression and classification systems [66,67,18,61].
– The experiences and suggestions reported by the competitors to enhance a common framework to compare IPSs.

The rest of this paper is organized as follows. Section 2 describes the importance of a framework for benchmarking IPSs. Section 3 introduces the off-site Wi-Fi Competition. Section 4 shows the results. Finally, some conclusions are given in Section 5.

2. Benchmarking of indoor positioning systems

This section describes the importance of a common and public available framework for benchmarking IPSs, since open available frameworks allow the mean-

ingful comparative analysis of IPSs. Moreover, it also introduces the current competition details and its relation to previous ones.

2.1. *The importance of comparative analysis*

Although physical experimentation is used in 77% of indoor positioning research [1], most of these experiments are carried out in areas that are easily accessible to the researcher, such as his own offices [29]. In fact, the typical evaluation scenario covers one, or more, university departments or buildings. This diversity in evaluating IPSs does not allow for comparative analysis, because the situations can be very different. For instance, the experiments carried at the Tampere University of Technology covered part of the Tietotalo building (10.000 m^2 approx.) [47]. In this work, they collected fingerprints at 96 reference points with a Nokia N900. Their reference dataset contained a total of 206 different WAPs. In contrast, the experimental setup done in [49] covered two three-storey buildings of the University of Minho. In this work, they collected fingerprints at 392 reference points with a laptop computer equipped with three similar USB Wi-FI adapters. Their reference dataset contained data from a total of 101 different WAPs. Although the research, results and conclusions shown in works where experiments are done at the own university are valid, we cannot directly compare them because the scenario, the mapping strategy, the equipment and the Wi-Fi environment radically differ. To make meaningful comparisons between IPSs, they must be tested in the same situation(s).

When IPSs are tested in the same situation, we can rank them given certain criteria. In the vast majority of works present in the literature, the evaluation metric is based on the two-dimensional euclidean distance between the estimated location and the current one. Although the mean error is the typical choice to evaluate IPSs, it can be very dependent on outliers. Some works alternatively use a certain percentile score, which indicates how likely the system will perform below a given error. The hit or miss rate of an identification of a building, floor, or room can be introduced in the error functions through penalty terms. These rates may be selected as the main ranking criteria in some situations, where spatial error is less important, such as in-home monitoring. In other situation, the computation speed may be crucial.

The performance of the IPSs on these aspects can only be fairly compared if the IPSs are deployed in the same situation.

2.2. Open access datasets

In order to compare different IPSs, we must find a common input. Jain [31] calls the workload selection the most crucial part of a system analysis. Particularly, it must be representative of the real application, it must include the impact of external components, and it must allow for repeatability. Open access datasets are ideal for comparing different IPSs, given that they are collected in a way that would also happen in a real application. Since the datasets are static, they allow different researchers to apply their or others' systems to the same data without much effort.

Two important open access datasets were contributed by the competition organizers. The first one was used in the competition [63] and is discussed in detail in Section 3.2. The other was published recently [65] and covers a single large room, where magnetometer data was collected, in addition to acceleration and orientation. This dataset features trajectories through the room, so that also tracking algorithms can be tested. Another important source of common data is the EVARILOS project [29]. From here, samples can be obtained from different locations, including offices and an industrial-like environment. Finally, some datasets available at Crawdad[5], e.g. [57,53,56], are also of interest.

2.3. The EvAAL competition at IPIN 2015

The EvAAL-ETRI Indoor Location Competition aimed at establishing benchmarks and evaluation metrics for comparing AAL solutions[6]. In the 2015 edition, in conjunction with the 2015 IPIN Conference, the competition included three tracks:

- Track 1: "Smartphone based positioning" supported by ETRI;
- Track 2: "Foot-mounted pedestrian dead reckoning positioning" supported by ETRI;
- Track 3: "Wi-Fi fingerprinting in large environments" supported by UJI-INIT.

---

[5]http://www.crawdad.org/purpose-Positioning-Systems.html
[6]http://evaal.aaloa.org

Tracks 1 and 2 were on-site and they took place at the IPIN conference venue in parallel to the conference. Competitors were allowed to survey the area the day before the day of the competition, but they could not deploy any external element to support positioning. At the competition, an external actor had to follow a pre-defined competition path, which reproduces the way people move within a big indoor environment. This path was disclosed 30 minutes before the competition. For Track 1 "Smartphone based positioning", the competitors could use any sensor available on their smartphones. For Track 2, "Foot-mounted pedestrian dead reckoning positioning", the competitors could use Micro-Electro-Mechanical Systems (MEMS) sensors (inertial, compass, and pressure sensors) for positioning, and external electronic devices, such as tablets or notebooks, for control and monitoring.

Track 3 "Wi-Fi fingerprinting in large environments" was off-site, the competitors had access to a large Wi-Fi fingerprint database, the *UJIIndoorLoc* database [63], to which they had to apply their algorithms off-line. Competitors only had to implement and set up their localization systems with the open access database. For the competition, a private testing data without labels (ground-truth locations) was provided to the competitors. Competitors had 6 weeks to provide their predictions about indoor location private testing fingerprint. This track is detailed in Sections 3.1 and 3.2.

2.4. Other competitions

Evaluating AAL Systems Through Competitive Benchmarking (EvAAL) was the first international competition aimed at comparing indoor localization systems, with its first edition organized in 2011. In 2014, two new competitions were born: the *IPIN competition* and the *Microsoft Indoor Localization Competition - ISPN*. The current *EvAAL-ETRI* competition is the continuation of the previous editions of EvAAL and IPIN competitions.

2.4.1. Previous editions of the EvAAL competition

Evaluating AAL Systems Through Competitive Benchmarking is an international competition aimed at benchmarking both advanced prototypes and commercial products for the Ambient Assisted Living domain with particular attention to Indoor Positioning Systems. The EvAAL Indoor Location Competition

was launched in 2011 as an initiative proposed by the universAAL FP7[7] project and promoted by the AAL Open Association [58] in order to overcome some technical and economical issues regarding AAL area.

In the first edition (2011) [12], seven competitors demonstrated their systems at the CIAmI Living Lab in Valencia, Spain in July 2011 [5,58,4]. Each competitor had three hours to install their system, calibrate it, log the measurements and lastly to unmount it and answer a short interview on the system's details. The second edition (2012) [2] was organized in two tracks. Eight competing systems participated in the first track, which was focused on *Indoor Localization and Tracking for AAL*, and it was held in July 2012 at the Smart House Living Lab of the Polytechnic University of Madrid, Spain. The second track, with five competitors, focused on Activity Recognition for AAL, and it was held the next week at the CIAmI Living Lab in Valencia, Spain. The third edition (2013) [8] had the same formula as the previous one. Seven competitors participated in the first track, and five competitors in the second track.

This edition also included a demo on Companion Robots for AAL, held on July 2013 at the Peccioli Living Lab in Pisa, Italy.

In general, the competing systems were evaluated by the Evaluation Committee members and staff members who were present during the competition. They gather all the information (accuracy, availability, installation complexity, user's acceptance and integrability [5]) that was going to be used to compute the final scores. Each edition was officially closed at the annual AAL forum. The forums included a session of short presentations by the competitors and the organizers, followed by a round table for freely discussing localization issues from both theoretical and implementation points of view.

2.4.2. The IPIN Competition

The on-site Indoor Positioning and Navigation Competition was held during the IPIN 2014 Conference at the BEXCO Exhibition Center in Busan, Korea. Similarly to the first three editions of the EvAAL competition, the IPIN competition wanted to establish a well-agreed performance evaluation method for Indoor Positioning Systems as well as to provide opportunities for the participants to learn about the evaluating methods of positioning systems. This competition

---

[7]http://universaal.aaloa.org/

consisted of two well-differentiated tracks, *Smartphone Based Positioning* and *Foot-mounted Pedestrian Dead Reckoning Positioning*, where ten competitors had to use their solutions, making use of the existing environment. During the first day of the competition, competitors set up their IPSs. In the second day, the competition took place with the participants carrying their device along a predefined path marked in the corridors and rooms available in the BEXCO exhibition centre.

Although the previous EvAAL editions and the IPIN 2014 competition have some common features, the IPIN competition was restricted to Smartphone and Pedestrian Dead-Reckoning (PDR) systems. Therefore, the evaluation criteria was based only on the accuracy.

2.4.3. The Microsoft Indoor Localization Competition

The Microsoft Indoor Localization Competition is done in conjunction with the International Conference on Information Processing in Sensor Networks (IPSN). It aims to compare IPSs in the same environment since 2014.

The first edition (2014) [45] took place during the IPSN conference in Berlin, Germany. The evaluation scenario consisted of two $90m^2$ attached rooms and a hallway. In the second edition (2015) [46], they repeated the same formula and the competition was held at IPSN 2015 in Seattle, USA. However, the evaluation scenario was larger and covered an area of $1250m^2$, it consisted of one exhibition room and an open challenging area (open space with automatic stairs, structural columns & beams, elevator shaft, among other structural elements).

In both editions, the competitors were divided into two main categories: Infrastructure-based and Infrastructure-less technologies. During the first day of the competition, competitors set up their IPSs and deployed their custom hardware if required. Some particular restrictions and conditions were applied for those competitors using custom Wi-Fi in the Infrastructure-free category. Moreover, the valid WAPs for Infrastructure-less category were provided by organizers. During the second day, the organizers evaluated the devices by carrying the device (phone, tablet or laptop) above 20 evaluation points whose position were disclosed the day before. The evaluation criteria adopted was based on the mean error, computed as the Euclidean distance between estimated and real positions over the 20 testing points.

For the 2016 edition, the organizers have introduced some interesting changes. Firstly, the upcoming edition will divide systems into three main categories: Commercial off-the-shelf Technologies, Commercial off-the-shelf Technologies with initialization and Modified Commercial off-the-shelf Technologies. Secondly, the evaluation area will include different elevation characteristics so competitors will be required to report the estimated position in three dimensions.

3. Off-site Wi-Fi fingerprinting Competition

This section introduces the details of the Off-site Wi-Fi fingerprint competition, the datasets used in the competition, the competing teams, the description of the competing IPS, and the description of a simple ensemble that tries to effectively combine the different competing approaches.

3.1. Competition Details

In contrast to earlier EvAAL, IPIN and MS competitions, the 2015 EvAAL-ETRI competition featured an off-site track. Participants in this off-site track applied their IPSs to a large scale database containing Wi-Fi fingerprints (see Section 3.2). This allowed participants to prepare their systems, off-line and off-site, in advance. Moreover, the competitors had the same data to generate, configure and tune their IPSs. The idea was to estimate the locations of a private set of Wi-Fi fingerprints, whose ground truth was unknown for competition participants. The competition main aim was to fairly compare different approaches for multi-building, multi-floor positioning using a common reference database.

Each competing team had to register by submitting a short description of their IPS. This abstract was evaluated by the independent Technical Program Committee to asses its feasibility. The private test set (see Section 3.2) was released to all competitors in August 7th 2015 and the organizers scheduled the submission of the results (estimated positions) to September 20th 2015. All competitors had one month and a half to prepare their IPSs and submit up to five different sets of estimations. These estimations were evaluated by the competition organizers, who selected the best performing approach according to the mean positioning error for each team. We define the positioning *error* as the Euclidean distance between the *real position* where the fingerprint was taken and the *estimated position* provided by the competing IPS. Although this distance is computed in two

dimensions, penalties are added for floor error and building error. In particular, we add 4 m. for each wrong floor (absolute difference between the real and estimated floors) and 50 m. if the building was not correctly estimated. The results were presented at the IPIN 2015 conference.

The four finalist teams that participated in the off-site competition were:

– The **RTLS@UM team**: Adriano Moreira, Maria João Nicolau, Filipe Meneses and António Costa. University of Minho, Portugal [52].
– The **HFTS team**: Stefan Knauth, Martin Storz, Habiburrahman Dastageeri, Athanasios Koukofikis, and Nora A. Mähser-Hipp. Stuttgart University of Applied Sciences, Stuttgart, Germany [36].
– The **MOSAIC team**: Rafael Berkvens, Maarten Weyn, and Herbert Peremans. University of Antwerp - iMinds. Antwerp, Belgium [7].
– The **ICSL team**: Seungwon Choi, Jaehyun Yoo, and H. Jin Kim. Seoul National University, Seoul, Korea [13].

The **Technical Program Committee** was composed by independent experts: Machaj Juraj (University of Žilina, Slovakia), Christos Laoudias (University of Cyprus, Cyprus), Antoni Perez (Universit Oberta de Catalunya, Spain) and Robert Piché (Tampere University of Technology, Finland). The **Organizing Committee** was composed by: Joaquín Torres-Sospedra, Raul Montoliu, Sergio Trilles and Oscar Belmonte (Universitat Jaume I, Castellón, Spain).

3.2. The datasets for the off-site competition

The off-site competition's datasets were collected at the Jaume I University [63]. Researchers of the INIT research group collected fingerprints in three buildings: two buildings with four floors, and one building with five floors. The total covered area is almost $110.000 \ m^2$, with 520 WAPs scattered through the environment. The fingerprints were stored in three datasets, two of which are publicly available for training and validation purposes, while the third dataset was provided as input data for the competitors' systems. Ground truth and some information about the user was omitted to perform a realistic evaluation of the competing positioning systems.

The first dataset is the training dataset. It contains 19937 fingerprints, collected at 933 distinct locations. Each fingerprint includes its location. On average, 18 access points are visible in any fingerprint. At 71% of the locations there are

20 fingerprints or more, with up to 80 fingerprints at a single location. At only 2.36% of the locations there are less than 9 fingerprints. The average received signal strength for an access point is $-83.24\ dBm$. For the training dataset, users collected data in 6 days within the period running from May 30th to June 20th of 2013.

The second dataset is the validation dataset. It contains 1111 fingerprints, collected at 1074 distinct locations, where these locations do not exactly correspond to the locations of the training dataset. Each fingerprint includes its location. On average, 16 access points are visible in any fingerprint. At 98% of the locations there is only one fingerprint, with at most 8 fingerprints at a single location. The average received signal strength for an access point is $-77.66\ dBm$. For the validation dataset, users collected data in nine days within the period running from September 19th to October 8th 2013.

The third dataset is the one that was private to the competition organizers. It contains 5179 fingerprints, and only the organizers have access to the locations of where the fingerprints were collected. On average, 16 access points are visible in a fingerprint. The average received signal strength for an access point is $-72.02\ dBm$. For the private test dataset, users collected data in four days divided into two periods: 2 days around the end of November 2013, and 2 days at the end of March 2015. A total of 1395 testing fingerprints, more than 25%, were collected 19 months after the data collected for the radio map (training data).

3.2.1. Detailed analysis of datasets

As mentioned in [63], the database contains unprocessed data. So "strange" fingerprints, non-conventional WAPs, and samples taken by low cost-devices were not removed from the datasets. Thus the IPSs accuracy may be considered more general and less device-dependent, in contrast to those works and competitions in which the same device is used for profiling an evaluating a single IPS within one day.

The analysis of the three datasets shows that the training dataset included 367 "strange" fingerprints. RSSI values higher than $-15dBm$ were detected in 291 fingerprints. In fact, 118 fingerprints and 25 WAPs contain RSSI values equal to $0dBm$, which is an uncommon value in smartphone based fingerprinting. Most of these 118 fingerprints were provided by the *Nexus 4 with Android 4.2.2 device* (86.6%). Moreover, no single WAP was detected in 76 fingerprints, most of them

collected by the *Celkon A27 with Android 4.0.4* device (77.6%) and the *GT-I8160 with Android 2.3.6 device* (18.42%). Although the same procedure was used to collect the three sets, the "strange" fingerprints did not appear in the validation and private testing datasets.

Moreover, the number of observed WAPs is also different across the three datasets: 465 WAPs in the training dataset, 367 in the validation dataset, and 270 in the private test dataset. The private test dataset includes 24 WAPs that are not observed in the training dataset. The analysis of the training and validation datasets also suggests that some WAPs were relocated during the data collection process, that some observations of mobile hotspots were included in the samples, or that some observations of point-to-point networks were also included in the samples. This suspicion is supported by the fact that some WAPs were observed in locations with large spatial separation. In particular, 13 WAPs were observed in locations more than 200 meters apart.

Furthermore, 30 different devices (considering device model and Android version) were used to collect the fingerprints and to emulate the problem of samples diversity (see Table 1). To generate the private test set, 3 new devices with recent Android Versions (4.4 or 5.0) and 2 devices, that were already used for training and validation, with a higher Android version were used. These 5 devices collected 40% of the private test set samples.

The previous facts and database features are of special interest because the competing IPSs have been built up and evaluated with the data collected over a period that extends almost 2 years, and provided by different devices and different people.

3.3. The UM Team competing system

The approach adopted by the RTLS@UM team to address the competition challenge comprises the creation of the radio map and the process used to estimate the positions associated to each one of the samples in the private testing dataset [52].

Table 1

Correspondence between *PhoneID* and real device. Real device's information includes the model description and Android version. *TR* stands for Training dataset; *VL* for Validation dataset and *PT* for Private test dataset.

PhoneID	Android Device	Android Version	TR	VL	PT
0	Celkon A27	4.0.4(6577)	✗	✓	✗
1	GT-I8160	2.3.6	✓	✗	✗
2	GT-I8160	4.1.2	✗	✓	✗
3	GT-I9100	4.0.4	✓	✗	✗
4	GT-I9300	4.1.2	✗	✓	✗
5	GT-I9505	4.2.2	✗	✓	✗
6	GT-S5360	2.3.6	✓	✗	✗
7	GT-S6500	2.3.6	✓	✗	✗
8	Galaxy Nexus	4.2.2	✓	✗	✗
9	Galaxy Nexus	4.3	✗	✓	✗
10	HTC Desire HD	2.3.5	✓	✗	✗
11	HTC One	4.1.2	✓	✗	✗
12	HTC One	4.2.2	✗	✓	✗
13	HTC Wildfire S	2.3.5	✓	✓	✓
14	LT22i	4.0.4	✓	✓	✗
15	LT22i	4.1.2	✗	✓	✓
16	LT26i	4.0.4	✓	✗	✗
17	M1005D	4.0.4	✓	✗	✗
18	MT11i	2.3.4	✓	✗	✗
19	Nexus 4	4.2.2	✓	✗	✗
20	Nexus 4	4.3	✗	✓	✗
21	Nexus S	4.1.2	✗	✓	✗
22	Orange Monte Carlo	2.3.5	✓	✗	✗
23	Transformer TF101	4.0.3	✓	✗	✗
24	bq Curie	4.1.1	✓	✗	✗
25	Nexus 5	5.0.1	✗	✗	✓
26	Orange Rono	4.4.2	✗	✗	✓
27	D2303	4.4.4	✗	✗	✓
28	Wildfire S A510e	4.2.2	✗	✗	✓
29	GT-I9505	4.4.2	✗	✗	✓

The following names and conventions are used in the next paragraphs:

T: The training dataset

V: The validation dataset

U: The final testing dataset

R: The radio map used for positioning estimation

fp_i: Denotes fingerprint i of a radio map $(i > 0)$

fp_0: Denotes a test fingerprint (unknown position)

WAP_i^n: Denotes the nth strongest Wireless Access Point in fingerprint fp_i

$rssi_i^j$: Denotes the RSSI value of the ith WAP in fp_j

k: The number of neighbours in k-nearest neighbours approaches

f: The estimated floor

b: The estimated building

Since each team was given access to a training (T) and validation (V) datasets, both datasets including the position (building, floor and latitude/longitude coordinates) associated to each sample, joining both datasets ($T \bigcup V$) to build the radio map maximizes the available information about the covered area. Moreover, the samples in V are more recent that the samples in T, include information about 55 new WAPs not observed in T, and map some new locations not mapped in T.

Since both datasets (T and V) are made from samples collected by a multitude of different devices, an approach inspired on the work of Laoudias et al. [40] was used to normalize the RSSI values. The normalization process first computes $rssi_d$ as the average RSSI value across all the samples in the radio map, for all WAPs, as measured by device d. This is called the representative RSSI value for device d. The mean of all representative RSSI values, $RSSI_D$, can then be used to compute the deviation of each device d from the mean: $\Delta rssi_d = RSSI_D - rssi_d$. By subtracting this deviation value $\Delta rssi_d$ from all the RSSI values measured by device d, a normalized radio map is obtained.

The position estimation process is based on a hierarchical approach designed for large-scale multi-building, multi-floor settings where computational effort is to be minimized. Given a fingerprint fp_0 taken at an unknown location, the process starts by estimating the building (b) where it was collected, as follows:

1. Take WAP_0^1, the strongest WAP observed in fp_0.
2. Build R', a subset of the radio map R, with all the samples where the strongest WAP is WAP_0^1 (filtering).
3. If R' is an empty set, repeat steps 1 and 2 for the 2nd, 3rd, ..., strongest WAP in fp_0.
4. Count the number of samples in R' associated to each building and set b to the most frequent building (majority rule).

This process resulted in a 100% hit rate in estimating the correct building. One advantage of this solution is that it can be implemented by a single SQL query over the radio map database, thus avoiding the effort of computing the similarity between the given fingerprint and all the samples in the radio map. The SQL query that implements steps 1, 2 and 4 above is simply:

```
SELECT building FROM radiomap WHERE strongAP1=AP01
```

```
GROUP BY building ORDER BY count(*) DESC LIMIT 1;
```

and requires that each entry in the radio map be appended with the index of its strongest WAPs (`strongAP1, strongAP2, ...`), which is done only once while building the radio map. This query might have to be repeated (step 3 above), by replacing `strongAP1` by `strongAP2` and so on, if the result set is empty, which is quite unlikely.

Floor (f) estimation is achieved by a combination of filtering, k-NN classification, and majority rule-based selection operations, as described by the following procedure:

1. Build R', a subset of R, with all the samples from the building b estimated in the previous process (building estimation) (filtering).
2. Build R'', a subset of R', with all the samples where $(RSSI_0^1 - \Delta RSSI) \leq RSSI_i^1, \leq (RSSI_0^1 + \Delta RSSI)$, with $\Delta RSSI$ being a parameter (filtering).
3. If $\#(R'') < n$, then $R'' = R'$, where $\#(.)$ denotes the cardinality of a set, and n is a parameter.
4. Compute the similarity, $S()$, between fp_0 and all the fingerprints in R''.
5. Take the $k1$ samples in R'' that are the most similar to fp_0.
6. Count the number of samples, from within the $k1$, associated to each floor, and set f to the most frequent floor (majority rule).

In step 2, the idea is to use only the samples in the radio map where the RSSI value is somehow similar (within $\Delta RSSI$) to the RSSI value of the strongest AP in the given fingerprint ($RSSI_0^1$). A value of $\Delta RSSI = 12$ showed to provide good results. Again, computational efficiency is gained here since it reduces the number similarity calculations needed in step 4. Additionally, steps 1 to 3 can also be implemented by a simple SQL query over the radio map database. The similarity function $S()$ referred in step 4 is a variant of the Manhattan distance defined as:

$$S(fp_1, fp_2) = \frac{1}{N} \sum_{i=1}^{N} |rssi_i^1 - rssi_i^2| - 2C \qquad (1)$$

where N is the total number of WAPs observed in fp_1 and/or fp_2, and C is the number of WAPs that were observed in both fp_1 and fp_2 (common WAPs). For missing WAPs, in fp_1 or fp_2, a default RSSI value of -90 was used. For

the datasets made available to this competition, a value of $k1 = 50$ showed to maximize the floor hit rate [52].

After estimating the floor, the geometric coordinates are estimated based on a simple k-NN classification procedure as follows:

1. Build R''', a subset of R'', with all the samples where the floor f is the one estimated in the previous process (floor estimation) (filtering).
2. Compute the similarity, $S()$, between fp_0 and all the fingerprints in R'''.
3. Take the $k2$ samples in R''' that are the most similar to fp_0.
4. Compute the estimated coordinates as the centroid of the $k2$ samples.

Note that R''' is a subset of R'', and that the similarity between fp_0 and all the samples in R'' has already been computed during the floor estimation procedure (at step 4). Therefore, the more computational intensive task associated to step 2 can be avoided, thus speeding up the estimation process. A value of $k2 = 7$ showed to provide good results [52].

3.4. The HFTS Team competing system

The HFT Team competed with two algorithms: The "*Fingerprint Calibrated Weighted Centroid*" (FCWC) method and the "*Scalar Product Correlation Fingerprinting*" (SPCF) algorithm.

The first approach uses weighted centroid to calculate the rover position. For this approach, the position of the WAPs are needed. Since these WAP positions were not known, they were estimated using the calibration dataset of the competition: from the existing RSSI readings of a certain WAP at different test points, the position of this WAP was calculated using weighted centroid. This "reverse positioning" is performed once to generate an WAP position database.

The second approach uses processed fingerprints to estimate positions via k-NN. The SPCF algorithm performed slightly better in the competition. Therefore all numbers and scores in the paper, which belong to the HTF team, have been obeyed with the SPCF algorithm.

Both algorithms are model based: The Frijs equation [21] with modified propagation exponent n (see for example [28]) is used to translate an RSSI value into an inverse distance estimation:

$$w = \begin{cases} 10^{S[dBm]/(10 \cdot n)}, \text{if } S \geq RSSI_{min} \\ 0, \qquad \text{otherwise} \end{cases}, \tag{2}$$

where w is the inverse distance, S is the measured RSSI value, and $RSSI_{min}$ is a lower RSSI threshold. The equation does not contain a constant factor representing the transmission power, as both algorithms would eliminate that constant anyways.

For the FCWC algorithm, the inverse distance is used as weight for the summing of the WAP positions. For the SPCF algorithm, inverse distance vectors are compared using the scalar product as comparison norm: U, V are two vectors of weighted RSSI values $U : u_q = w(RSSI_{U,q})$ respectively $V : v_q = w(RSSI_{V,q})$, with q denoting the WAP index. For missing values a weight of zero is assigned to u_q respectively v_q. The normalized scalar product, eq.3, is used as a measure for the equality of two RSSI vectors.

$$nsp(U, V) = cos(\alpha) = \frac{\sum u_q v_q}{|U| \, |V|} \tag{3}$$

Where nsp values vary between [-1..1], with higher values indicating higher correlation or "nearness".

The position estimate is obtained by using k-NN with that norm: Those three positions R_i of the training points with the highest correlation values are averaged with a weight of 2,1 and 1, in descending correlation order, and the result is the position estimate. The propagation exponent n and $RSSI_{min}$ are determined manually for both algorithms, globally as well as for each building, by minimizing the positioning error for the training dataset. A more detailed description is given in [35].

An important competition task is the the building and floor estimation. For both described algorithms, positions are always three dimensional coordinates and comprise (x,y,z), i.e. northing, easting and height. The z component for the calibration measurement positions is generated using the floor ID. The algorithms, i.e. FCWC and SPCF, operate on these three-dimensional coordinates and therefore produce not only a (x,y) result but also a z result. No special al-

gorithm for building- and floor-ID determination is applied. Instead, a bounding box is defined for each building, and the building ID for a certain measurement is obtained by matching the calculated x and y coordinates with the boxes. The floor ID is given by the z component of the position estimate.

3.5. The MOSAIC Team competing system

The MOSAIC competition system is based on the Wi-Fi positioning system that was included in the Opportunistic Seamless Localization system [68]. This system defines weights for three distinct situations in Wi-Fi fingerprint comparison. The comparison is between a fingerprint collection that describes the radio map, and a fingerprint that is measured for localization. Each access point in the measured fingerprint is compared with the corresponding access point of a fingerprint in the collection that describes the radio map, for each fingerprint in this collection. The fingerprints in the collection are the mean values of all fingerprints measured at that location during the training phase. In the competition, the team used only the training dataset values.

The three situations are: first, the access point in the measured fingerprint was also measured with some received signal strength in the collection's fingerprint—defined as a *hit*; second, an access point in the collection's fingerprint does not occur in the measured fingerprint—defined as a *miss*; third, an access point in the measured fingerprint does not occur in the collection's fingerprint—defined as an *extra*. The MOSAIC team additionally defined the situation where an access point does not occur in both the measured fingerprint and the collection's fingerprint—we call this a *none*.

The MOSAIC system calculates the likelihood of the given situation for each access point in the measured fingerprint. To compute the likelihood of the measured fingerprint, the algorithm multiplies the likelihoods of the access points, assuming that they are independent [7]. By normalizing the likelihood of the measured fingerprint over all locations in the training dataset, the algorithm can calculate the posterior probability distribution of the location given the measured fingerprint. It then selects a location by using k-NN, with $k = 3$, weighted by the posterior value.

Adapting the MOSAIC system to the competition data representation required little effort. There was a slight difference in location specification: the MOSAIC

system uses a relative system without any floors, while the competition data uses an absolute system in different buildings and floors. This only influences the interpretation of the results, it has no influence on the MOSAIC likelihood calculation, which is by design agnostic to the position representation. However, since a k-NN is also implemented, the floors and buildings have to be taken into account. The building and floor identification numbers are weighted by the posterior value, just like the position is, and then rounded to again obtain a building and floor identification number.

3.6. The ensemble approach

Ensembles are commonly used in regression and classification tasks to combine a set of individual estimators (classifiers or regressors) [18,61]. The error may be reduced when the individual estimators have a high degree of diversity [66,67].

Here a simple ensemble approach has been developed as a first attempt to fuse the three competing systems in an advanced IPSs. The main idea was to try to combine the strengths of all the competing systems. Although there were four finalists in the competition, only three of them accepted to include their system in the ensemble and participate in the elaboration of this paper. The ensemble works as follows:

1. Estimate the position with the three competing systems.
2. Apply a voting procedure to estimate the building. The most often voted building among the three competing systems' estimations is assigned as the ensemble's estimated building. The three competing systems' estimations are equally weighted.
3. Apply a voting procedure to estimate the floor. The most often voted floor among the competing systems' estimations is assigned as the ensemble's estimated floor. Only the competing systems' estimations which belong to the ensemble's estimated building are considered in this voting. In case of tie in this second voting procedure, the HFTS system has the highest priority followed by the MOSAIC team.
4. The ensemble's estimated coordinates correspond to the RTLS@UM's IPS if its estimated building and floor correspond to the ensemble's estimated building and floor. If not, the estimated coordinates correspond to the

HFTS's IPS if its estimated building and floor correspond to the ensemble's estimated building and floor. If not, the estimated coordinates correspond to the MOSAIC's IPS.

5. Return ensemble's estimated coordinates, floor and building.

The results -error and building&floor hit detection rate- of the three competing IPSs before the competition on the validation set, included in the short description or implemented by the organization committee, have been used to set the priorities in Steps 3 and 4. In particular, the hit rate is used to establish the priorities in Step 3, whereas the error in positioning is used to establish the priorities in Step 4.

4. Results and Discussion

This section introduces the results of a simple baseline method, the competing IPSs and the ensemble approach. Moreover, some discussion about the results and the competition is introduced.

4.1. Baseline results

A baseline based on the *k-Nearest Neighbor* rule (k-NN) [15] is introduced for comparison purposes. In particular, the k-NN algorithm with $k = 1$ and the *Manhattan distance* (or *city block*) as base distance to calculate the closest(s) neighbors is used for the baseline. The reference dataset is generated from the training and validation subsets of the *UJIIndoorLoc* database using the *Positive values* data representation [64] (see eq.4), where the new low values stand for low signal (0 means that the WAP has not been detected) and the higher values indicate that the signal is stronger. Finally, all the WAPs (520 in this work) are considered to calculate the *Manhattan distance* between two fingerprints in our implementation of the 1-NN algorithm, it even compute those WAPs which are not detected (value 0) in any of the two compared fingerprints. The results of this baseline are shown in Table 2, where the mean error and the percentile values are based on the positioning error with penalties described in Section 3.1.

$$NewRSSI = \begin{cases} OldRSSI + 105 & \text{if } OldRSSI < 0 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

Table 2

Baseline Results

	Baseline
Mean Error	8.46m
25th percentile	3.39m
50th percentile	6.50m
75th percentile	11.72m
95th percentile	21.41m
100th percentile	73.30m
Building Hit rate	100%
Floor Hit rate	85.34%

Table 3

Competition and Ensemble Results

	HFTS	MOSAIC	RTLS@UM	ICSL	Ensemble
Mean Error	8.49m	11.64m	6.20m	7.67m	6.10m
25th percentile	3.69m	3.26m	2.51m	3.10m	2.51m
50th percentile	6.99m	6.72m	4.57m	5.88m	4.56m
75th percentile	11.60m	12.12m	8.34m	10.87m	8.24m
95th percentile	19.93m	21.54m	15.81m	19.68m	15.41m
100th percentile	40.70m	313.33m	52.27m	39.14m	52.27m
Building Hit rate	100%	98.65%	100%	100%	100%
Floor Hit rate	96.25%	93.86%	93.74%	86.93%	96.43%

According to the results shown in Table 2, the baseline mean error is 8.46 m. but the floor hit rare is 85.34%. In all the cases, the building has been correctly estimated.

4.2. Competition results

The results of the competing teams are shown in Table 3. Again, the mean error and the percentile values are based on the positioning error with penalties described in Section 3.1. Moreover, the results of the simple ensemble approach (see Section 3.6) are also included in the table.

According to the results shown in Table 3, the RTLS@UM IPS provided the best competing results according to the mean positioning error (including the floor and building error penalties) and, therefore, this team was the winner of the off-site track in the 2015 EvAAL-ETRI competition. Moreover, this team also reported the lowest error in most of percentile values. Although the competition metric was the mean error, the EvAAL website reports ranks based on the third

Table 4

Contribution of the competing systems on estimating the correct floor.

Systems	Floor Hit rate
Only HFTS	0.23%
Only MOSAIC	0%
Only RTLS@UM	0%
HFTS + MOSAIC	2.86%
HFTS + RTLS@UM	3.98%
MOSAIC + RTLS@UM	0.66%
ALL	88.70%
TOTAL	96.43%

percentile for all the tracks. In the off-site track, ranks provided by the mean error and third quartile are identical.

Although the HFTS team reported a mean error similar to the baseline, it provided the overall highest floor hit rate (96.25%). The HFTS's IPS reduces the wrong-floor errors in a 2.5% with respect to the teams with highest hit rate.

The MOSAIC team provides the third best 25th and 50th (median) percentiles and also the second best Floor hit rate among the competing IPSs. In contrast to the other competing IPSs, it reports a Building hit rate of 98.65% which had a negative impact on the mean error.

The ICSL team reports the second lowest mean error. However, it was the competing team reporting the lowest floor hit rate.

The ensemble approach, which combines the three competing IPSs, provides the best overall results. The three competing systems equally contributed on estimating the building and floor (see Table 4). The coordinates were mainly provided by the RTLS@UM system (96.97% of cases) and the rest were provided by HFTS system (3.03% of the cases). Although the ensemble improvements with respect to the best errors and hit rates are marginal, they are not when the ensemble is compared to the individual competing systems. For RTLS@UM system, the ensemble and the competing system provide similar errors in positioning but the ensemble provides a hit floor detection rate 3% higher. For HFTS system, both provide similar hit detections rates but the ensemble provides a positioning error about 2.4 m lower. The ensemble takes benefit from the three independent competing systems, it provides the highest floor hit rate and the lowest positioning error.
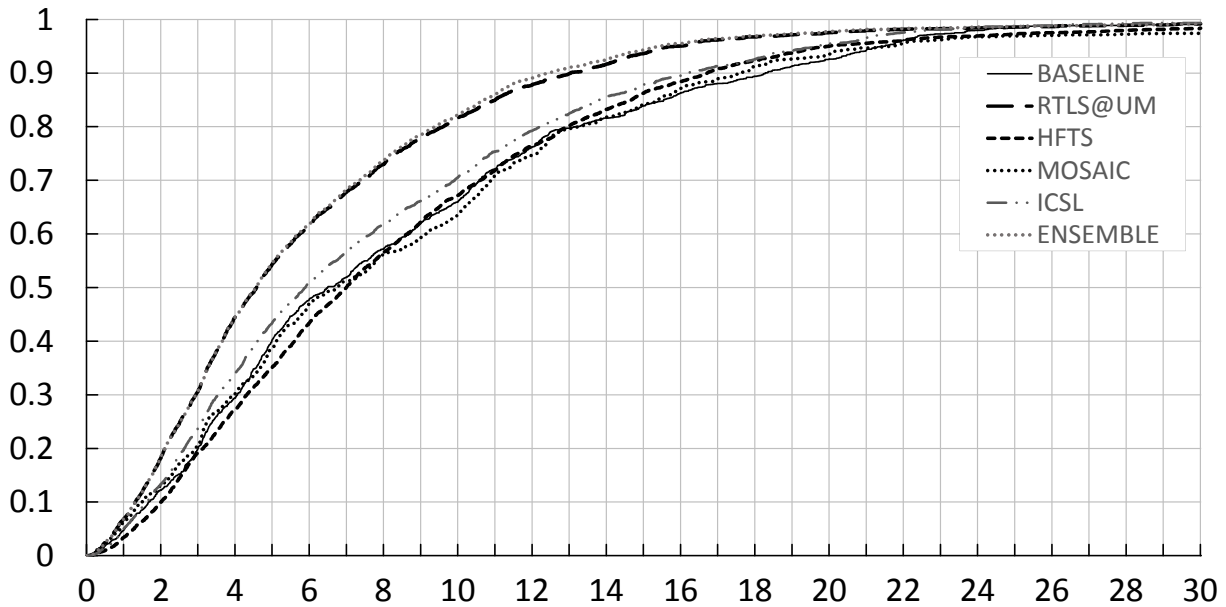
Fig. 1. Distribution of the errors in positioning with penalties for the baseline, competing IPSs and ensemble

Figure 1 shows the Cumulative Distribution Function (error in positioning with penalties) for the baseline, the competing IPSs and the ensemble approach. In the figure, it can be seen that there are two differentiated groups of IPSs. The first group reports the lowest error and it is formed by the IPS developed by RTLS@UM IPS and the ensemble, whereas the baseline and the IPSs developed by MOSAIC and HFTS form the second group. On the second group, the IPS developed by ICSL reports the lowest error. Among the other systems, the IPS developed by MOSAIC tends to be the one until the 35th percentile and the IPS developed by HTFS tends to be the best from 60 to 95 percentile. Although the differences between two IPSs are marginal in some cases, the error function does not explicitly include the floor hit rate. As stated in [10] wrong-floor errors may not be acceptable at all because it is much easier to move within the same floor than among floors. Although the results reported by the ICSL team and the baseline are similar to the IPS provided by HFTS and MOSAIC, they should be discarded due their low floor hit rate.

4.3. Discussion

This section introduces the discussion derived from the competing systems, the results shown and the experiences reported by the competitors.

This competition has allowed the meaningful comparison of different systems. As pointed out by competitors after the closing session, the results suggested that a hybrid technique might lead to significant improvements in Wi-Fi fingerprinting.

An ensemble approach, which combines three competing systems, has been used and it reports the best overall results according to the positioning error and floor hit rate. Combining diverse systems is useful and tends to reduce the errors provided by the individual systems.

The results of the ensemble approach encourages to develop advanced positioning systems that include different approaches to provide more accurate positioning results.

4.3.2. Metrics used in the competition

The EvAAL-ETRI competition used the mean positioning error with penalties as main performance metric. They also provided information on the cumulative distribution of the positioning error, improving the performance description of the different IPS. Furthermore, the floor and building hit rates were also provided. In practice, these are very relevant metrics, but even more details on the systems could be used to evaluate the performance.

For example, the latency and energy efficiency of the systems could be taken into account, as discussed in [29]. To compare these metrics, they must be deployed on the same or very similar hardware. This increases the complexity of the competition, unless a dedicated testbed is available. However, there is no consensus on what to include when measuring the localization latency.

Additionally, there is research on quantifying the information provided by probabilistic sensor model in localization [6]. The mean mutual information or the conditional entropy can be used as a metric for the performance of the measurement model in a given environment. This separates the measurement model from the ultimate location estimation step. It is necessary, however, that the IPS is probabilistic.

4.3.3. General experiences reported by competitors

In general, the competition was of great interest for the IPIN conference attendees (the session was quite well visited) and the competing teams reported that their experiences at the competition were positive. The RTLS@UM team

used the opportunity provided by the competition to evolve their previous estimation algorithms. The MOSAIC team enjoyed joining the EVAAL-ETRI track 3 competition, because of its accessible approach, using the huge fingerprinting database. The HFTS team extended their developments in order to provide the 3D positioning (lat,lot, & floor) with building estimation according the *UJIIndoorLoc* format. In general, all the teams had to adapt their position estimation systems to the format of the provided datasets, so information and restrictions about other coordintate representations had to be considered.

The competitors suggested some minor improvements for further editions that can be summarized as follows:

- All competitors should be more involved in the elaboration and review of the competition technical annex. After registration deadline, all competitors could provide ideas to improve the competition via, for instance, the competition mailing list.
- The process to join the competition and report the results should be simplified. In the current competition, the competitors had to submit an extended abstract and a short paper before the conference, which introduced some delay in the evaluation process. Submitting an extended abstract detailing the competing IPS and results on the public datasets should be enough to join the competition. Writing a paper reporting the results should be a post-conference optative step.
- The dates for delivering the datasets should be known in advance. In order to analyze the scenario and IPS fine-tunning with public datasets, they should be released as soon as possible.
- The private test dataset should include more information about the specific device and the user who collected it. According to the competitors, it would be realistic and enable the use of more advanced estimation techniques, such as Predicted k-Nearest Neighbours [43], which uses recent past information of users to improve the accuracy of the localization algorithm.
- The public and private datasets should be extended to include more recent data, new buildings, new devices to collect the fingerprints and, even, new sequences of fingerprints emulating user trajectories at different speeds.

5. Conclusions

This paper introduces the results of the 2015 EvAAL-ETRI competition, Track 3 "Wi-Fi Fingerprinting in large environments". The data used to train and fine-tune the IPS were open access and they were generated using information provided by different users and devices to emulate the problem of samples diversity. As far as we know, it has been the first off-site competition on Wi-Fi fingerprinting where all the participants competed in equal conditions in such large indoor scenario, which was formed by three multi-storey buildings.

Another advantage of the off-site competition was that the competing teams did not have to deploy and configure their systems on-site. Moreover, competitors had enough time to fine-tune their systems with public datasets. In contrast to other on-site competitions, all the competing teams completed their tasks and none of them had to quit the competition.

To allow meaningful and comprehensive comparisons of IPSs, it is necessary to have a common comparative framework. In fact, the use of common public databases and evaluation metrics made possible the fair comparison of the competing systems and the development of an ensemble, which reported the lowest overall error and the highest overall floor hit rate.

The competition had a winner, the RTLS@UM team, which reported the lowest mean error. Other measures, such as the floor hit rate and the distribution of errors, indicate that the other competing systems also reported interesting results. The competitors employed different approaches to develop their respective positioning systems. In fact, three different competing approaches have been successfully combined in an ensemble, which obtains the best overall mean error and the best overall floor hit rate. A set of diverse classifiers has outperformed any of the individual systems that composes it.

Finally, off-line competitions, such as the current one, are useful according to the experiences reported by the competitors. The competitors had not only the opportunity to compete and rank their systems, but also to analyze and compare different alternative and complementary ways to deal with Wi-Fi fingerprinting.

References

[1] Stephan Adler, Simon Schmitt, Katinka Wolter, and Marcel Kyas. A Survey of Experimental Evaluation in Indoor Localization Research. In *Indoor Positioning and Indoor Navigation (IPIN), 2015 International Conference on*, page 10, Banff, AB, Canada, oct 2015. IPIN.

[2] Juan Antonio Álvarez-García, Paolo Barsocchi, Stefano Chessa, and Dario Salvi. Evaluation of localization and activity recognition systems for ambient assisted living: The experience of the 2012 evaal competition. *Journal of Ambient Intelligence and Smart Environments*, 5(1):119–132, 01 2013.

[3] P. Bahl and V.N. Padmanabhan. Radar: an in-building rf-based user location and tracking system. In *Proceedings of the Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings.*, volume 2, pages 775–784, 2000.

[4] Paolo Barsocchi, Stefano Chessa, Francesco Furfari, and Francesco Potortì. Evaluating AAL solutions through competitive benchmarking: the localization competition. *IEEE Pervasive Computing Magazine*, 12(4):72–79, October–December 2013.

[5] Paolo Barsocchi, Francesco Potortì, Francesco Furfari, and Alejandro M. Medrano Gil. Comparing AAL indoor localization systems. In *Evaluating AAL Systems Through Competitive Benchmarking. Indoor Localization and Tracking*, volume 309 of *Communications in Computer and Information Science*, pages 1–13. Springer Berlin Heidelberg, 2012.

[6] Rafael Berkvens, Herbert Peremans, and Maarten Weyn. Localization Performance Quantification by Information Theory. Technical report, University of Antwerp, 2015.

[7] Rafael Berkvens, Maarten Weyn, and Herbert Peremans. Localization Performance Quantification by Conditional Entropy. In *Proceedings of the Sixth International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2015.

[8] Juan A. Botia, Juan Antonio Alvarez-Garcia, Kaori Fujinami, Paolo Barsocchi, and Till Riedel, editors. *Evaluating AAL Systems Through Competitive Benchmarking: International Competitions and Final Workshop, EvAAL 2013*, volume 366 of *Communications in Computer and Information Science*. Springer, 2013.

[9] Luca Calderoni, Matteo Ferrara, Annalisa Franco, and Dario Maio. Indoor localization in a hospital environment using random forest classifiers. *Expert Systems with Applications*, 42(1):125 – 134, 2015.

[10] Rafael Saraiva Campos, Lisandro Lovisolo, and Marcello Luiz R. de Campos. Wi-fi multi-floor indoor positioning considering architectural aspects and controlled computational complexity. *Expert Systems with Applications*, 41(14):6211 – 6223, 2014.

[11] Yin Chen, Dimitrios Lymberopoulos, Jie Liu, and Bodhi Priyantha. Indoor localization using fm signals. *IEEE Transactions on Mobile Computing*, 12(8):1502–1517, 2013.

[12] Stefano Chessa, Francesco Furfari, F. Potortì, Paolo Barsocchi, Mohammad-Reza Tazari, and Reiner Wichert. Evaluating AAL Systems through Competitive Benchmarking (EVAAL)-Technical Aspects of the First Competition. In *Partnerships for Social Innovation in Europe, Ilse Bierhoff, Henk Herman Nap, Wil Rijnen, Reiner Wichert editors, ISBN 978-90-819709-0-7 (2012) Proceedings of the AAL Forum, Lecce, Italy, September 26-28*, pages 617–623.

[13] Seungwon Choi, Jaehyun Yoo, and Hyoun Jin Kim. Machine learning for indoor localization: Deep learning and semi-supervised learning. In *USB On-Site Proceedings of the Sixth International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2015.

[14] Jaewoo Chung, Matt Donahoe, Chris Schmandt, Ig Jae Kim, Pedram Razavai, and Micaela Wiseman. Indoor location sensing using geo-magnetism. In *Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services*, MobiSys '11, pages 141–154, New York, NY, USA, 2011. ACM.

[15] T. Cover and P. Hart. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theor.*, 13(1):21–27, September 1967.

[16] A. Dammann, R. Raulefs, and Siwei Zhang. On prospects of positioning in 5g. In *Communication Workshop (ICCW), 2015 IEEE International Conference on*, pages 1207–1213, June 2015.

[17] Antonio del Corte-Valiente, José Manuel Gómez-Pulido, and Oscar Gutiérrez-Blanco. Efficient techniques and algorithms for improving indoor localization precision on wlan networks applications. *International Journal Communications, Networks and System Sciences*, 2(7):645–651, 2009.

[18] Thomas G. Dietterich. Ensemble methods in machine learning. In Josef Kittler, editor, *Multiple Classifier Systems. First International Workshop, MCS 2000*, number 1857 in Lecture Notes in Computer Science, pages 1–15. Springer Berlin / Heidelberg, 2000.

[19] Alberto García Estevez and Niklas Carlsson. Geo-location-aware emulations for performance evaluation of mobile applications. In *Proceedings of the IEEE/IFIP Conference on Wireless On-demand Network Systems and Services*, 2014.

[20] Silke Feldmann, Kyandoghere Kyamakya, Ana Zapater, and Zighuo Lue. An indoor bluetooth-based positioning system: Concept, implementation and experimental evaluation. In *Proceedings of the International Conference on Wireless Networks*, pages 109–113. CSREA Press, 2003.

[21] H.T. Friis. A note on a simple transmission formula. *Proceedings of the IRE*, 34(5):254–256, May 1946.

[22] Sebastian Gansemer, Uwe Großmann, and Syuzanna Hakobyan. Rssi-based euclidean distance algorithm for indoor positioning adapted for the use in dynamically changing wlan environments and multi-level buildings. In *Proceedings of the 1st the International Conference on Indoor Positioning and Indoor Navigation (IPIN'2010)*, 2010.

[23] Enrique García, Pablo Poudereux, Álvaro Hernández, Juan Jesús García, and Jesús Ureña. DS-UWB indoor positioning system implementation based on FPGAs. *Sensors and Actuators A: Physical*, 201:172 – 181, 2013.

[24] Sergio Garcia-Villalonga and Antoni Perez-Navarro. Influence of Human Absorption of Wi-Fi Signal in Indoor Positioning with Wi-Fi Fingerprinting. In *Proceedings of the Sixth International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2015.

[25] S. Gezici, Zhi Tian, G.B. Giannakis, Hisashi Kobayashi, A.F. Molisch, H.V. Poor, and Z. Sahinoglu. Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks. *Signal Processing Magazine, IEEE*, 22(4):70–84, July 2005.

[26] David Gualda, Jesús Ureña, Juan C. García, and Alejandro Lindo. Locally-referenced ultrasonic - lps for localization and navigation. *Sensors*, 14(11):21750, 2014.

[27] Yansha Guo, Yiqiang Chen, and Junfa Liu. Indoor location estimation based on local magnetic field via hybrid learning. In Fuchen Sun, Kar-Ann Toh, Manuel Grana Romay, and Kezhi Mao, editors, *Extreme Learning Machines 2013: Algorithms and Applications*, volume 16 of *Adaptation, Learning, and Optimization*, pages 189–207. Springer International Publishing, 2014.

[28] H. Hashemi. The indoor radio propagation channel. *Proceedings of the IEEE*, 81(7):943–968, Jul 1993.

[29] Tom Van Haute, Eli De Poorter, Filip Lemic, Vlado Handziski, Niklas Wirstrom, Thiemo Voigt, Adam Wolisz, and Ingrid Moerman. Platform for benchmarking of RF-based indoor localization solutions. *Communications Magazine, IEEE*, 53(9):126–133, 2015.

[30] F. Ijaz, Hee Kwon Yang, A.W. Ahmad, and Chankil Lee. Indoor positioning: A review of indoor ultrasonic positioning systems. In *Proceedings of the Advanced Communication Technology*, pages 1146–1150, Jan 2013.

[31] Raj Jain. *The art of computer systems performance analysis. techniques for experimental design, measurement, simulation, and modeling*. New York, N.Y., Wiley, 1991.

[32] Antonio R. Jimenez Ruiz, F.S. Granja, J.C. Prieto Honorato, and J.I.G. Rosas. Accurate pedestrian indoor navigation by tightly coupling foot-mounted imu and rfid measurements. *Instrumentation and Measurement, IEEE Transactions on*, 61(1):178–189, Jan 2012.

[33] Guang Yao Jin, Xiao Yi Lu, and Myong Soon Park. An indoor localization mechanism using active rfid tag. In *Proceedings of the IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing*, volume 1, pages 4 pp.–, June 2006.

[34] Mohsen Kavehrad, M. I. Sakib Chowdhury, and Zhou Zhou. *Indoor Positioning Methods Using VLC LEDs*, pages 225–262. John Wiley & Sons, Ltd, 2015.

[35] Stefan Knauth, M. Storz, H. Dastageeri, A. Koukofikis, and N.A. Mahser-Hipp. Fingerprint calibrated centroid and scalar product correlation rssi positioning in large environments. In *Indoor Positioning and Indoor Navigation (IPIN), 2015 International Conference on*, pages 1–6, Oct 2015.

[36] Stefan Knauth, Martin Storz, Habiburrahman Dastageeri, Athanasios Koukofikis, and Nora Mähser-Hipp. Fingerprint calibrated centroid and scalar product correlation rssi positioning in large environments. In *Proceedings of the Sixth International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2015.

[37] R.J. Kuo, M.C. Shieh, J.W. Zhang, and K.Y. Chen. The application of an artificial immune system-based back-propagation neural network with feature selection to an RFID positioning system. *Robotics and Computer-Integrated Manufacturing*, 29(6):431 – 438, 2013.

[38] Ye Sheng Kuo, Pat Pannuto, Ko Jen Hsiao, and Prabal Dutta. Luxapose: Indoor positioning with mobile phones and visible light. In *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking*, MobiCom '14, pages 447–458, New York, NY, USA, 2014. ACM.

[39] Kun Chan Lan and Wen Yuah Shih. On calibrating the sensor errors of a pdr-based indoor localization system. *Sensors*, 13(4):4781–4810, 2013.

[40] Christos Laoudias, Robert Piche, and Christos G. Panayiotou. Device self-calibration in location systems using signal strength histograms. *Journal of Location Based Services*, 7(3):165 – 181, 2013.

[41] Erin Ee Lin Lau and Wan Young Chung. Enhanced rssi-based real-time user location tracking system for indoor and outdoor environments. In *Proceedings of the 2007 International Conference on Convergence Information Technology*, pages 1213–1218, Washington, DC, USA, 2007. IEEE Computer Society.

[42] Wenhua Le, Zhanbin Wang, Jingcheng Wang, Guanglei Zhao, and Haoxuan Miao. A novel wifi indoor positioning method based on genetic algorithm and twin support vector regression. In *Control and Decision Conference (2014 CCDC), The 26th Chinese*, pages 4859–4862, May 2014.

[43] Binghao Li, James Salter, Andrew G. Dempster, and Chris Rizos. Indoor positioning techniques based on wireless lan. Technical report, 2006.

[44] Honggui Li. Low-cost 3d bluetooth indoor positioning with least square. *Wireless Personal Communications*, pages 1–14, 2014.

[45] Dimitrios Lymberopoulos, Romit Roy Choudhury, Xue Yang, and Souvik Sen. Microsoft Indoor Localization Competition - IPSN 2014. Accessed: November 2015.

[46] Dimitrios Lymberopoulos, Jie Liu, Xue Yang, Ayman Naguib, Anthony Rowe, Niki Trigoni, and Nader Moayeri. Microsoft Indoor Localization Competition - IPSN 2015. Available at http://research.microsoft.com/en-us/events/indoorloccompetition2015 Accessed: November 2015.

[47] J. Machaj, P. Brida, and R. Piché. Rank based fingerprinting algorithm for indoor positioning. In *Proceedings of the 2nd the International Conference on Indoor Positioning and Indoor Navigation (IPIN'2011)*, pages 1–6, 2011.

[48] Markets and Markets. Indoor Location Market by Solution (Tag-based, RF-based, Sensor-based), by Application (Indoor Maps & Navigation, Indoor Location-based Analytics, Tracking & Tracing, Monitoring & Emergency Management), by Service, by Vertical, & by Region - Global Forecast Up to 2019.

[49] Nelson Marques, Filipe Meneses, and Adriano Moreira. Combining similarity functions and majority rules for multi-building, multi-floor, wifi positioning. In *Proceedings of the 3th the International Conference on Indoor Positioning and Indoor Navigation (IPIN'2012)*, 2012.

[50] José V. Martí, Jorge Sales, Raúl Marín, and Ernesto Jiménez-Ruiz. Localization of mobile sensors and actuators for intervention in low-visibility conditions: The zigbee fingerprinting approach. *International Journal of Distributed Sensor Networks*, page 10 pages, 2012.

[51] Ali Montaser and Osama Moselhi. RFID indoor location identification for construction projects. *Automation in Construction*, 39(0):167 – 179, 2014.

[52] Adriano Moreira, Maria João Nicolau, Filipe Meneses, and António Costa. Wi-Fi Fingerprinting in the Real World - RTLS@UM at the EvAAL Competition. In *Proceedings of the Sixth International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2015.

[53] Klara Nahrstedt and Long Vu. CRAWDAD dataset uiuc/uim (v. 2012-01-24). Downloaded from http://crawdad.org/uiuc/uim/20120124, January 2012.

[54] Ana Régia de M. Neves, Álvaro Marcos G. Carvalho, and Célia G. Ralha. Agent-based architecture for context-aware and personalized event recommendation. *Expert Systems with Applications*, 41(2):563 – 573, 2014.

[55] José A. Paredes, Teodoro Aguilera, Fernando J. Álvarez, Jesús Lozano, and Jorge Morera. Analysis of Doppler effect on the pulse compression of different codes emitted by an ultrasonic LPS. *Sensors*, 11(11):10765–10784, 2011.

[56] Aveek Purohit, Shijia Pan, Kaifei Chen, Zheng Sun, and Pei Zhang. CRAWDAD dataset cmu/supermarket (v. 2014-05-27). Downloaded from http://crawdad.org/cmu/supermarket/20140527, May 2014.

[57] Injong Rhee, Minsu Shin, Seongik Hong, Kyunghan Lee, Seongjoon Kim, and Song Chong. CRAWDAD dataset ncsu/mobilitymodels (v. 2009-07-23). Downloaded from http://crawdad.org/ncsu/mobilitymodels/20090723, July 2009.

[58] Dario Salvi, Paolo Barsocchi, MariaTeresa Arredondo, and JuanPabloLÃázaro Ramos. Evaal, evaluating aal systems through competitive benchmarking, the experience of the 1st competition. In *Evaluating AAL Systems Through Competitive Benchmarking. Indoor Localization and Tracking*, volume 309 of *Communications in Computer and Information Science*, pages 14–25. Springer Berlin Heidelberg, 2012.

[59] Vladimir Savic and Erik G. Larsson. Fingerprinting-Based Positioning in Distributed Massive MIMO Systems. In *Proc. of IEEE 82nd Vehicular Technology Conference (VTC2015-Fall), Sept. 2015.*, 2015.

[60] O.E. Segou, S.A. Mitilineos, and S.C.A. Thomopoulos. Dale: A range-free, adaptive indoor localization method enhanced by limited fingerprinting. In *Proceedings of the 1st the International Conference on Indoor Positioning and Indoor Navigation (IPIN'2010)*, pages 1–8, 2010.

[61] Joaquin Torres-Sospedra. *Ensembles of Artificial Neural Networks: Analysis and Development of Design Methods*. PhD thesis, Department of Computer Science and Engineering, Universitat Jaume I, 2011.

[62] Joaquín Torres-Sospedra, Joan Avariento, David Rambla, Raúl Montoliu, Sven Casteleyn, Mauri Benedito-Bordonau, Michael Gould, and Joaquín Huerta. Enhancing integrated indoor/outdoor mobility in a smart campus. *International Journal of Geographical Information Science*, 29(11):1955–1968, 2015.

[63] Joaquín Torres-Sospedra, Raul Montoliu, Adolfo Martínez-Usó, Joan P. Avariento, Tomás J. Arnau, Mauri Benedito-Bordonau, and Joaquín Huerta. UJIIndoorLoc: A new multi-building and multi-floor database for WLAN fingerprint-based indoor localization problems. In *Proceedings of the Fifth Conference on Indoor Positioning and Indoor Navigation*, pages 261–270, 2014.

[64] Joaquín Torres-Sospedra, Raul Montoliu, Sergio Trilles, Óscar Belmonte, and Joaquín Huerta. Comprehensive analysis of distance and similarity measures for Wi-Fi fingerprinting indoor positioning systems. *Expert Systems with Applications*, 42(23):9263 – 9278, 2015.

[65] Joaquín Torres-Sospedra, David Rambla, Raul Montoliu, Oscar Belmonte, and Joaquín Huerta. UJIIndoorLoc-Mag: A New Database for Magnetic Field-Based Localization Problems. In *Proceedings of the Sixth Conference on Indoor Positioning and Indoor Navigation*, 2015.

[66] Kagan Tumer and Joydeep Ghosh. Analysis of decision boundaries in linearly combined neural classifiers. *Pattern Recognition*, 29(2):341–348, 1996.

[67] Kagan Tumer and Joydeep Ghosh. Error correlation and error reduction in ensemble classifiers. *Connection Science*, 8(3-4):385–403, 1996.

[68] Maarten Weyn. *Opportunistic Seamless Localization*. PhD thesis, University of Antwerp, 2011.

[69] Jaegeol Yim. Introducing a decision tree-based indoor positioning technique. *Expert Systems with Applications*, 34(2):1296 – 1302, 2008.

[70] Feng Yu, Minghua Jiang, Jing Liang, Xiao Qin, Ming Hu, Tao Peng, and Xinrong Hu. 5g wifi signal-based indoor localization system using cluster k-nearest neighbor algorithm. *International Journal of Distributed Sensor Networks*, 2014.

[71] Mu Zhou, Zengshan Tian, Kunjie Xu, Xiang Yu, Xia Hong, and Haibo Wu. SCaNME: Location tracking system in large-scale campus wi-fi environment using unlabeled mobility map. *Expert Systems with Applications*, 41(7):3429 – 3443, 2014.

[72] Mu Zhou, Qiao Zhang, Zengshan Tian, Feng Qiu, and Qi Wu. Integrated location fingerprinting and physical neighborhood for wlan probabilistic localization. In *Computing, Communication and Networking Technologies (ICCCNT), 2014 International Conference on*, pages 1–5, July 2014.