2015 International Conference on Indoor Positioning and Indoor Navigation (IPIN), 13-16 October 2015, Banff, Albeta, Canada

Evaluating Indoor Localization Solutions in Large Environments Through Competitive Benchmarking: The EvAAL-ETRI Competition

Francesco Potortì*, Paolo Barsocchi*, Michele Girolami*†, Joaquín Torres-Sospedra‡ and Raúl Montoliu‡
*CNR-ISTI, Via G. Moruzzi 1, 56124, Pisa, Italy
†Department of Computer Science, Largo B. Pontecorvo 3, 56127, Pisa, Italy
‡Institute of New Imaging Technologies, Universitat Jaume I, Avda Vicente Sos Baynat S/N, Castellón, Spain
*Email: {Potorti, paolo.barsocchi}@isti.cnr.it - ‡Email: {jtorres, montoliu}@uji.es

***Abstract*—The increasing demand for services and higher comfort levels inside buildings, together with the rise in time spent indoor, ensure an upward trend in indoor localization demand for the future. Evaluation of indoor localization systems is particularly challenging due to the complexity of such systems and to the variety of solutions adopted and services offered. EvAAL is an international competition aimed at evaluating and assessing indoor localization systems. The fifth edition of EvAAL promotes competitions on indoor localization in large environments. This paper describes its technical aspects, the competing systems and the results.**

I. INTRODUCTION

As people spend most of their time inside buildings, we expect an upward trend in demand of Indoor Localization (IL) services, especially given the increasing demand for services and higher comfort levels inside buildings.

Localization of devices and people has been recognized as one of the main building blocks for indoor services [1], [2], [3], because the user position can be used for detecting the user's activities and provide services based on them. While in outdoor scenarios the Global Positioning System (GPS) constitutes a reliable and easily available technology, for indoor scenarios GPS is largely unavailable. For this reason, several systems have been proposed for IL, but no winner technology has still emerged: ILs are still in their infancy, and a matter of research. As such, evaluating IL solutions is of paramount importance, both from a research and a commercial point of view. To ensure validity and usability of IL systems researchers must reach consensus on a set of standard evaluation method, otherwise the scientific advantages on the state of art will remain unclear.

Driven by this objective, and building on the experience of the first three EvAAL competitions in small environments [4], we organized, in conjunction with the IPIN conference on Indoor Positioning and Indoor Navigation, an annual international competition for IL systems in large environments. The Korean research institute ETRI was the main sponsor to the competition, with two 2000 USD cash prizes for the winners of the on-site tracks and two paid invitations to hold a seminar in Korea during 2016. The winner of the off-site track was awarded with a 500 USD worth prize provided by UJI-INIT. The competition aims at creating an environment where researchers, students and industries can compare their solutions and exchange ideas, and where the comparison of IL systems may become feasible and standardised.

The main objective of this paper is to describe the comparison criteria and the results from the fifth (2015) edition of EvAAL-ETRI.

The idea of initiating EvAAL was inspired by successful competitions such as the Trading Agent Competition [5] and DARPA Grand Challenge [6]. Beyond supporting the growth of the IL community, the main technical objectives of the competitions organized by EvAAL are to:

- enable the comparison of different IL solutions
- experiment with benchmarking and evaluation methods
- identify relevant IL problems, requirements and issues
- identify new, original solutions for IL and software tools for the evaluation of IL systems

EVAAL aims at enabling the comparison of different IL solutions, by establishing suitable benchmarks and evaluation metrics that will be progressively refined and improved with time. EvAAL's objective is to fill the gap between research and practical applications by tackling the evaluation issue and offering researchers an arena to try, test and experiment not only IL solutions but also benchmarks and evaluation methods. In the long term, EvAAL will develop evaluation methodologies, criteria and tools (including software, benchmarks etc.) from which communities interested in IL can harvest. Making these techniques open, available, and easy to use will enable comparative evaluation between similar components across systems and, in the end, of whole IL systems.

The rest of this paper has been organized as follows. Section II describes previous editions of the EvAAL competition, the IPIN competition and the Microsoft Indoor Localization Competition. Sections III and introduce the on-site and off-site tracks of the the fifth (2015) edition of EvAAL-ETRI competition. Section V describes the evaluation criteria for the competition and shows the competition results. Finally, some conclusions are given in Section VI.

II. EXISTING INDOOR LOCALIZATION COMPETITIONS

EvAAL was the first international competition aimed at comparing indoor localization systems, with its first edition organized in 2011. Three years later, in 2014, two new competitions were born: the IPIN competition for big environments, which built on EvAAL's experience as far as the comparison criteria were concerned, and the IPSN Microsoft competition, with a more easy-going stance.

A. *EvAAL competition's first three editions*

The EvAAL competition[1] aims at establishing benchmarks and evaluation metrics for comparing Ambient Assisted Living solutions with particular attention to indoor positioning systems. International competitions on specific aspects of AAL systems have been organized since 2011, with the long-term goal of evaluating complete AAL solutions.

EvAAL aimed from the start at rigorous, well-defined measurement criteria in a realistic setting, and publication of as much data as possible so that researchers could benefit from the experiments carried during the competition. The actor carrying the competing device moves in a realistic way, equal for all competitors, following a precisely defined, step-by-step path aiming at a reproducibility in the order of 20 cm in space and 250 ms in time [4]. In order to allow for a precise and reliable real-time measurement of the accuracy performance, competing systems were required to integrate with the organizers' system through the universAAL middleware [7].

In the first three editions, all technologies were allowed, with the only limits being one hour's time to install the system inside the living lab area, and one hour to complete the required task. Given the purpose of the competition, which focused on AAL systems, and the need to compare vastly different technologies ranging from Wi-Fi to Zigbee, ultrasound, IR, magnetic and more, accuracy was not the only performance measure; in fact, it accounted for less than half the final score [4]. Other performance criteria were installation time, system reliability measured as the number of estimation samples produced (2 Hz was the target rate), and other subjective measures like use of standard protocols and libraries, integrability with other systems, ease of use when carried, ease of installation in a normal domestic environment, maintainability and so on [4].

The accuracy was computed as the third quartile of the error [4], a significant and robust statistic.

1) 2011 edition: The first EvAAL competition [8] was focused on indoor localization and tracking and its main objective was to enable the comparison of different localization solutions, by establishing suitable benchmarks and evaluation metrics. Seven competitors demonstrated their systems at the CIAmI Living Lab in Valencia, Spain in July 2011 [9], [10], [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 competition was not a public event, but all results were made public. Eight Evaluation Committee members plus four staff members were present during the two days and half of the competition, to gather all the information that was going to be used to compute the final scores. The first EvAAL competition officially closed at the AAL forum in Lecce, Italy, in September 2011. The forum 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) 2012 edition: The aim of the second edition [11] was to award the best indoor localization system from the point of view of Ambient Assisted Living (AAL) applications [12], [13]. The automatic and unobtrusive identification of users activities was considered one of the challenging goals of context-aware computing. Real-time monitoring of human movements can be a useful tool for many purposes and future applications such as lifelog, healthcare or entertainment. The second EvAAL competition was organized in two tracks. The first track 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 focused on Activity Recognition for AAL, and it was held the next week at the CIAmI Living Lab in Valencia, Spain. Similarly to the previous edition, Evaluation Committee members were present during the two days and half of both competitions, and the official end was during the AAL forum in Eindhoven, Netherlands, in September 2012 [14].

3) 2013 edition: The third edition [15] had the same formula as the previous one, and is described in detail in [16]. This edition also included a demo on Companion Robots for AAL, held on July 2013 at the Peccioli Living Lab in Pisa, Italy. The third EvAAL competition officially closed at the AAL forum in Norrköping, Sweden, in September 2013.

B. *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 on 27th to 29th October 2014 [17]. The competition consisted of two tracks: Smartphone Based Positioning and Foot-mounted Pedestrian Dead Reckoning Positioning.

The IPIN competition inherited the basic criteria from the first three EvAAL competitions: a rigorous and well-defined measurement method, a realistic setting, and real-time measurements. The main difference was that only very specific technologies were allowed, and as a consequence the scoring criterion was based on accuracy only.

The purpose of the IPIN 2014 competition was to assess and measure the ability of competing systems to accurately identify their position inside a large, public indoor area. Competing systems had to be engineered or prototype systems intended to be carried by an actor without impairing her or his movements. Each competing system was carried by an actor and continuously communicated real-time estimates of its position to a measurement app provided by IPIN.

---

[1] http://evaal.aaloa.org

Competitors could use any sensor available on the smartphones used. No instrumentation of the area by competitors was allowed. Competitors were able to survey the competition area by themselves during the day preceding the competition. In order to help integrate their solution with the measurement app provided by IPIN, the IPIN team assisted the competitors and a competitor's integration package was delivered in advance to them.

C. The Microsoft Indoor Localization Competition

The annual Microsoft Competition is held in conjunction with the International Conference on Information Processing in Sensor Networks (IPSN), aiming to bring together real-time or near real-time indoor location technologies and compare their performance in the same space since 2014.

The ideas behind this competition are quite different from EvAAL's. The aim is to lower as much as possible the barriers for competitors, by allowing any kind of technology in a small-to-medium environment. Measurements are done with the competitors standing at a set of predefined points, with no attempt at realistic real-time measurements, and ranking of results is done using the average error, putting together systems with vastly different technologies.

1) 2014 edition: The first edition took place at the venue of the 2014 IPSN conference in Berlin, Germany [18], [19]. The evaluation scenario consisted of two 90 m^2 attached rooms and a hallway and the competition lasted two days. During the first day, competitors had 7 hours to set up their indoor localization systems and deploy their custom hardware (up to 10 items per team). Competitor systems were divided into two main categories: Infrastructure-based and Infrastructure-less technologies. Since most of the technologies were based on Wi-Fi, the organizers deployed 10 Wi-Fi Access Points (WAPs) in the evaluation area to be used in the localization algorithms. The use of other generic WAPs, such as the Hotel's Internet connectivity ones, was not allowed for localization purposes. In case of the necessity of deploying specialized WAPs, competitors were allowed to deploy their own WAPs under request to organizers. On the second day, each competitor had pre-assigned a time-slot for evaluation. The organizers carried the device (phone, tablet or laptop) above 20 evaluation points whose positions were disclosed the day before. The evaluation criteria adopted was based on the mean error, computed as the Euclidean distance between estimated and current positions over the 20 testing points.

2) 2015 edition: The second edition was held at IPSN 2015 in Seattle, USA [20]. The evaluation scenario was on the third floor of the venue and it consisted of one exhibition room and an open challenging area, covering an area of 1250 m^2. The competition lasted two days and the competing systems were again divided into the two categories. In the first day, competitors had at least 5 hours to set up their indoor localization systems and deploy their custom hardware (up to 10 anchor points per team) simultaneously. The organizers also considered the possibility of small time windows about 10-15 minutes where competitors could set up their system without interferences from other systems. Competitors leveraging on Wi-Fi in the infrastructure-less category had to use the 8 WAPs already deployed in the conference venue for Internet connectivity. However, some particular conditions were applied for those competitors using custom Wi-Fi in the Infrastructure-free category. On the second day, each competitor had a pre-assigned time-slot for evaluation. The organizers carried the device (phone, tablet or laptop) above 20 evaluation points, that were unknown to the competitors the day before. Averaging and smoothing techniques were explicitly allowed. The main metric and the procedure to establish the winner remained the same as in 2014.

3) 2016 edition: The third edition will be held by IPSN 2016 in Vienna, Austria and it was officially announced on October 2015 [21]. Some changes with respect to previous editions have been introduced. Firstly, the upcoming edition will divide systems into three new main categories: (i) Commercial off-the-shelf Technologies, (ii) Commercial off-the-shelf Technologies with initialization and (iii) 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.

III. THE ON-SITE TRACKS

The purpose of the on-site tracks is to assess and measure the ability of competing systems to accurately identify their position inside a large, public indoor area. To this purpose we proposed two competition tracks:

- Track 1: Smartphone-based;
- Track 2: Foot-mounted Pedestrian Dead Reckoning (PDR).

For both tracks we requested that the competing systems have to be engineered or prototyped so that they could be carried by an actor without impairing her movements.

Concerning Track 1 we allowed to use only one commercial smartphone. The participant had the possibility of using any sensor available on the phone, such as GPS, accelerometer, magnetometer and barometer sensors.

For what concerns Track 2 each participant implemented a localization system based on MEMS sensors (such as inertial, compass and pressure sensors) that must be mounted no higher than the ankle articulations. We did not limit the number of devices to use.

Competing teams could not install any kind of instrumentation on the area, only the existing Wi-Fi access points already installed could be used during the competition. Furthermore, competitors were allowed to survey the area the day before the competition day. We disclosed the competing path half an hour before the start of the competition.

We provided the competitors with a detailed geo-referenced map of the area composed by 2 buildings (see Section III-C) connected by a secondary road.

During the competition a number of volunteers played the role of actor. The actor had the role of testing the applications of the competitors following a pre-defined path. The actor was not trained in advance to use the applications, this ensured us to obtain performance measurements not biased from the actor itself. This aspect is particularly critical for PDR systems. In fact the way the actors move and specifically their gait affects the accuracy of the estimated position especially while the actor steps over the stairs.

Each competing team was scheduled for performing a number of test runs during the competition day. During each run the actor tested one competing application by means of the StepLogger application (see Section III-B). StepLogger recorded continuously performance of the application in order to compute the final score of the team.

A. *Reference localization system*

A reference localization system is essential to measure the accuracy of the competing applications. More precisely, the accuracy is defined as a statistic associated to the distance between the real position of the user and the estimated position of the application.

To this purpose, the reference localization system was composed by a number of markers stuck on the floor with predefined coordinates, as shown in figure 1. The actor had to follow the markers sequentially and to step over each of them in a natural way. The synchronization between the actor positions and the estimated position was guaranteed by the StepLogger application that logged the time when the actor stepped over a marker as well as the estimated position of the application with a temporal marker.

B. *StepLogger: the evaluation system*

We developed a software package for the competitors, which included the StepLogger and StepLoggerClient applications and the developer handbook.

StepLogger is a simple Android-based application with a minimalist graphical interface as shown in figure 2. The interface is designed to show the sequence of markers in the right order that the actor must follow during each run. We asked the actors to press the button when stepping over the marker with the same label shown by StepLogger.

StepLogger provides two logging functionalities:

- log the button pressures;
- log the estimated positions.

The first log is generated as soon as the actor presses the button corresponding to the marker on the floor. StepLogger logs the following information: `[timestamp, markerID]`, where `timestamp` is in millisecond from the Unix epoch and `markerID` is the button label.

The second log records the estimated position of the competing applications. StepLogger logs the following information each time the application estimates the position: `[timestamp, x, y, z]` where `x,y` are respectively longitude, latitude (expressed in WGS84 reference system) and `z` floor number of the estimated position.

We asked competitors to integrate their applications with StepLogger using AIDL formalism (Android Interface Definition Language). In particular, StepLogger implements the `IStepLoggerInterface` defining only one method: `logPosition(long timestamp, double x, double y, double z);`. The competing applications have firstly to fetch the implementation of `IStepLoggerInterface` and, then to invoke the `logPosition(···)` method every time a new position is estimated.

Fig. 1. The reference localization system: when the actor steps over the green mark, he presses the button of the logging application (section III-B).

The software package also included the StepLoggerClient application demonstrating how to integrate with StepLogger and a handbook for developers. The StepLoggerClient implements a simple localization application that, with a frequency of 2 Hz, invokes the `logPosition(···)` method. StepLoggerClient shows a simple graphical interface with two buttons: start/stop the localization application, as shown in figure 2. The software is available at http://evaal.aaloa.org.

C. *The chosen path*

One of the distinguishing features of the EvAAL competition is the challenging path. The path has been defined with the goal of reproducing realistically the way how people move

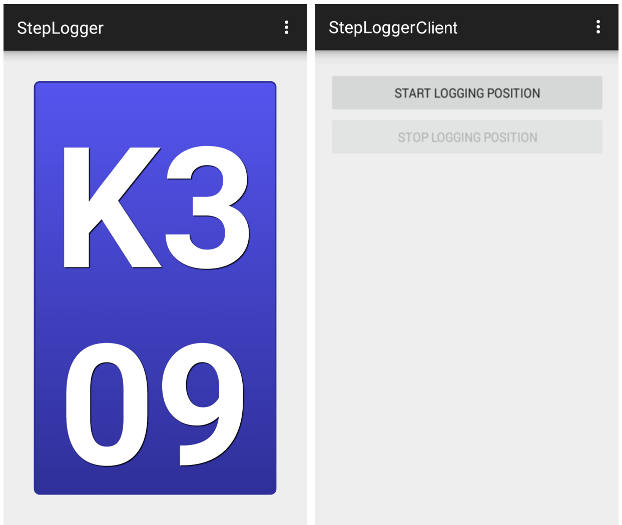


Fig. 2. StepLogger and StepLoggerClient applications.

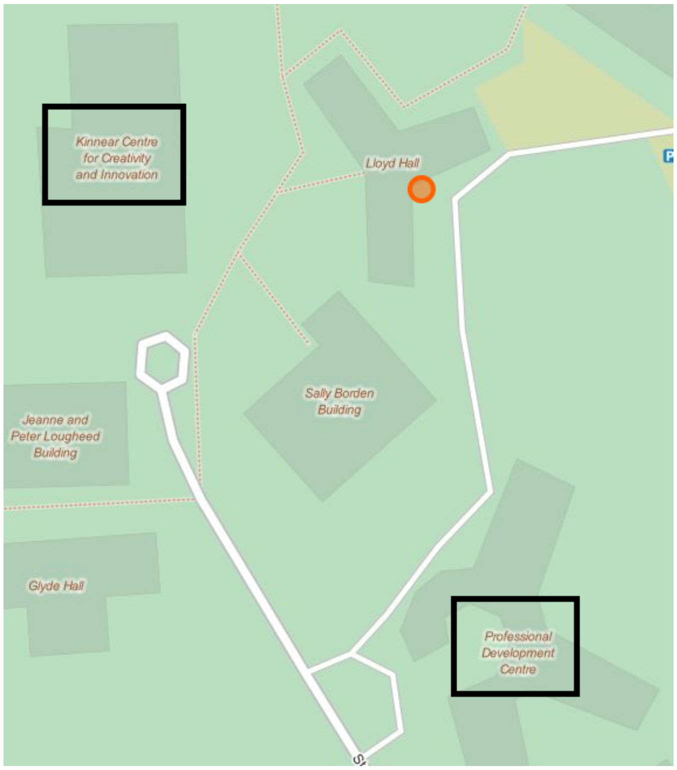


Fig. 3. Map of the Banff Centre with building outlines. The KCCI and PDC buildings are used for the on-site tracks.

within a big indoor environment. To this purpose we applied the following rules:

- the path comprises two buildings connected through a secondary road;
- inside each building, the path traverses 3 floors;
- floors are traversed both with stairs and with a lift;
- along the path the actor has to rest for few seconds in 3 locations in order to reproduce a natural behavior of humans while moving;
- the actor moves with a typical pedestrian speed ranging from 1 m/s to 2 m/s;
- the time needed to complete the path exceeds 20 minutes in order to stress enough the competing applications.

We surveyed the Banff Centre[2] and we selected the KCCI and PDC buildings shown in figure 3 with black boxes.

The KCCI building is composed of three floors connected both with stairs and with a lift. Floors have a similar layout, in particular a floor is composed by a long corridor about 40 m long and 6 m wide. On the left side and right side of the corridor a number of doors provide access to medium sized conference rooms. The KCCI structure is mostly composed by concrete with around 15 pillars holding up the whole structure. The upper and lower corners of each floor had wide glass walls.

The PDC hosts one of the hotels inside the Banff Centre. The PDC shape is similar for the three floors: a wide round open space gives access to three wings where most of guest's rooms are located. The PDC is a wooden structure, also in this case a number of pillars hold up the building.

Wi-Fi access was available both inside KCCI and PDC buildings, while along the small road connecting the two buildings the network coverage was only partial.

The path consisted of 62 markers. The path started from KCCI level 3, down to level 2 and down to level 1 by means of stairs, then the actor moved to the PDC building through a short outdoor path of about 200 meters. Inside PDC the actor had to step over the markers starting from level 1, up to level 2 and to level 3 by using the lift. Of 62 markers, 2 were placed outdoors and 42 in KCCI (14 markers for each floor), the remaining 18 markers were placed on PDC, in particular 13 on PDC level 1, 3 markers on level 2 and 2 markers on level 3. Figure 4 shows the competition path, outdoor markers are not shown.

The positioning of the marked followed these criteria:

- we labeled markers with a tag in this form: `[buildingID,floor,markerID]`, where `buildingID` is `K` for KCCI and `P` of PDC, the `floor` ranges from 1 to 3 and the `markerID` ranges from 1 to 14 in KCCI and from 1 to 13 in PDC (see Figure 1);
- we placed markers in easily accessible places where people usually step over;
- we placed markers with a distance ranging from 3 to 5 meters appart;
- we tried to make the upcoming marker always visible from the position of the previous one;
- when possible, we placed the markers on the floor. At level 3 and 2 of KCCI the floor was carpeted, so we stuck the markers on the walls.

---

[2] coordinates: $(Long : 51.172, Lat : -115.562)$

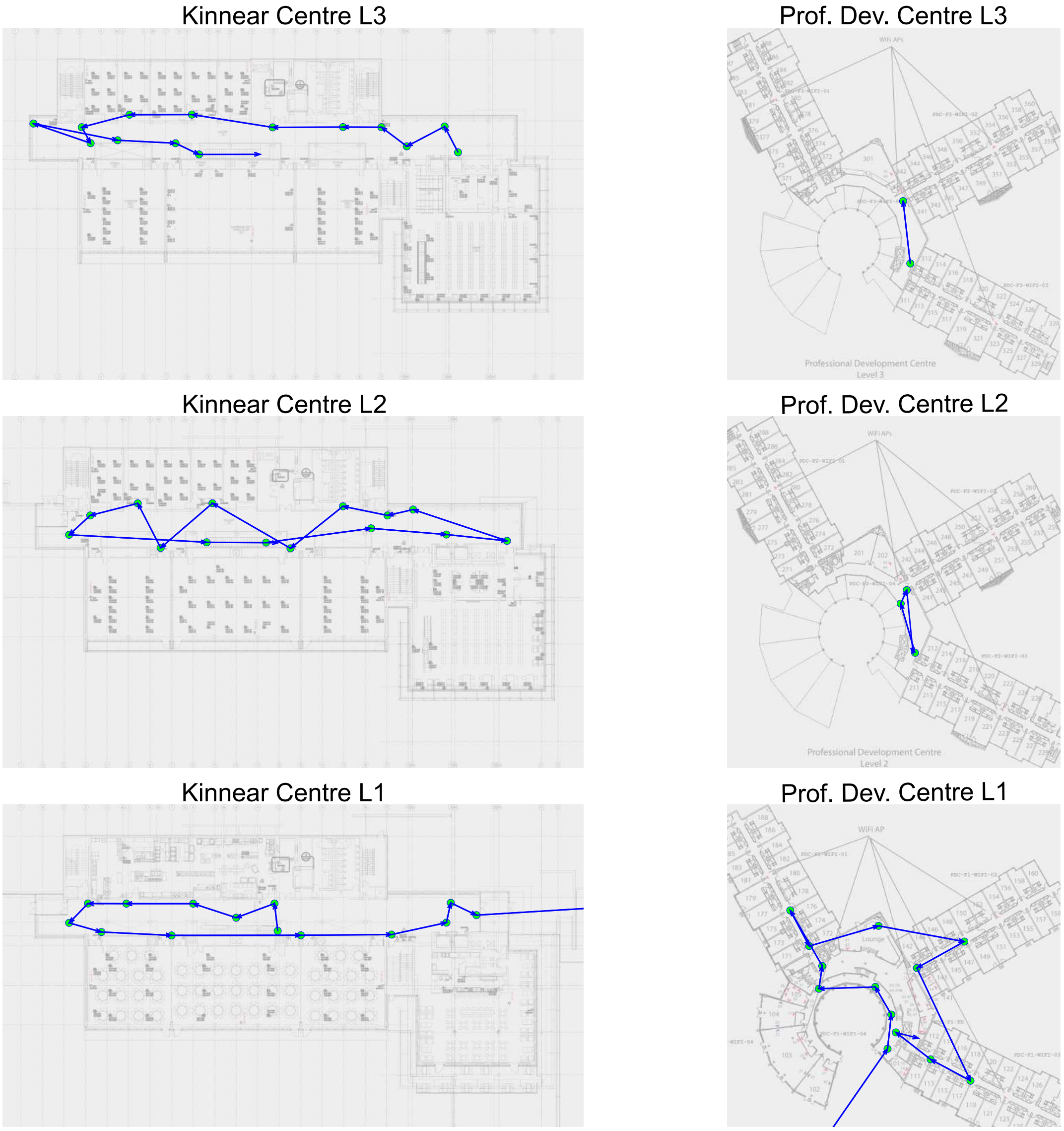


Fig. 4. The EvAAL competition path for the on-site tracks. The path starts on the third floor of KCCI building (top left map). It ends on the first floor of KCCI for track 2. For Track 1 it continues outdoor traversing two marks, not depicted here, and ends on the third floor of PDC.

IV. THE OFF-SITE TRACK

The purpose of the off-site track is to assess and measure the ability of competing systems to accurately identify their position inside a large, public indoor area by using Wi-Fi fingerprinting. Participants to the off-site track received the *UJIIndoorLoc* database [22], [23] and we asked them to implement and set up their localization systems. A web based viewer[3] was developed to show the training and validation sample locations.

For the evaluation purpose, we provided a set of private samples without labels (ground-truth locations) to competitors for testing. This data set was not publicly available to guarantee that all competitors had the same time to optimize and validate their indoor localization systems.

Competitors applied their locations algorithms to the private final test data and they sent us their predicted locations. We accepted up to five different independent results for each participant. Only the best performing participant's alternative was considered to compute the final score. The evaluation was performed off-line before the conference.

A. The UJIIndoorLoc Database

Organizers of Track 3 made the *UJIIndoorLoc* database available to the scientific community by publishing it in the UCI Machine Learning Repository. This is a Wi-Fi fingerprint database that contains well-differentiated Wi-Fi samples for training and validation/test.

---

[3] http://indoorloc.uji.es/webviewer/

TABLE I

CORRESPONDENCE BETWEEN *PhoneID* AND REAL DEVICE. REAL DEVICE'S INFORMATION INCLUDES THE MODEL DESCRIPTION AND ANDROID VERSION.

PhoneID	Android Device	Android Version
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

Each database sample is directly related to a Wi-Fi fingerprint capture and it contains the following 529 numeric elements:

`001-520`: RSSI levels
`521-523`: Real world coordinates of the sample points
`524`: BuildingID
`525`: SpaceID
`526`: Relative position with respect to SpaceID
`527`: UserID
`528`: PhoneID
`529`: Timestamp

The main features of the *UJIIndoorLoc* database are:

- it covers a surface of 108703 m^2 including 3 buildings with 4 to 5 floors;
- the number of reference points for training is 933;
- 21049 samples (individual fingerprints) were captured: 19938 for training/learning and 1111 for validation/testing;
- independence was assured by taking the validation samples 4 months after training ones;
- the database collects information gathered from 520 different access-points;
- data were collected by more than 20 users using 25 different models of mobile devices, some users used more than one model;
- samples where timestamped.

The full description of this public database, the indoor scenario and how the fingerprints were collected can be found in [22], [23], [24].

B. *The final data set*

Since the *UJIIndoorLoc* database is publicly available, we provided the participants with a private final test set with the same structure as *UJIIndoorLoc*, but information about location and users (fields 521 to 527) was removed. This new set is not public and it is composed by 5179 new fingerprints that were collected in different places of the three buildings between 29 November 2013 and 31 March 2015 with seven different devices by 6 people. Five of the devices were new and they had not been used in the public *UJIIndoorLoc* database, and three of these people had not taken any measurements for the public UJIIndoorLoc database. The intention of this test set was to be as realistic as possible. So, *new* users and *new* devices were considered.

TABLE II

A FLOOR PENALTY IS ADDED TO THE x, y EUCLIDEAN DISTANCE FOR EACH WRONG FLOOR ESTIMATE. IN TRACK 3 ONLY, A BUILDING PENALTY IS ADDED IF THE ESTIMATED BUILDING IS WRONG.

Track	Floor penalty	Building penalty
Track 1	15 m	0
Track 2	15 m	0
Track 3	4 m	50 m

We asked the competitors to perform up to five different estimations, to collect results in CSV files and to send back the data. Each CSV file had the same number of lines as samples included in the final test set, so the location provided in the i$^{\text{th}}$ line corresponded to the i$^{\text{th}}$ fingerprint of the private test set. Each result line had the following format: `longitude,latitude,FloorID,BuildingID`

V. SCORING CRITERIA AND RESULTS

The evaluation criteria are common to all tracks. We discuss the reasons behind the chosen criteria and present the competitors and the results obtained for those who completed the competition.

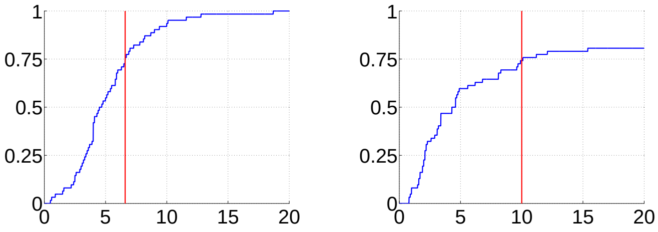
A. *Evaluation criteria*

Each track has a strict definition, chosen in such a way that competing systems are homogeneous and directly comparable. For this reason, the only performance criterion used is *accuracy*, without any attempt at considering other performance measures such as cost, power consumption, ease of use, scalability, ease of maintenance and so on.

Indeed, accuracy is the classical measurement of the goodness of a localization systems, based on samples of the distance between the point where the system locates the user and the point where the user really is. Accuracy is computed by reading the data produced by the competing systems and comparing it with reference data (the ground-truth).

We define the *error* as the Euclidean distance between the *real position* where the actor presses the button (over the markers on the floor) and the last *estimated position* produced by the competing system before the button is pressed. Distances are computed in two dimensions, penalties are added for floor error and building error (see table V-A).

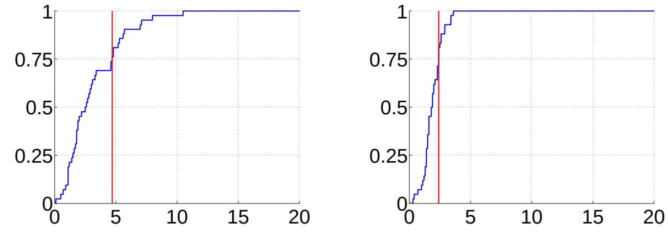
In order to rank the competing systems, the error series should be reduced to a scalar score, and the literature is rich in methods to reach this result. In [4] 195 papers from the first edition of the Indoor Positioning and Indoor Navigation (IPIN 2010) Conference are analyzed and the criteria used for describing system performance are discussed and compared. The metrics taken into account in these works range from path comparison or error CDF comparison to statistics such as mean, a quantile or error variance. We are aiming at the same objectives as [4], that is removing subjectivity and using a robust statistic while considering that we are mostly dealing with experimental systems: we follow the same reasoning, reach the same conclusion and we adopt the third quartile as our score to rank the accuracies of the competing systems.



(a) MMSS team.

(b) SAMS team.

Fig. 5. Results of MMMS and SAMS teams.



(a) With untrained actor.

(b) With trained actor.

Fig. 6. Results of the NESL team.

In summary:

- each error is the x, y Euclidean distance from the estimated position to the real position of the marker, with the addition of:
  - a penalty in case the building is not correctly identified;
  - a floor penalty that is proportional to the difference between estimated floor number and real floor number;
- the score is computed as the third quartile of errors at marks.

B. The on-site tracks

The fifth edition of EvAAL had 6 competing teams in the on-site tracks: 4 teams in Track 1 and 2 teams in Track 2. We briefly describe the solutions proposed only for those teams that have successfully completed at least one run of the path. We also report the best results obtained by each team among all the completed runs.

1) Smartphone-based track: The **MMSS** team was formed of researchers from University of Calgary (Canada) and from Wuhan University (China). They proposed an indoor navigation system that uses multiple kinds of sensors and technologies: inertial sensors, magnetic sensors, barometer and Wi-Fi [25]. The heading from the attitude-determination module is fed into the PDR-based position-tracking module. Then, PDR is used for providing continuous position estimates and for the blunder detection of both Wi-Fi fingerprinting and magnetic matching. Meanwhile, Wi-Fi fingerprinting utilizes a point-by-point matching technology, while magnetic matching is based on profile-matching. Finally, Wi-Fi and magnetic matching results are passed into the position-tracking module as updates to a Kalman filter. The distribution of error obtained by the MMSS team is shown in figure 5(a).

The **SAMS** team is composed of researchers from Samsung R&D (Poland). Their system relies on information provided by inertial sensors, barometer (not used during this competition), Wi-Fi and BLE (not used during this competition) [26]. Information gathered from sensors are fused using a particle filter. The floor plan is used to refine and smooth the walked paths. The distribution of the error obtained by the SAMS team is shown in figure 5(b).

2) Foot-mounted PDR track: The **NESL** team was composed of researchers from Seoul National University (South Korea). Their system is a purely PDR-based solution, as required by the rules of Track 2. In order to reduce the heading error, they assume that walls and corridors are straight and either parallel or orthogonal to each other in usual building [27]. The NESL team refers to the typical directions of walls and corridors as the dominant directions or cardinal directions. This method has limitations when the pedestrian is not walking along the corridors for a long period. In these cases, azimuth errors arise by only matching with the dominant directions. To overcome these limitations, the system implements the INS-EKF-ZUPT (IEZ) based Advanced Heuristic Drift Elimination (AHDE) which helps removing azimuth drift error in indoor environments using an extended Kalman filter. Results concerning the distribution of the error are reported in figure V-B2. In particular, figure 6(a) shows the distribution of the error when the actor walked along the path, while figure 6(b) shows the much better results obtained with a trained actor, that is, one of the system's developers.

Table V-B2 summarizes the results of the on-site tracks.

TABLE III
RESULTS FOR TRACKS 1 AND 2.

Track	Team	Median	Mean	RMS	3^{th} quartile
1	MMSN	4.6	5.3	3.1	6.6
1	SAMS	4.4	9.3	11.4	10
2	NESL (untrained actor)	2.5	3.1	2.2	4.7
2	NESL (trained actor)	1.8	1.9	0.7	2.4

C. The off-site track

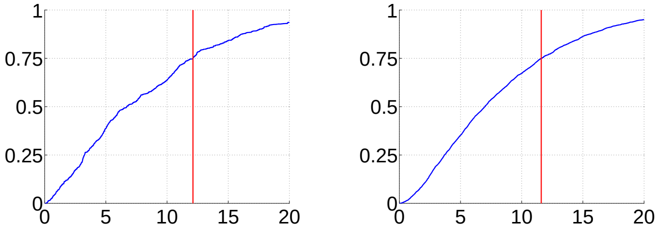
The off-site track had 4 competing teams. Below we briefly describe the Wi-Fi fingerprinting solution proposed as well as the best results obtained from the teams among five different results submitted.

The **MOSAIC** team was composed of researchers from the University of Antwerp, Belgium [28]. Their system quantified the localization performance of exteroceptive sensors solely by virtue of their sensor model using information theory. For the Wi-Fi signals they used a probabilistic version of the sensor model used in [29] and they defined four mutually exclusive events when comparing the RSS values: *hit*, *miss*, *extra* or *none*. Then, they applied a localization algorithm based on k-Nearest Neighbors (kNN) and the Maximum Likelihood Estimator (MLE) that select the location with the highest likelihood. They submitted both alternatives and their best competing system was the one based on MLE. The distribution of the error obtained by the best MOSAIC system is shown in figure 7(a).

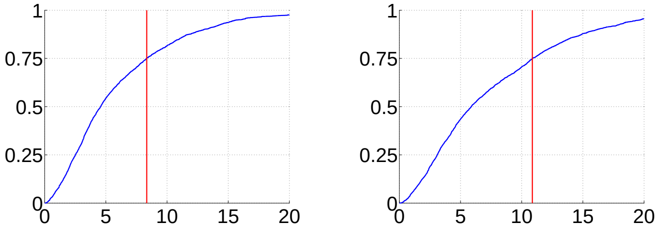
The **HFTS** was composed of researchers from the Stuttgart University of Applied Sciences, Germany [30]. Their system introduced a Wi-Fi fingerprint calibrated weighted centroid localization algorithm. In the first stage, calibration, the virtual positions of the access points are determined using the weighted centroid algorithm. Those virtual positions do not necessarily need to closely match the real positions of the access points. The second stage corresponded to the localization task. They used a weighted centroid and a scalar product fingerprinting to estimate the position where each testing sample was taken. They submitted both alternatives and their best competing system was based on the scalar product fingerprinting algorithm, which reported a high rate on estimating the correct building and floor. The distribution of the error concerning the best MOSAIC system is shown in figure 7(b).

The **RTLS@UM** team was composed of researchers from the University of Minho, Portugal [31]. Their system applied a sequence of filtering and majority rules to a centroid localization algorithm. They used the Real Time Location Service (RTLS) [32] that was deployed at University of Minho. In a first step, the RTLS engine builds a filtered radio map that contains all the fingerprints where the strongest access point is the same as the strongest access point in the testing (operational) fingerprint. Then, this radio map was used to sequentially estimate the building, floor and room by using majority rule of the most similar fingerprints. They computed the geometric position (x, y) as the centroid (kNN-based solution) or the weighted centroid (WkNN solution) for the most similar fingerprints belonging to the estimated building, floor and room. Moreover, they also considered a Predicted K-Nearest Neighbors version on their system (PkNN). They submitted five alternatives and their best solution was based on the first alternative, kNN. The distribution of the error concerning the best RTLS@UM system is shown in figure 7(c).

The **ICSL** team was composed of researchers from the National University of Seoul, Korea [33]. Their system introduced a wireless access point selection to retain useful measurements and machine learning techniques for indoor positioning. First, an additional binary set is created based on the strongest signals. This new set is combined to the existing one to reduce the noise inherent to Wi-Fi fingerprinting. A semi-supervised dimensionality reduction based on Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA) was applied as a base indoor location algorithm. Moreover, they also used a single layer neural network and extreme learning machine as a base indoor location algorithm. For the neural network based approach, they applied k-Nearest Neighbor in the output layer. They submitted five alternatives and their best solution was the one based on EML with access point selection. The distribution of the error concerning the best ICSL system is shown in figure 7(d).



(a) MOSAIC team. (b) HFTS team.



(c) RTLS@UM team. (d) ICSL team.

Fig. 7. Results for Track 3.

Table IV summarizes the results for Track 3. In particular, the system proposed by RTLS@UM team provided the best results.

TABLE IV
RESULTS FOR TRACK 3.

Team	Median	Mean	RMS	3^{th} quartile
MOSAIC	6.72	11.64	31.75	12.12
HFTS	6.99	8.49	10.65	11.60
RTLS@UM	4.57	6.20	8.29	8.34
ICSL	5.88	7.67	9.76	10.87

VI. Conclusions

The 5th EvAAL competition, in conjunction with IPIN 2015, raised a lot of attention during the conference. The reasons were essentially that we showed the results obtained from real working systems in a realistic environment open to scrutiny and with a realistic usage pattern: no simulations, no simplifying assumptions, no small-controlled environments, only real working systems.

Competitors for the on-site tracks were required to provide a working system to an actor who followed a predefined path equal for all competitors and unknown to them until the day of the competition. Competitors only had a single day to set up and adapt their system to the environment, without knowing the path in advance but only the map of the area. This simulates what would happen with the first installation of a generic localization system.

For the on-site smartphone-based track, the path involved two separate three-floor buildings. The actor carrying the system walked along the path with a natural pace, without artificially stopping at every measurement point. The actor used both staircases and elevators for moving between floors, for a total time exceeding 20 minutes of walk. Two out of four competitors were able to complete the path.

For the on-site foot-mounted PDR track, only the first building was used, because the drift when walking from one building to the next (over 200 m) was enough to offset the measurements in the second building by about 20 m. It is worth to notice that results with a trained and an untrained actor were significantly different. Only one competitor out of two was able to complete the path.

For the off-site off-site track, competitors were provided a huge fingerprinting database in a big multi-building, multi-floor environment. All four competitors were able to complete the assigned task.

Next editions will build on the experience gained in 2015 and previous editions, with an eye to maintaining the rigorous testing procedure adopted until now.

REFERENCES

[1] P. Barsocchi, S. Lenzi, S. Chessa, and F. Furfari, “Automatic virtual calibration of range-based indoor localization systems,” *Wireless Communications and Mobile Computing*, vol. 12, pp. 1546–1557, 2012.

[2] F. Palumbo, P. Barsocchi, C. Gallicchio, S. Chessa, and A. Micheli, “Multisensor data fusion for activity recognition based on reservoir computing,” in *Evaluating AAL systems through competitive benchmarking*. Springer Berlin Heidelberg, 2013, pp. 24–35.

[3] P. Barsocchi, “Position recognition to support bedsores prevention,” *Biomedical and Health Informatics, IEEE Journal of*, vol. 17, no. 1, pp. 53–59, 2013.

[4] P. Barsocchi, S. Chessa, F. Furfari, and F. Potortì, “Evaluating AAL solutions through competitive benchmarking: the localization competition,” *IEEE Pervasive Computing Magazine*, vol. 12, no. 4, pp. 72–79, Oct.–Dec. 2013.

[5] “Trading Agent Competition,” 2010. [Online]. Available: http://www.sics.se/tac/

[6] “DARPA Grand Challenge,” 2007. [Online]. Available: http://archive.darpa.mil/grandchallenge

[7] E. Ferro, M. Girolami, D. Salvi, C. Mayer, J. Gorman, A. Grguric, R. Ram, R. Sadat, K. M. Giannoutakis, and C. Stocklöw, “The universaal platform for aal (ambient assisted living),” *Journal of Intelligent Systems*, 2015.

[8] “The first EvAAL competition,” 2011, accessed: November 2015. [Online]. Available: http://evaal.aaloa.org/2011-competition/2011

[9] P. Barsocchi, F. Potortì, F. Furfari, and A. M. M. Gil, “Comparing AAL indoor localization systems,” in *Evaluating AAL Systems Through Competitive Benchmarking. Indoor Localization and Tracking*, ser. Communications in Computer and Information Science. Springer Berlin Heidelberg, 2012, vol. 309, pp. 1–13. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-33533-4_1

[10] D. Salvi, P. Barsocchi, M. Arredondo, and J. 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*, ser. Communications in Computer and Information Science. Springer Berlin Heidelberg, 2012, vol. 309, pp. 14–25. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-33533-4_2

[11] “The second EvAAL competition,” 2012, accessed: November 2015. [Online]. Available: http://evaal.aaloa.org/2012/2012-competition

[12] P. Barsocchi, “Evaluating indoor localization systems for aal environments,” in *Evaluating AAL Systems Through Competitive Benchmarking*, ser. Communications in Computer and Information Science. Springer Berlin Heidelberg, 2013, vol. 362, pp. 1–5. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-37419-7_1

[13] “Evaal’12 awards,” *Journal of Ambient Intelligence and Smart Environments*, vol. 4, no. 6, pp. 565–567, 2012.

[14] J. A. Álvarez-García, P. Barsocchi, S. Chessa, and D. 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*, vol. 5, no. 1, pp. 119–132, 01 2013.

[15] “The third EvAAL competition,” 2013, accessed: November 2015. [Online]. Available: http://evaal.aaloa.org/2013/2013-competition

[16] J. A. Botía, J. Antonio, Álvarez-García, K. Fujinami, P. Barsocchi, and T. Riedel, Eds., ser. Communications in Computer and Information Science, vol. 386. Springer Berlin Heidelberg, 2013. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-41043-7

[17] “The indoor positioning and indoor navigation (ipin) competition,” 2014, accessed: November 2015. [Online]. Available: http://www.ipin2014.org/sub/sub17.asp?sub_param=17

[18] D. Lymberopoulos, R. R. Choudhury, X. Yang, and S. Sen, “Microsoft Indoor Localization Competition - IPSN 2014,” accessed: November 2015. [Online]. Available: http://research.microsoft.com/en-us/events/ipsn2014indoorlocalizatinocompetition

[19] D. Lymberopoulos, J. Liu, X. Yang, R. R. Choudhury, V. Handziski, and S. Sen, “A realistic evaluation and comparison of indoor location technologies: Experiences and lessons learned,” in *Proceedings of the 14th International Conference on Information Processing in Sensor Networks*, ser. IPSN ’15, 2015, pp. 178–189.

[20] D. Lymberopoulos, J. Liu, X. Yang, A. Naguib, A. Rowe, N. Trigoni, and N. Moayeri, “Microsoft Indoor Localization Competition - IPSN 2015,” accessed: November 2015. [Online]. Available: http://research.microsoft.com/en-us/events/indoorloccompetition2015

[21] D. Lymberopoulos, J. Liu, Y. Zhang, P. Dutta, X. Yang, and A. Rowe, “Microsoft Indoor Localization Competition - IPSN 2016,” accessed: November 2015. [Online]. Available: http://research.microsoft.com/en-us/events/msindoorloccompetition2016

[22] K. Bache and M. Lichman, “UCI machine learning repository,” 2013. [Online]. Available: http://archive.ics.uci.edu/ml

[23] J. Torres-Sospedra, R. Montoliu, A. Martinez-Uso, J. Avariento, T. Arnau, M. Benedito-Bordonau, and J. Huerta, “UJIIndoorLoc: A new multi-building and multi-floor database for WLAN fingerprint-based indoor localization problems,” in *Proceedings of the Fifth International Conference on Indoor Positioning and Indoor Navigation*, Oct 2014, pp. 261–270.

[24] J. Torres-Sospedra, R. Montoliu, S. Trilles, O. Belmonte, and J. Huerta, “Comprehensive analysis of distance and similarity measures for Wi-Fi fingerprinting indoor positioning systems,” *Expert Systems with Applications*, vol. 42, no. 23, pp. 9263 – 9278, 2015. [Online]. Available: http://dx.doi.org/10.1016/j.eswa.2015.08.013

[25] Y. Li, P. Zhang, Y. Zhuang, H. Lan, X. Niu, and N. El-Sheimy, “Real-time indoor navigation with smartphone sensors,” in *USB On-Site Proceedings of the Sixth International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2015, short paper.

[26] P. Wilk, J. Karciarz, and J. Swiatek, “Indoor positioning engine,” in *USB On-Site Proceedings of the Sixth International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2015, short paper.

[27] H. J. Ju, M. S. Lee, S. Y. Park, Y. Choe, and C. G. Park, “Implementation of advanced heuristic drift elimination for indoor pedestrian navigation,” in *In USB On-Site Proceedings of the Sixth International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2015, short paper.

[28] R. Berkvens, M. Weyn, and H. Peremans, “Localization performance quantification by conditional entropy,” in *Proceedings of the Sixth International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2015, long paper at IEEE Xplore.

[29] R. Berkvens, A. Jacobson, M. Milford, H. Peremans, and M. Weyn, “Biologically inspired slam using wi-fi,” in *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*, Sept 2014, pp. 1804–1811.

[30] S. Knauth, M. Storz, H. Dastageeri, A. Koukofikis, and N. 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, long paper at IEEE Xplore.

[31] A. Moreira, M. J. ao Nicolau, F. Meneses, and A. Costa, “Wi-Fi Fingerprinting in the Real World - RTLSUM at the EvAAL Competition,” in *Proceedings of the Sixth International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2015, long paper at IEEE Xplore.

[32] N. Marques, F. Meneses, and A. Moreira, “Combining similarity functions and majority rules for multi-building, multi-floor, wifi positioning,” in *Indoor Positioning and Indoor Navigation (IPIN), 2012 International Conference on*, Nov 2012.

[33] S. Choi, J. Yoo, and H. J. 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, short paper.