8

Study and Evaluation of Selected RSSI-Based Positioning Algorithms

Stefan Knauth

HFT STUTTGART, UNIVERSITY OF APPLIED SCIENCES, STUTTGART, GERMANY

1 Introduction

This chapter reports on the authors' work on radio signal strength indicator (RSSI) Smartphone indoor positioning and discusses ideas and theory behind it. It documents and extends the authors' presentation at the First International Workshop on Challenges of Fingerprinting in Indoor Positioning and Navigation, 2016 held at Barcelona and the participation at the IPIN EvAAL competition 2015 which was collocated with the IPIN 2015 International Conference on Indoor Positioning and Indoor Navigation.

Smartphone indoor positioning is a substitution for GNSS as typically GNSS signals are to weak to be used inside buildings. A wide span of technologies and schemes are investigated to provide location estimations and tracking inside buildings. Back to 2000, Microsoft RADAR (Bahl and Padmanabhan, 2000) was among the first Wi-Fi RSSI-based positioning experiments. Since then, countless approaches have been proposed and developed for the Wi-Fi RSSI-based positioning problem. Upon the more prominent solutions there are proximity, centroid, fingerprinting, multilateration, and radio tomography. Unfortunately propagation properties in indoor scenarios are quite individual and results of a certain algorithm may vary considerably between different deployments and locations or laboratories.

RSSI is of course not the only possible solution for indoor positioning. There are other technologies suitable for smartphones like light modulation (Philips Lighting Holding B.V., 2017), sound and ultrasound (Smith et al., 2004; Ward et al., 1997; Knauth et al., 2009, 2015a), magnetic fingerprinting, and dead reckoning (Kang and Han, 2015; Willemsen et al., 2015; Knauth and Koukofikis, 2016) using the built-in sensors of smartphones like barometer, compass, acceleration, gyroscope, etc. These methods may also use RSSI as supporting technology.

RSSI-based indoor positioning using standard smartphones is particularly interesting because users may apply this technology in existing environments where Wi-Fi access points (APs) are deployed. Smartphone-based localization may be performed even if the

Geographical and Fingerprinting Data for Positioning and Navigation Systems. https://doi.org/10.1016/B978-0-12-813189-3.00006-X

Wi-Fi installation has not been specifically prepared for that use case and, for example, characteristics and positions of the Wi-Fi nodes are unknown.

In Section 2 some important features of the indoor radio propagation are recalled. Section 3 is on proximity and centroid methods which most of them do not need effortful recording of reference maps, at the cost of limited accuracy. Wi-Fi fingerprinting principles and the authors' algorithms are presented in Section 4. Section 5 focuses on fingerprint calibrated weighted centroid, which is a special method combining fingerprinting and centroid. In Section 6 running algorithms of the presented methods are discussed and the results are compared against each other and "third-party" solutions. Section 7 outlines some more exotic methods, and Section 8 closes the chapter with a short summary.

2 Indoor Radio Propagation

2.1 The Free Space Model

For the smartphone signal strength-based positioning problem, it is interesting how the signal at a receiver depends on the environment and the distance to a transmitter, for example, a Wi-Fi AP. In free space, radio waves expand more or less undisturbed on straight trajectories. For frequencies below 20 GHz and ranges up to some kilometers, the atmospheric absorption can be neglected. So at typical Wi-Fi frequencies, the received power P_r is given by geometric considerations: A part of the transmitted radio power P_t reaches the receiver antenna, the rest of it beams in other directions. In the isotropic case, the received energy is given by the effective antenna cross-section (which may differ from the physical one) with the sphere with radius d around the transmitter. Here d denotes the distance between transmitter and receiver. The ratio P_r/P_t is named "path loss." One should keep in mind that path loss does not automatically mean that the power is absorbed by some means, in the described case the antenna captures less power because the radio beams diverge.

The effective antenna area of typical antennas is proportional to the square of the wavelength λ. With the sphere surface being $4\pi d^2$ this leads to the Friis equation (Friis, 1946):

$$P_r/P_t = G_r G_t \frac{\lambda^2}{4\pi d^2} \tag{1}$$

The Friis equation describes the relation between distance and path loss in free space, where P_r is the received power and P_t expresses the transmitted power, in nonlogarithmic units. Further variables are the antenna gains G_r and G_t. The formula can be written also as

$$P_r/P_0 = d^{-2} \tag{2}$$

with P_0 being the received power at a distance of 1 m and d given in meters.

2.2 Indoor Propagation

Unfortunately, the free space model performs quite poor in indoor scenarios. Strong deviations from the model arise mostly due the following effects (Rappaport, 2002):

- Radio waves are heavily reflected in the inside of modern buildings. Iron and steel, copper, concrete walls, and a lot more of modern building materials strongly reflect radio waves, making the inner of a building like a mirror cabinet for the latter. Long corridors act as wave guides and considerably reduce the path loss.
- Multipath fading is caused by interference of signals from different path, for example, by reflection. Interference may be constructive or destructive, the destructive interference leading to strong fading up to vanishing of a signal, even when close to the source. Multipath fading caused by reflections is a major cause for the difficulties in reliably predict path losses in buildings with the Friis formula.
- Absorption of radio power is caused, for example, by the bodies of people and by certain building materials, mainly those who can keep moisture. Absorption and reflection leads to attenuation of the signals when traversing through walls and ceilings.
- The orientation of the mobile antenna depends highly on the users pose and the position of the smartphone. Real existing antennas may be isotropic in two dimensions, but not in all three dimensions, and will therefore always have some directional behaviors.

The effects can to some extend be approximated by using an exponent n different from the free space value of 2 for the distance d in Eq. (2) (see e.g., Hashemi, 1993), leading to

$$P_\mathrm{r}/P_0 = d^{-n} \tag{3}$$

$$d = \left(\frac{P_0}{P_r}\right)^{1/n}, \tag{4}$$

or, with logarithmic units

$$d = d_\mathrm{RSSI} = 10^{-S\,[\mathrm{dBm}]/(10\cdot n)} \cdot C, \tag{5}$$

where S is the received RSSI value in dBm and C is a constant expressing P_0. Measured in linear units like watt, the path loss is the quotient of transmitted power and received power. Going to logarithmic units, for example, dBm, the path loss L formula changes to $L = P_\mathrm{t}\,[\mathrm{dBm}] - RSSI + c$.

In Fig. 1 curves for several values of n are shown. In buildings, typical values are between 2.5. Depending on antennas and transmission power, there may be a different offset C and the curves may have to be shifted up or down, but the characteristics remain the same. Considering a noise level of, for example, 10 dBm it is obvious that a range estimation is most useful for distances of a few meters only.

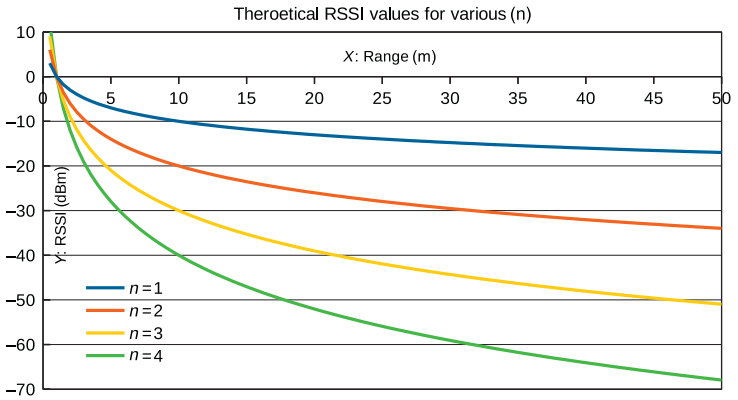

FIG. 1 Received RSSI value for different propagation exponent n. Depending on transmitter power, all curves may have to be shift up or down.

2.3 The RSSI Measure

In the following chapters mostly the RSSI will be used. The RSSI represents the path loss in logarithmic units. From experience, the authors assume that typical values reported from smartphones lie in the interval $0, \ldots, -95$ dBm. 0 dBm would be regarded as a very high RSSI value, it is likely that a smartphone detecting this value is very close to an Wi-Fi AP. -50 dBm could indicate a 5.25-m distance or a signal which passed some walls. Values of -95 dBm indicate a possible high distance, or strong attenuation, for example, by waves crossing several concrete ceilings or walls. Already these estimates show that application of the modified Friis path loss model is of limited use for a single position estimate and must be seen as a statistical approach.

3 Wi-Fi Positioning by Centroid Methods

3.1 The Centroid Method

The problem of RSSI positioning typically introduces a "rover," also referred to as "mobile node," whose position R shall be determined using path loss measurements between this point and a set of fixed nodes, typically Wi-Fi APs at a known or unknown position R_q. It is assumed that the fixed nodes emit radio packets and the rover determines the received signal strength (RSSI), but operating vice versa would not change the path loss. As we are discussing mainly smartphone Wi-Fi RSSI positioning, the rover can be identified with the smartphone, and the reference points are the AP.

A quite simple approach for estimating the position is to report the position of the AP, which is received with the highest RSSI value, as the rovers position. This is often referred to as the proximity method. The accuracy of this method is obviously related to the density

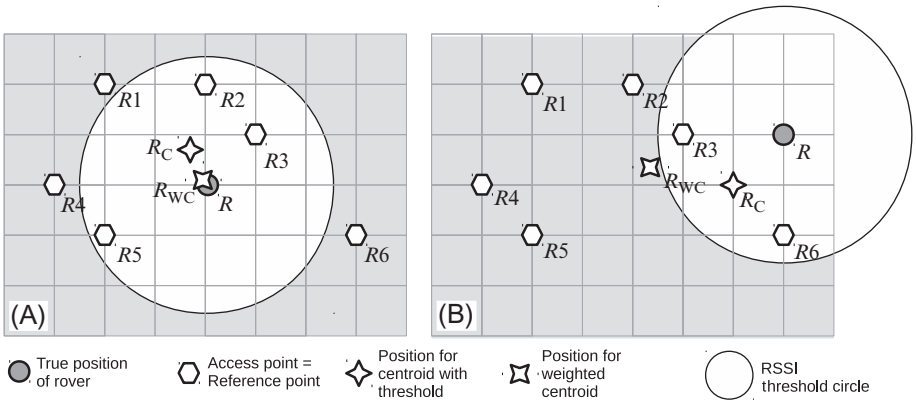

FIG. 2 Two sample scenarios, for the centroid and weighted centroid methods each. *Threshold circle* marks the area over which summation of access point positions is performed for the unweighted centroid. (A) A scenario where a rover is surrounded by a set of access points. Obtained positions more or less match the rover position, for both methods. (B) An anisotropic scenario where the rover position lies outside of the convex hull of the access point set. For the threshold-based centroid, the result is still reasonable, while the weighted centroid is "clamped" by the access point positions.

of APs. Looking at Fig. 2, for case (A) this would be R_3, for case (B) this would be either R_3 or R_6, whichever would be received with a slightly higher RSSI value.

Better results are obtained when calculating the "average" position of all fixed points R_q for which the path losses to the rover are below a certain threshold. Mathematically this is performed by averaging the coordinates of the respective points (the centroid):

$$R_{\mathrm{C}} = \frac{\sum_q R_q}{\sum_q 1}, \tag{6}$$

where q runs over all fixed points fulfilling the threshold criterion.

3.2 Weighted Centroid Method

A more sophisticated approach is the weighted centroid:

$$R_{\mathrm{WC}} = \frac{\sum_q R_q \cdot w(q)}{\sum_q w(q)} \tag{7}$$

$$w(i) \propto |R_q - R_{\mathrm{WC}}|^{-1} \tag{8}$$

Each reference point R_q is given a weight $w(q)$ with respect to the current measurement at R. Eq. (8) defines the "ideal" weight, that is, a weight which is inverse proportional to the distance d_q between the rover R and a fixed point R_q. Unfortunately, the rover position and the distances are not known in advance. A real weight function shall estimate a range based on the observed path loss or RSSI value.

In Fig. 2 two sample scenarios for the centroid and weighted centroid methods are sketched. APs cover a certain area. The real position of a rover is indicated. For the

threshold-based nonweighted centroid, the "threshold circle" marks the area over which the summation of AP positions is performed. For case (A), k would be in {1, 2, 3, 5} while in case (B) this would be {3, 6}. The positions obtained using threshold-based centroid (R_C) according to Eq. (6) and by weighted centroid (R_{WC}) (Eq. 7) are indicated for the two different scenarios. In Fig. 2A the rover is surrounded by APs. In this case, both methods deliver a more or less reasonable result. Obviously, the centroid method always delivers a result, which lies within the convex hull of the set of APs, as long as no negative weights are considered. So the true position R should also be within the convex hull, otherwise reasonable errors will be introduced.

For the situation outlined in Fig. 2B, the convex hull condition is not fulfilled. For the threshold-based centroid, the result will at least lie within the threshold circle. For the weighted centroid the situation is different. The spatial distribution of the APs is quite anisotropic, with respect to the position of the rover. The obtained result is then "compressed," that is, shifted more to the center of gravity of the APs. As it can be already estimated from Fig. 2, centroid is useful especially under conditions where an area is regularly covered by APs, and the rover position lies within that area.

If the convex hull condition for the centroid method is not given, for example, positioning shall also occur in areas outside of an AP "cloud," centroid results may not be satisfying. A possible solution could be the usage of multilateration algorithms operating on the ranges obtained by Eq. (5). However, classic multilateration does not take into account weights. This is not feasible for RSSI, since low RSSI values are typically quite noisy, and the obtained range is therefore quite error prone. Nevertheless, optimized multilateration has been applied (Gau et al., 2008; Mautz et al., 2007) to wireless sensor networks. An alternative approach is probability map-based algorithms. An example is given in Section 7.

4 Wi-Fi Fingerprinting

4.1 The Radio Map

The last sections described methods for position estimation based on the Friis model or else model. While the positions of the APs have to be known, no site surveying, etc., is used here. The expected accuracy will be, for example, limited by the density of APs. As a rule of thumb, the expected accuracy will not exceed 0.5,...,1 times the average spacing of the APs.

Higher accuracy may be obtained by surveying the area under interest: Before the system can determine the position of a rover, a database of "fingerprints" needs to be created. A single fingerprint is created by recording RSSI values for all receivable APs at a certain position R_i, where i is an index for the reference positions, and the number of reference positions is n. The received RSSI values at position R_i form a row vector

$$S_i = \{s_{i,1}, s_{i,2}, \ldots, s_{i,q}, \ldots, s_{n,m}\} \tag{9}$$

with q being an index for a certain AP, m being the total number of APs, and $s_{i,q}$ being the RSSI value received at R_i from AP with index q. The APs index is an enumeration. The actual identification of the AP is performed via the APs' MAC address. Special handling is needed if a certain AP is not received at a certain position. Since a vector needs values for each coordinate, some special value may be used to indicate that no signal has been received. For example, if measured values lie in the interval 0,..., −100 dBm, a "null" value of −101 could indicate that no value has been recorded for that AP. The row vector consisting of all surveyed reference point vectors S_i forms the radio map which can also be seen as a matrix S

$$\mathrm{S} = \begin{array}{c} \\ S_{i=1} \\ S_2 \\ \vdots \\ S_n \end{array} \begin{array}{c} \begin{array}{cccc} AP_1 & AP_2 & \dots & AP_m \end{array} \\ \begin{pmatrix} s_{1,1} & s_{1,2} & \dots & s_{1,4} \\ s_{2,1} & s_{2,2} & \dots & s_{2,4} \\ \vdots & \vdots & s_{i,q} & \vdots \\ s_{n,1} & s_{n,2} & \dots & s_{n,4} \end{pmatrix} \end{array}, \tag{10}$$

where the row vectors S_i are the reference vectors measured at position R_i (Eq. 9), and the column vectors AP_q are the measured RSSI values for the qth AP. When a radio map is available, position estimates can be calculated by performing the k-Nearest Neighbor (kNN) method with the following steps:

- Observe RSSI values of APs at the unknown position R_x and create a measurement vector S_x.
- For each reference vector S_i calculate the similarity $\sigma_{x,i}$ between S_x and S_i.
- Take the set of the k reference vectors S_i which are most similar to S_x.
- Calculate a final estimate based on the positions R_i of the k most likely S_i.

The final estimate will typically be again a centroid method. Weights could be based on the similarity measure σ or could just be fixed. The authors typically use $k = 3$ and use fixed weights of 2–1–1 in descending similarity order.

4.2 RSSI Vector Similarity Measures

Popular algorithms like the above-discussed kNN fingerprinting compare measured RSSI vectors with the RSSI vectors in a radio map database. A crucial step in this algorithm is of course the comparison itself. There is quite a choice for similarity measures. The most common are

$$\text{Euclidian distance: } \sigma_{x,i} = \sqrt{\sum_q (s_{x,q} - s_{i,q})^2}, \tag{11}$$

$$\text{Manhattan distance: } \sigma_{x,i} = \sum_q (s_{x,q} - s_{i,q}), \tag{12}$$

$$\text{Euclidian or cosine norm: } \sigma_{x,i} = \frac{\sqrt{\sum_q (s_{x,q} \cdot s_{i,q})^2}}{|S_x| \cdot |S_i|}. \tag{13}$$

While the Euclidian distance is widely used as comparison norm, there are quite some more choices like, for example, the Manhattan distance, the cosine norm and others. Focusing on the three mentioned, they have advantages and disadvantages. First, it must be said, that for kNN, the similarity measure is used mainly for ranking of the RSSI vectors. Since the measures are monotone, if there is a reasonable difference in the top-similar fingerprints and a small is k used, either norm will produce the same ranking. In particular when using fixed weighting and a small k, for the three mentioned norms it is likely that the same result will be obtained for the estimated position on either norm. However, this is not generally guaranteed. A comprehensive analysis of the Wi-Fi similarity problem is given, for example, in Torres-Sospedra et al. (2015).

There is a principal problem that different phones even of the same model may have a different offset in the RSSI values. This may be caused, for example, by different Wi-Fi chipsets, protection cases around the phone and different pose of the phone during measurement. Such an offset in the RSSI values may, in particular for the distance-based norms, lead to failures in the ranking of the vectors. Avoidance is typically performed either by include already different phones and poses in the mapping phase or by normalizing procedures, for example, normalizing of the map vectors and the measured vectors. The normalizing is self-included when using the cosine norm. Actually the cosine norm is more or less performing a correlation of the vectors more than a difference calculation.

There are other peculiarities when dealing with the similarity:

- In practical measurements, depending on the scale of the deployment, a majority of APs will not be received at all. The formulas have to be modified to cope with such "null" values. One solution is to replace "null" values with the lowest overall obtained RSSI value. But in situations, where an AP is taken out of operation, or during the scan period has not been captured, this would lead to reasonable error in similarity estimation. Another approach is to compare only those RSSI values, which are not null in both, in the radio map data and in the online received data. Of course, this would have to be taken into account also into the normalization such that the normalization should only count values, which are available on both sides.
- The RSSI-based similarity does not take into account the "credibility" of the RSSI value. As stated in Section 2, RSSI values are typically quite noisy. A difference of ± 10 dBm relates to a small range difference when being at a level of -20 dBm, but gives a high spatial range difference at -70 dBm. This can be addressed by performing a weighting of the RSSI values before comparison. A natural choice for the weighting function is the inverse distance calculated from Eq. (5):

$$d^{-1} = 10^{S\,[\text{dBm}]/(10 \cdot n)} \tag{14}$$

with the propagation exponent n. The constant C expressing the transmitter power and antenna gain has been omitted since it will vanish on the normalization step in the similarity measure, visible, for example, in Eq. (13).

5 Fingerprint Calibrated Weighted Centroid

Wi-Fi fingerprinting algorithms do not need to know the position of the APs, since all necessary information is captured by the radio map generation. In some situations, however, the size of the radiomap may be a burden for a specific implementation. The radio map contains n reference vectors, each of it containing an array of q values for the APs. The data set of the 2015 IPIN EvAAL competition had about 500 APs and more than 800 reference points leading to about 500,000 values, leading to a database size of about 1 MB, in compressed format. Also, each position estimate has a complexity of about 400,000 calculations. While this is not a serious task for a smartphone, embedded devices based on small microcontrollers reach their resource limits with these numbers.

The fingerprint calibrated weighted centroid (FCWC) algorithm (Knauth et al., 2015b) processes the radio map data to estimate the positions of the APs. A rover may then use these positions to calculate a position estimate, for example, via weighted centroid. The size of the database is thereby collapsed to the positions of the APs, which will be not more than 500 coordinates in the described case, thus reducing the size of the database down to some kilobyte. As the estimates are performed by weighted centroid (Eq. 7), the computational effort reduces to a number proportional to the AP count, in the given case to 500.

The method works by first estimating the positions of the APs offline by "reverse positioning": By transposing of the RSSI matrix (10), the qth row vector contains now a set of RSSI values for the qth AP. Each RSSI value corresponds to a certain known reference position R_i. The path loss is independent of the propagation direction, that is, the path loss from a transmitter to a receiver is the same if the roles of transmitter and receiver are exchanged. So the unknown position R_q of the AP can be estimated by the weighted centroid method using Eqs. (7), (8).

Once the positions of the APs have been estimated, they may be used for online position detection. A position estimate can now be based on the RSSI report S_i of an unknown location and the AP positions R_q using again Eqs. (7), (8), but substituting R_i with R_q. It should be noted that the estimated AP positions not necessarily have to be close to their real positions. The estimated positions are called "virtual positions" and might be interpreted as just being parameters for a Wi-Fi-based parametric approximation, based on the radio map data.

A comparable approach has been described for ultrasound positioning in Knauth et al. (2013).

For both, the fingerprinting and the FCWC algorithm, in multifloor scenarios also the floor has to be identified. Therefore, each reference point is assigned a z-coordinate

based on the floor ID given in the dataset. The centroid calculation also processes the z coordinates such that the virtual APs and the position estimates contain a z-coordinate. This coordinate is then used to assign a floor to the result.

6 Validation of the Described Fingerprint and FCWC Schemes

6.1 Validation Data

The validation is performed using publicly available measurement data provided by the UJIIndoorLoc database (UJIIndoorLoc Data Set at UCI Machine Learning Repository, 2014; Torres-Sospedra et al., 2014). This data formed the base for the EvAAL 2015 competition on indoor localization (EvAAL: Evaluating AAL Systems through Competitive Benchmarking, 2015) track on "Wi-Fi fingerprinting in large environments (off-site)." The presented algorithms participated at the competition as one of four competitors, allowing to compare the results of the two algorithms (fingerprinting and FCWC) between each and also to other schemes and implementations (Torres-Sospedra et al., 2017).

The provided data comprises a dataset for training and adjusting of the algorithms, which is enriched with ground truth information, and a competition dataset for which the ground truth is not known to the competitors. The UJIIndoorLoc training table comprises about 20,000 measurements (rows) collected with a variety of mobile equipment at 933 different positions, distributed within 3 huge university buildings each having 4 to 5 floors. The area is equipped with 520 Wi-Fi APs. Each data row consists of 520 RSSI values, one for each AP, and additional information like the position of the measurement, the used equipment, a code for the person performing the measurement, a building ID and a floor ID.

6.2 Algorithm Implementations

As the described algorithms run on local metric coordinates, latitude/longitude position data are converted into a local x–y coordinate with respect to a reference point in the middle of the spawned area and a reference direction, which is aligned to the main structures of the given data. A z-coordinate is also introduced, which is set to floor ID times 5 m. Note that, due to the linearity of Eq. (7), the kNN calculation used in both algorithms is independent of the used coordinate system and could actually be performed with latitude/longitude values instead of local coordinates. The local coordinates are convenient for debugging purposes and calculation of deltas between given true and estimated positions. Also the floor spacing value does not need to match the true floor spacing, as it is only used to determine the floor ID, not the real z position of an item.

A fingerprint-based algorithm "SPCF" (scalar product correlation fingerprinting) has been set up as described in Section 4. The algorithm employs the cosine norm (Eq. 13) on

weighted RSSI values (Eq. 14) as similarity measurement and uses kNN with fixed weights for final position estimation.

The parameters used were 2–1–1 as fixed weights, $k = 3$, and n in [2.5,..., 4], depending on the building. These values have been determined manually by iteratively optimizing the algorithm output against the provided true ground positions in the validation data. The optimization criteria were the average position error not regarding floor and building penalties.

A further optimizing has been reached by removal of "bad" APs: For each AP, the algorithm was run one time with a particular AP included, one time with the specific AP excluded. If the result was better when the specific AP was not included, it was considered as "bad AP" and its data were ignored in the processing. This led to some decimeters of overall accuracy gain.

Also an FCWC algorithm was implemented, as described in Section 5: A reverse positioning of the APs is performed. The training table with the calibration data is transposed such that for each AP, the RSSI values measured at different positions R_i are available. The yet unknown position R of an AP is now calculated using Eq. (7) with the weight function (14). The parameters used were mostly same as for the SPCF algorithm. Also, the described optimizing was applied.

However, some parameter settings were altered for FCWC: the different parameter settings evaluated were the exponent n and the lower RSSI threshold $RSSI_{\min}$. Obtained values were $RSSI_{\min} = -85\,\text{dBm}$ and $n = 1.0$. This value is surprisingly low as typical indoor propagation exponents are above 2.0. In the given scenario the low exponent favors high RSSI values, because weak readings will be assigned a particularly low weight. This seemed to be important in the FCWC case, where the weighted centroid is applied twice.

6.3 Results SPCF and FCWC

The overall mean delta of FCWC against the provided validation data file is about 9.7 m (see also Fig. 3). This accuracy is x–y based and does not include the height displacement. Using the calculated positions the floor and building ID is estimated. It is found that about 94% of floor IDs are reported correctly, and the building ID has been assigned correctly to all measurements. For all measurements, a position has been estimated. The discussed and implemented SPCF (scalar product correlation fingerprinting algorithm) delivered an average delta of 7.7 m and a floor ID detection rate of about 96% against the "validationData.csv" file and thus performed reasonably better on the validation data. As it will be seen in the next section, for the private test set of the competition, the difference in performance between the two algorithms was much lesser.

6.4 Validation Against Competing Algorithms

Four teams have participated at the offsite track of the 5th EvAAL competition. Detailed results have been published at the competition website (EvAAL: Evaluating AAL Systems through Competitive Benchmarking, 2015). Table 1 lists the results of the four teams.

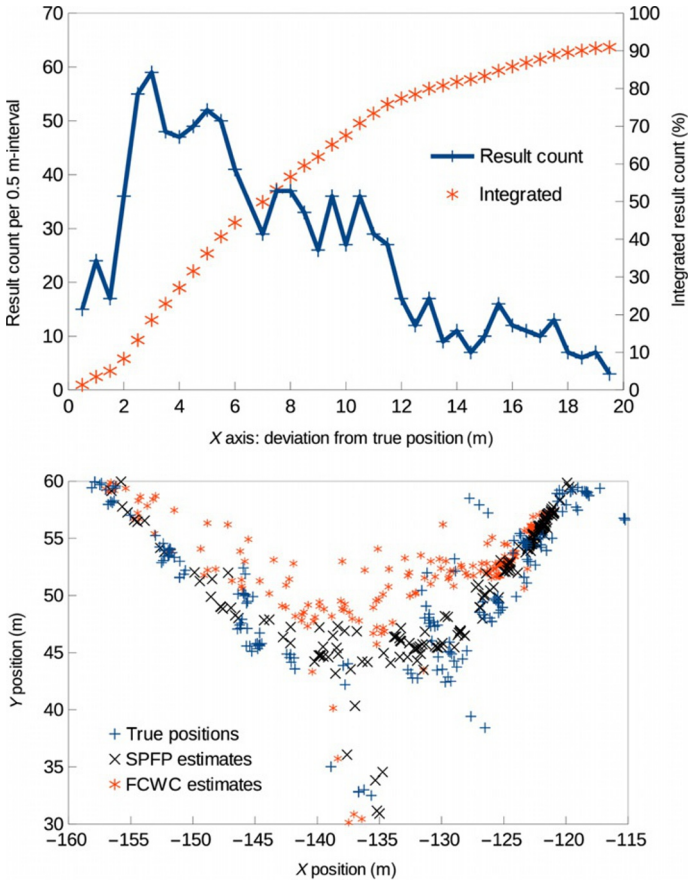

FIG. 3 *Top*: Delta between calculated positions and true positions, for the public test set, for FCWC. The *solid line* displays the number of measurements obeying the indicated (*x*-axis) delta. The interval for the value count is 0.5 m. The total measurement count is 1110, of which about 10% have a delta of 20 m and above and lie outside of the figure. The *asterisk line* indicates the integrated number of measurements, in % of the total number. The 50th percentile delta value is about 7 m, the 75th percentile about 11.5 m. *Bottom*: FCWC and SPCF results. Results of the centroid algorithm always lie within the convex hull of the reference points. Therefore, they more accumulate in the centers of covered areas.

The authors team "HFTLoc" is listed in the second and the third columns. The ranking of the teams has been performed according to the row "third-quartile error," with better results listed more right. For the "HFTLoc" team, the "official" results have been obtained using the SPCF algorithm. The extra column lists the results obtained with the FCWC algorithm and is included to allow comparison of FCWC and SPCF competition performance. A topological overview of the FCWC and SPCF results is given in Fig. 3.

Team "MOSAIC" uses a probabilistic Wi-Fi sensor model and calculates the mean mutual information. Maximum likelihood estimation and *k*NN are then applied as localization scheme (Berkvens and Weyn, 2015). Team "ICSL" (Choi et al., 2015) uses machine

Table 1 EvAAL 2015 Offsite Track Results

Team Name	MOSAIC	HFTLoc SPCF	HFTLoc FCWC	ICSL	RTLSUM
Building success rate (%)	98.65	100.00	100.00	100.00	100.00
Floor success rate (%)	93.86	96.25	93.71	86.93	93.74
Floor error rate (%)	6.14	3.75	6.29	13.07	6.26
Floor ± 1 rate (%)	98.88	99.98	100.00	100.00	99.79
B&F success rate (%)	92.59	96.25	93.71	86.93	93.74
Mean error if B&F success	7.97	8.40	8.55	6.80	5.71
Mean error	11.64	8.49	8.69	7.67	6.20
First quartile error	3.26	3.69	4.43	3.10	2.51
Second quartile error	6.72	6.99	7.41	5.88	4.57
Third quartile error	12.12	11.60	11.79	10.87	8.34

learning techniques and a deep learning approach. Team "RTLSUM" (Moreira et al., 2015) employed besides fingerprinting also temporal filtering. Details can be found in Torres-Sospedra et al. (2017).

An important finding is that, for the three left columns, the results are quite close together, while the approaches taken are quite different. This may be an indication that the information content of the dataset does not allow to come to a more accurate result. For the convincing results in the rightmost column, the team among others applied temporal filtering.

The presented SPCF approach turned out to be particularly successful in estimating the floor. This may be attributed to the usage of 3D coordinates, that is, using x, y, and z, when averaging positions for calculating of a result. The FCWC algorithm scored quite comparable, which is an amazing result since it means that in large deployments it is possible to use only one vector describing the virtual positions of the APs instead of the whole fingerprint database, and still get reasonable results. The described algorithms can well compete with algorithms published by other groups and may therefore be helpful for implementing fingerprint smartphone indoor location systems.

7 Wi-Fi Probability-Based Positioning and BLE

7.1 The Probability Density Function

A drawback of the centroid and the kNN step in fingerprinting is that obtained positions are always reported within the convex hull of the APs, as averaging can not go "out" of the convex hull. This restriction is not present for Multilateration. However, the latter is heavily disturbed by noise, offsets, etc. The parametric approach presented in this section merges to some extend the concepts of proximity and multilateration to obtain accuracies more close to those of fingerprinting algorithms. There is only minor mapping effort needed, in comparison to fingerprint-based approaches.

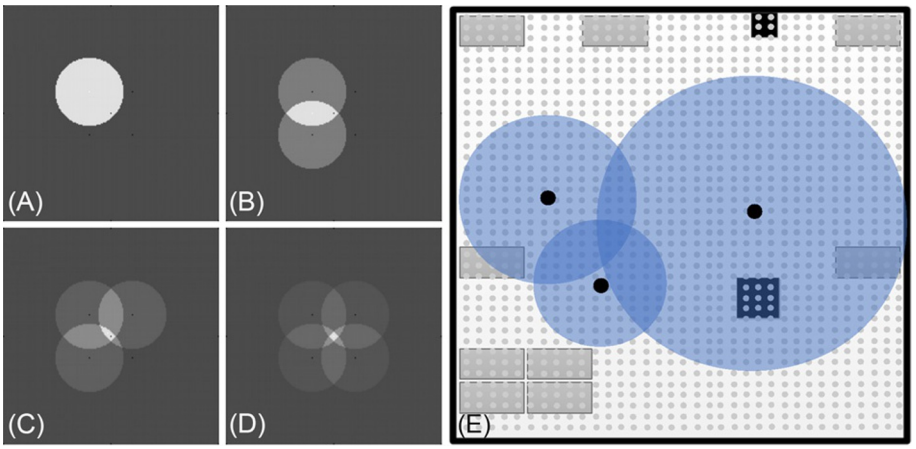

FIG. 4 (A) Probability map (representation of a probability density function) of a single receiver, for an RSSI value corresponding to the diameter of the circle. (B)–(D) Merged maps for several receivers. (E) Merged map showing also the underlying calculation raster, and some orientation marks in the room. Overall size of (E) is about 6 × 6 m. *Circle areas with small radius* correspond to high RSSI values/low path losses.

The probability-based positioning algorithm makes use of a parametric probability density function $p(d, s)$. The function describes, for a received RSSI value s, the probability p for the rover to be at a distance d from the reference position. In Fig. 6B, a 2D gray scale map representation of such a function is displayed where d is the distance between a map pixel and the position of a receiver R_i. R_i is the dark point in the center of the rings.

Fig. 4 sketches a simple variant of $p(d, s)$: here the probability is 1 for points R within a certain range $d = |R - R_i| < d_{\text{RSSI}}$ around the fixed node position R_i, and 0.3 for other positions. The figure represents this function as a gray scale image. d_{RSSI} is given by the observed RSSI value according to Eq. (5). Already this easy probability density function respects some important properties of the indoor path loss, namely that the distance between fixed node and mobile node can be small, even if only low RSSI values are reported (high attenuation), but for a high RSSI value, it is unlikely that there is a large distance between transmitter and receiver. The figure sketches the principle of operation: for each AP, there is a probability density function $p(d, s)$ describing the probability of the mobile node to be at a certain position, based on the RSSI information of that AP. Fig. 4B–D represents subsequent introduction of further APs. The probability maps of the APs are merged by multiplying thus creating an overall merged map. The merged map indicates the overall probability of presence of the mobile device for a given coordinate. This map could, for example, be used to seed new particles in a particle filter. In the presented work we estimate the position of the mobile node by finding the position of highest probability. Fig. 4E displays again the merging of three probability functions, and indicates the discrete points for which probability values are calculated. In practice, more complex probability density functions are used. For example, the function may model

the directional behavior of the fixed node antenna. A more smooth function will typically result in smoother absolute maximum regions in the merged probability map.

The choice of a suitable $p(d, s)$ is important for the success of the method. For the example shown in Fig. 6, the function was defined such that at distances related to d_{RSSI} (Eq. 5) the function has the highest probability. A typical function will look like

$$p(d, s) = \frac{1}{(d - d_{\text{RSSI}}(s))^2 + c}, \tag{15}$$

where $d_{\text{RSSI}}(s)$ is the RSSI-based distance estimation for an RSSI value s, and c is a constant defining the "sharpness" of the function. An example for $d(s) = 2$ m and $c = 0.5$ is plotted in Fig. 6C.

Eq. (15) gives the probability from one AP. In order to get a probability map which considers all APs, the probabilities are multiplied. The residual probability including m APs is

$$p(R, S) = \prod_{q=1,\ldots,m} p(|R_q - R|, s_q) \tag{16}$$

where R is the position of the mobile node, R_q is the position of the qth AP, and S is the measured RSSI vector at position R. The above result delivers the probability distribution for all APs.

7.2 A Probability-Based Setup and Algorithm

A simple approach to get a position estimate is to use the position of the maximum of Eq. (16). For the validation of the presented probability-based positioning algorithm, the maximum is found by evaluating $p(R, S)$ on a grid of positions with a certain spacing. The spacing should be smaller than the expected accuracy and also smaller than the features of the probability density function $p(d, s)$. A typical value is 25 cm. For large deployments performance might be an issue. In that case, a two-phase approach will be useful: First, the maximum on a coarse grid is searched using a smooth $p_{\text{coarse}}(d, s)$. Then, around the obtained coarse position, a fine search is performed using a sharp $p(d, s)$. The algorithm may be used to calculate probability distributions or "maps." These maps can be used as one of the input data for probabilistic positioning schemes using Markov chains or for seeding in particle filter approaches.

The algorithm has been implemented in a small scale lab setup (Knauth et al., 2014). Instead of using Wi-Fi, a different setup consisting of smartwatches and microcontroller-based receivers was used. The principles are applicable also to Wi-Fi setups. A general operation schema is shown in Fig. 5. Fig. 6B sketches the actual setup in the Lab: Fixed receiver nodes were installed in a 6×6 m sized room. Six CC1110DK MINI nodes were positioned on tables in the lab. Eight ground positions were defined and marked on the floor. For these positions, measurements were recorded. The mobile node consisted of a student wearing an EZ430-Chronos smart watch. The student visited the eight marked positions. A second person recorded accurate timing information, that is, at what time

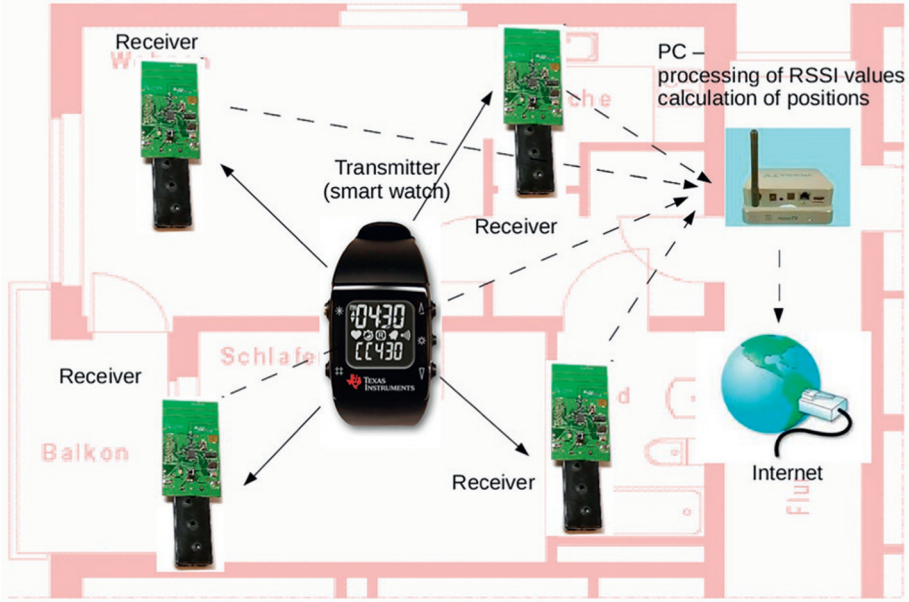

FIG. 5 Schematic setup used for validation of the probability-based parametric method: An EZ430 Smartwatch emits 868 MHz radio packets, which are received by CC1110DK MINI nodes (*solid arrows*). These fixed nodes transmit their obtained RSSI values to a central PC (*dashed arrows*), which logs the values and runs the algorithm. Note that in the experiment, six receiver nodes were used.

which position was visited. The smart watch emits 868 MHz RF Packets of about 3 ms duration each, at a rate of 30 Hz. The six receivers collect their observed RSSI values and transmit them to a PC equipped with an 868 MHz USB dongle receiver. A TDMA scheme is used to avoid collisions.

For each test position, the probability density is calculated according to Eq. (16) using Eq. (15) as probability density function. The calculation is performed over a grid of R values, each grid point position related to a pixel of a bitmap graphic, representing the probability map. The resulting distribution is shown in Fig. 6, inset 1–8, for each of the reference points. The maps indicate the probability for the mobile node to be at a certain position.

7.3 Probability-Based Results

Table 2 lists the positioning results of the algorithm. Obtained errors lie in the range of some centimeters up to about 2.5 m, with an average error of about 1.2 m. As it can be seen, high errors were obtained for positions 6, 7, and 8 (see also Fig. 6). This might be attributed to the fact that these points lie more or less outside of the area spanned by the reference nodes, which generally leads to less good algorithmic conditions for range-related positioning methods. As far as general statements can be deducted from such a

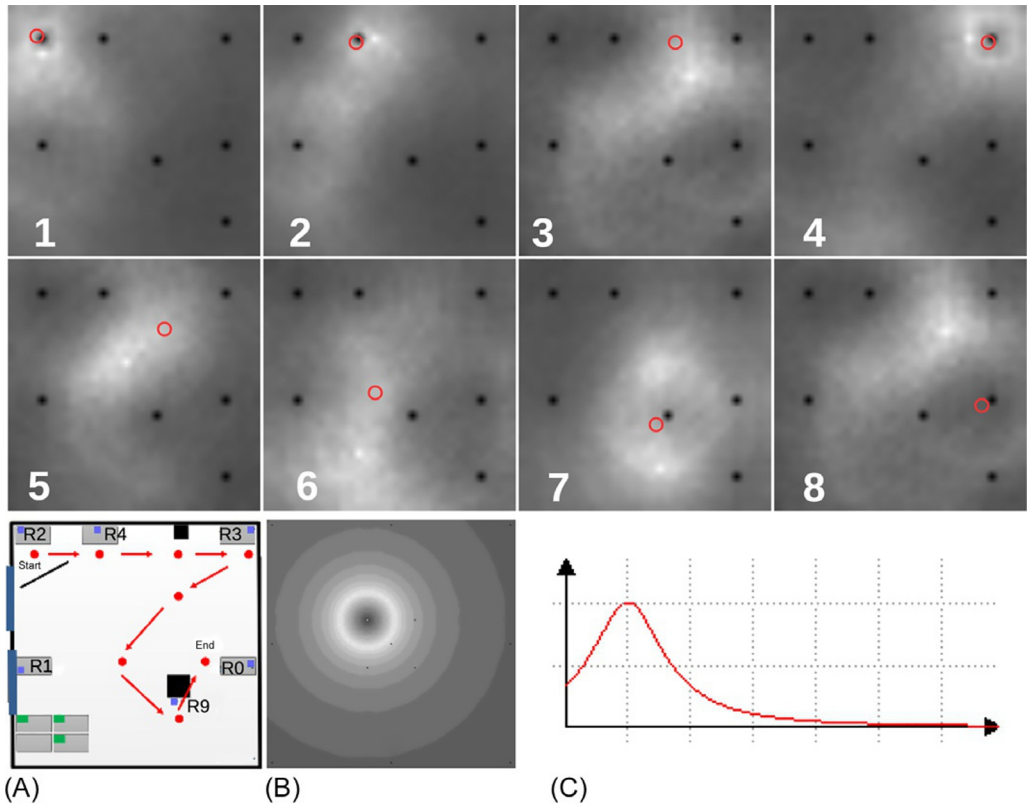

FIG. 6 *Top 1–8*: Obtained probability maps for eight measurement points. The map numbering corresponds to the sequence of the points in the path indicated in the *left lower inset*. The positions of the fixed nodes are indicated as *dark points* in the maps. *Bottom* (A): Situation in the lab. *Gray rectangles* indicate tables, receiver positions are marked with *asterisks*. The test path is indicated with *arrows*, measurement positions are indicated by *filled circles*. (B) Typical probability map as used in the experiments, in a gray scale representation. *Dark parts* indicate low probability, *bright parts* indicate high probability. The size of the map is 10 × 10 m, the d_{RSSI} value is 2 m. For lower RSSI values, the radius of highest probability (*ring with bright gray* values) would be larger, and vice versa. (C) Plot of a probability function $P(s, d)$ for an RSSI value s related to 2 m distance.

Table 2 Probability-Based Results

	Real Pos.	Estimated	Error
1	0.00/5.50	0.00/5.38	0.12
2	2.00/5.50	2.44/5.94	0.62
3	4.00/5.50	4.70/5.00	0.86
4	6.00/5.50	5.35/6.00	0.82
5	3.70/4.00	2.80/3.70	0.95
6	2.30/2.45	2.11/0.60	1.86
7	3.70/2.00	3.50/0.25	1.76
8	5.00/2.45	4.51/4.83	2.43

minimal proof-of-concept setup, obtained results indicate that the accuracy of the method could be comparable to weak-mapped fingerprinting, but besides a one-time calibration to resemble the internal path loss and antenna pattern of the reference node model (the constant C in Eq. (5)), no calibration has to be performed.

The method employs an RSSI-dependent probability density function, which is applied for each fixed node. The results are combined by multiplying to get the overall mobile nodes spatial probability distribution. Values were calculated for a grid of positions. The method has the advantage that no fingerprint maps have to be generated, certainly at the cost of a somewhat lower accuracy but allowing for less engineering effort during deployment. While in this experiment the "brightest point," that is, the most likely position has been chosen as the reported position estimate, the generated maps allow also for more sophisticated position estimators, especially when using them as seeding information in particle filters or other probabilistic methods.

7.4 BLE Beacon RSSI Weighted Centroid

Weighted centroid can be a good solution for Bluetooth Low Energy (BLE) beacon-based positioning. The following laboratory scale experiment was setup: A number of 14 BLE transmitters ("iBeacons") were installed at defined positions in a hall. BLE uses the same 2.4 GHz band as Wi-Fi, so no principal differences in radio propagation are to be expected between the both. A Samsung Galaxy S4 Mini smartphone was used for detection. Sixty-five test points were marked on the floor. A test person visited the test points and recorded measurements for about 10 s, at each test point. For each beacon and test position, on average eight RSSI values were recorded in that time.

In Fig. 7 the geometry and results for two test scenarios are shown: 14 transmitters are arranged around the about 300 m^2 test area. The weighted centroid algorithm is applied and in the left case, the average deviation was about 4 m. Note, that in this case, the orientation of the test person and smartphone was random.

In the experiment run demonstrated at the right-hand side of Fig. 7, the test person always kept herself and the smartphone in a certain pose, that is, facing the top border with respect to the image. The obtained average deviation for this case is about 4.5 m, but as it can be seen, this is not equally distributed over the test area but more pronounced at the lower parts. A quite likely explanation of that behavior is that the directional antenna pattern and the RF shading of the test persons body reduces weights of the "lower" transmitters, mainly IDs 1, 2, 7, 10, and 11.

The displayed results were actually not obtained with a pure weighted centroid approach, but by combining weighted centroid and introducing a lower RSSI threshold of −85 dBm such that only APs (here iBeacons) which have an RSSI value above this lower RSSI threshold are considered in the weighted centroid algorithm. So it could be called "proximity weighted centroid." The threshold value was determined by optimizing the mean error.

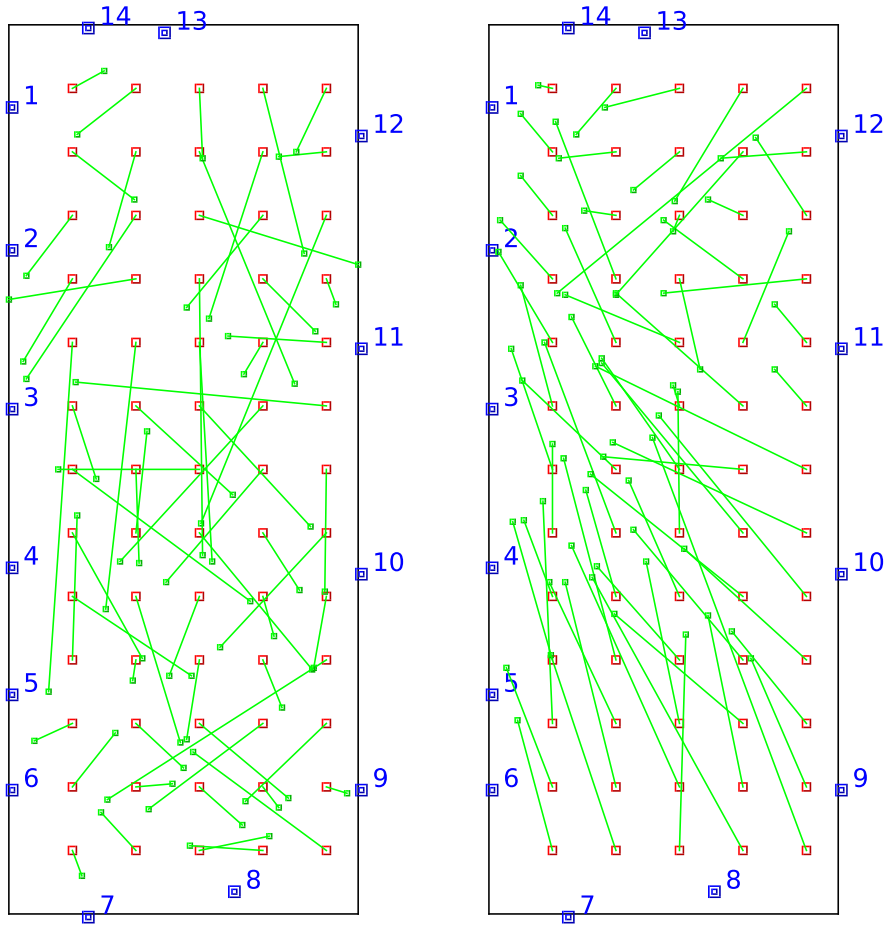

FIG. 7 Results of two weighted centroid measurements recorded in a laboratory test setup. *Larger double squares on the borders* indicate positions of radio transmitters (iBeacons). The regular grid marked with *squares* indicates the real positions of measurements. The spacing of these pints is 2 m. The size of the marked frame is about 28×11 m. The vectors starting at the test points indicate the calculated position. Both shown graphs differ in pose of the test person: (*right*) the test person rotated around herself while performing a measurement, (*left*) the test person kept herself and the phone in a certain orientation, that is, facing the top border.

8 Summary

Selected RSSI-based positioning algorithms have been studied. The backgrounds of radio propagation and the Friis model and modifications have been recalled, and aspects of indoor propagation have been discussed. Based thereon, centroid and weighted centroid methods have been explained. They form the base for more sophisticated approaches like fingerprinting. The fingerprinting method, the radio map concept and mathematic backgrounds like common similarity measures have been described.

In Chapter 5 the FCWC algorithm is described. It uses a radiomap for calculation of AP positions, which are then used for weighted centroid. This algorithm and the SPCF algorithm, which is a typical radiomap-based fingerprinting algorithm, have been described in detail and the performance of the algorithms has been assessed by comparing against each other. Furthermore, the algorithms participated at the IPIN EvAAL 2015 competition on indoor localization (offline track) (EvAAL: Evaluating AAL Systems through Competitive Benchmarking, 2015), which allowed to compare also against different approaches of other competitors, based on a very large dataset spawning four university buildings. Analysis showed that both algorithms well competed with other approaches, details have been outlined in Section 6.4.

Further examples of RSSI-based methods described are, for example, a BLE setup employing weighted centroid, and a probability map-based setup using proprietary 868 MHz communications of smart watches. Both algorithms are described in detail and can be regarded as alternative approaches, mostly useful in small-scale deployments where additional infrastructure is acceptable.

References

Bahl, P., Padmanabhan, V.N., 2000. RADAR: an in-building RF-based user location and tracking system. In: INFOCOM, vol. 2, pp. 775–784.

Berkvens, R., Weyn, M., 2015. Mean mutual information of probabilistic Wi-Fi localization. In: 2015 International Conference on Indoor Positioning and Indoor Navigation (IPIN).

Choi, S., Jaehyun, Y.O.O., Kim, H.J., 2015. Machine learning for indoor localization: deep learning and semi-supervised learning. In: 2015 International Conference on Indoor Positioning and Indoor Navigation (IPIN).

EvAAL: Evaluating AAL Systems through Competitive Benchmarking, 2015. EvAAL-ETRI on-site and off-site indoor localization competition in conjunction with IPIN 2015. Available from: http://evaal.aaloa.org/ (Accessed 20 June 2018).

Friis, H.T., 1946. A note on a simple transmission formula. Proc. IRE 34 (5), 254–256. ISSN 0096-8390. https://doi.org/10.1109/JRPROC.1946.234568.

Gau, Y.H., Chu, H.C., Jan, R.H., 2008. A weighted multilateration positioning method for wireless sensor networks. Int. J. Pervasive Comput. Commun. 3 (3), 289–303. https://doi.org/10.1108/17427370710856246.

Hashemi, H., 1993. The indoor radio propagation channel. Proc. IEEE 81 (7), 943–968. ISSN 0018-9219. https://doi.org/10.1109/5.231342.

Kang, W., Han, Y., 2015. SmartPDR: smartphone-based pedestrian dead reckoning for indoor localization. IEEE Sensors J. 15 (5), 2906–2916. ISSN 1530-437X. https://doi.org/10.1109/JSEN.2014.2382568.

Knauth, S., Koukofikis, A., 2016. Smartphone positioning in large environments by sensor data fusion, particle filter and FCWC. In: 2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp. 1–5.

Knauth, S., Jost, C., Klapproth, A., 2009. iLoc: a localisation system for visitor tracking and guidance. In: Proc. 7th IEEE Int. Conf. on Industrial Informatics INDIN2009, Cardiff, UK.

Knauth, S., Andrushevich, A., Kaufmann, L., Kistler, R., Klapproth, A., 2013. The iLoc+ ultrasound indoor localization system for AAL applications at EvAAL 2012. In: Chessa, S., Knauth, S. (Eds.), Evaluating

AAL Systems Through Competitive Benchmarking, vol. 362, Communications in Computer and Information Science. Springer, Berlin, Heidelberg, pp. 83–94.

Knauth, S., Griese, T., Tran, Y., Ortega, A.B., 2014. Towards smart watch position estimation employing RSSI based probability maps. Proc. First BW-CAR Baden-Württemberg CAR Symposium on Information and Communication Systems (SInCom 2014), Furtwangen, Germany, ISBN 978-3-00-048182-6, pp. 75–78.

Knauth, S., Kaufmann, L., Andrushevich, A., Kistler, R., Klapproth, A., 2015a. Evaluating the iLoc indoor localization system: competition outcomes and lessons learned. J. Ambient Intell. Smart Environ. 7, 287–300.

Knauth, S., Storz, M., Dastageeri, H., Koukofikis, A., Mäher-Hipp, N.A., 2015b. Fingerprint calibrated centroid and scalar product correlation RSSI positioning in large environments. In: 2015 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp. 1–6.

Mautz, R., Ochieng, W.Y., Brodin, G., Kemp, A., 2007. 3D wireless network localization from inconsistent distance observations. Ad Hoc Sensor Wirel. Netw. 3 (2–3), 140–170.

Moreira, A., Nicolau, M.J., Meneses, F., Costa, A.D., 2015. RTLS@UM—WiFi fingerprinting competition at IPIN 2015. In: 2015 International Conference on Indoor Positioning and Indoor Navigation (IPIN).

Philips Lighting Holding B.V., 2017. Available from: http://www.lighting.philips.com.au/systems/themes/led-based-indoor-positioning (Accessed 20 June 2018).

Rappaport, T.S., 2002. Wireless Communications: Principles and Practice, Prentice Hall Communications Engineering and Emerging Technologies Series. Prentice Hall PTR. ISBN 9780130422323. Available from: https://books.google.de/books?id=TbgQAQAAMAAJ.

Smith, A., Balakrishnan, H., Goraczko, M., Priyantha, N.B., 2004. Tracking moving devices with the cricket location system. In: 2nd International Conference on Mobile Systems, Applications and Services (Mobisys 2004), Boston, MA.

Torres-Sospedra, J., Montoliu, R., Martínez-Usó, A., Arnau, T.J., Avariento, J.P., Benedito-Bordonau, M., Huerta, J., 2014. 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, Busan, Korea.

Torres-Sospedra, J., Montoliu, R., Trilles, S., Belmonte, Ó., Huerta, J., 2015. Comprehensive analysis of distance and similarity measures for Wi-Fi fingerprinting indoor positioning systems. Expert Syst. Appl. 42 (23), 9263–9278. ISSN 0957-4174. https://doi.org/10.1016/j.eswa.2015.08.013. Available from: http://www.sciencedirect.com/science/article/pii/S0957417415005527.

Torres-Sospedra, J., Moreira, A., Knauth, S., Berkvens, R., Montoliu, R., Belmonte, O., Trilles, S., Nicolau, M.J., Meneses, F., Costa, A., Koukofikis, A., Weyn, M., Peremans, H., 2017. A realistic evaluation of indoor positioning systems based on Wi-Fi fingerprinting: the 2015 EVAAL-ETRI competition. J. Ambient Intell. Smart Environ. 9 (2), 263–279. https://doi.org/10.3233/AIS-170421.

UJIIndoorLoc Data Set at UCI Machine Learning Repository, 2014. Available from: https://archive.ics.uci.edu/ml/datasets/UJIIndoorLoc (Accessed 20 June 2018).

Ward, A., Jones, A., Hopper, A., 1997. A new location technique for the active office. IEEE Pers. Commun. 4 (5), 42–47.

Willemsen, T., Keller, F., Sternberg, H., 2015. A topological approach with MEMS in smartphones based on routing-graph. In: 2015 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp. 1–6.