WIRELESS COMMUNICATIONS AND MOBILE COMPUTING
Wirel. Commun. Mob. Comput. 2015; **15**:1342–1354
Published online 23 August 2013 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/wcm.2416

RESEARCH ARTICLE

Pseudo-3D RSSI-based WSN localization algorithm using linear regression

Frank Vanheel[1]*, Jo Verhaevert[1], Eric Laermans[2], Ingrid Moerman[2] and Piet Demeester[2]

[1] Faculty of Applied Engineering Sciences, University College Ghent, Ghent, Belgium
[2] Information Technology (INTEC), Ghent University - iMinds, Ghent, Belgium

ABSTRACT

Receiver Strength Signal Indication based Wireless Sensor Networks offer a cheap solution for location-aware applications. For a final breakthrough these systems need fast deployment and easy auto-configuration. In this study, we use the real-life iMinds test bed to expand a two-dimensional localization algorithm to the pseudo third dimension with very low additional computational time. Our experiments show that this fast three-dimensional algorithm has no outliers and avoids manual calibration. Our algorithm has lower position errors than a maximum likelihood algorithm with a mean square error cost function. Furthermore, with non-parametric statistical tests, we show that our previously designed two-dimensional preprocessing performs equally well in pseudo-three dimensions: the preprocessing reduces the position error in a statistically significant way. Copyright © 2013 John Wiley & Sons, Ltd.

KEYWORDS

Algorithm design and analysis; Correlation and regression analysis; Wireless sensor networks; Localization

***Correspondence**

Frank Vanheel, Faculty of Applied Engineering Sciences, University College Ghent, Ghent, Belgium.
E-mail: frank.vanheel@intec.ugent.be

1. INTRODUCTION

The amount of location-aware applications is still booming. The Receiver Signal Strength Indication (RSSI) based approach offers a cheap solution for the localization problem [1]. Indeed, because all sensors inherently need their own RSSI for demodulation, no additional hardware is needed. Localization in the presence of (indoor) multipath fading, however, remains a challenging task [2]. Statistical methods, like maximum likelihood (MLH) estimators [3] and Bayesian estimators [4] are widely used to improve the accuracy of the position. In [5], we presented an alternative statistical method: Linear Regression based Fast Localization Algorithm (LiReFLoA). This automated method optimizes and calibrates two-dimensional experimental data before offering it to our positioning tool. That tool uses the accuracy of the regression model to eliminate measurements with too much multipath fading before executing the final weighted multilateration process. Three-dimensional localization usually requires more (at least four) anchors (nodes knowing their own position) and in a multilateration algorithm, the computational cost rises exponentially with the number of anchors [6]. In the latter study, the authors further use existing two-dimensional approaches to simplify the complexity of pseudo-3D localization. In this paper, we follow this approach and expand LiReFLoA to obtain a fast pseudo-3D algorithm P3DLiReFLoA with the same number of anchors. Execution times barely change, enabling real-time localization. Pseudo-3D algorithms use two-dimensional projection techniques to find an object in a three-dimensional space.

This paper is organized as follows: In Section 2, related work is described. The hardware is described in Section 3. Our proposed algorithm can be found in Section 4. Section 5 follows with test results. The pseudo-3D algorithm is compared with the two-dimensional algorithm and with the more conventional minimum mean square error (MMSE) MLH algorithm. Finally, in Section 6, conclusions are drawn.

2. RELATED WORK

In [7], a survey of different application areas, ranging from military to civilian, is accompanied by their respective specific needs. A good starting point of the study of localization algorithms can be found in [8–12]. Our

work focusses on experimental RSSI-based Wireless Sensor Networks (WSN) indoors localization as in [13–15]. The shortage of experimental results obtained from real indoor test beds as outlined in [13] is well known. The latter work shows many similarities with ours. Like these authors, we also present a new localization algorithm. The environments, however, are difficult to compare because our test bed is larger (1512 vs. 23.2 square meter) with the same number of anchors (12 anchors). This results in an anchor density of only 0.008 (vs. 0.517) anchors per square meter. Furthermore, our automated calibration method is able to manipulate more measurements. During the selection and calibration method, each of the 41 nodes transmits 240 packets to the other nodes. The corresponding RSSI-measurements are reported and averaged. More than 380 000 RSSI-measurements are manipulated (somewhat less than $41 \times 40 \times 240$, because not all packets were above the noise floor of the receiver). This is an order of magnitude higher than 12 240 RSSI measurements reported in [13]. This increases the accuracy of the rough measurements, because the fast fading variation is averaged out.

The offline phase of the statistical indoor localization method, described in [16], is based on a local regression[17] fitting method to build a large RSSI database (called radio map) containing the distribution of the signal strength received at each known location. Local regression divides the independent statistical RSSI variable in small intervals and performs a regression on these binned data intervals. This offline phase tries to capture the complete distribution of the RSSI-distribution. Next, an online phase involves an MLH procedure on the distribution and the measured signal strength. A time-consuming bootstrapping method resamples the data (typically more than 1000 repetitions are needed), and gives 95% confidence intervals for the estimated position. Our work is also based on statistics, but takes a completely different approach: the underlying physical (and widely accepted [18,19]) relationship between the RSSI and the logarithm of the distance results in a regression that is simpler, because it is linear in the complete RSSI-variable and does not need data binning. Furthermore, our algorithm requires no radio mapping: the knowledge of two parameters (slope and intercept of the regression) is sufficient for the estimation of the position, further reducing the execution time. In this paper, statistics are also used for comparing results with non-parametric hypothesis testing, where no assumption needs to be made about the distribution of the position error. To our knowledge, this has not been encountered in WSN localization yet. More traditional research uses the cumulative distribution function (cdf) of the position error as well as parametric statistical metrics (mean value, average value, and standard deviation) to measure the localization performance [15]. Outliers can affect these parametric parameters substantially and make the tests and conclusions less reliable.

Very few authors [13,20] calibrate the propagation parameters to their individual values. In previous work [5], we used linear regression techniques to automate the selection and the individual calibration of the anchors. Here, this paper uses the same technique, but applied on a pseudo-3D algorithm.

The MLH is widely accepted in WSN positioning. A cost function is either minimized [21] or maximized [22–25] to find the most likely position. In [25], a linear regression based cost function has been compared with three other cost functions. In section 4, the results of our algorithm are compared with the most conventional widely used MMSE function.

Three-dimensional indoor positioning is complex and requires a combination of technologies. In [26], a three-dimensional algorithm is presented combining RSSI, time of arrival and sophisticated three-dimensional ray tracing. Ray tracing, which is a widely accepted technique for genuine-three-dimensional positioning, is based on geometrical optics. It can be applied as an approximate method for estimating the levels of high-frequency electromagnetic fields [27]. With the knowledge of the three-dimensional layout of the building and the used materials, path losses can be predicted. With this path loss, the distances can be calculated. Although this time-consuming task can be performed by the use of software tools as in [28], this procedure remains tedious. Therefore, we will not follow this methodology in this paper.

RADAR-based localization systems [29] and their two-dimensional fingerprinting method is widely known: in a time-consuming training phase, a database is filled with RSSI-measurements and in the online phase, a measurement is matched with these previously stored measurements. This two-dimensional fingerprinting method can be expanded to the third dimension. In a dynamically changing environment (e.g., changes in the position of furniture and presence of persons), however, the time-consuming training phase needs to be redone in order to get accurate results [30,31]. Therefore, [30] proposes an artificial neural network (*ANN*) incorporating not only a dynamic fingerprint, but also databases using a linear regression-based tree model mining technique. This approach trades in lack of accuracy for complexity. Three-dimensional fingerprinting incorporates not only RSSI, but also temperature, humidity, and light fingerprints [31]. In most cases, simple localization algorithms, like Weighted Centroid Location [13] are more robust against the variability of the investigated parameters [31].

Another three-dimensional localization system requires a full three-dimensional deployment. At least the double amount is then required: one node on the ceiling and one node on the floor. This solution is mostly used in multistory buildings as in [32]. It could be useful in buildings with extremely high ceilings. In these cases, the anchors are used efficiently. In most practical situations, however, vertical resolution is not always a primordial matter: for example, in a museum information system, it is more important to know that a person is in front of a particular painting, than the information that he is standing or kneeling. Therefore, most localization algorithms do not take the

third dimension into account and apply the procedures just as in the two-dimensional localization. In simple RSSI-based algorithms (like LiReFLoA [5]), however, the two-dimensional and three-dimensional propagation paths can differ significantly. When two-dimensional propagation paths are used for calibration in a three-dimensional environment, large errors occur, spoiling the accuracy of the underlying model. Therefore, [6] proposes a complexity-reduced multilateration for three-dimensional localization using super anchors (anchors with pairwise positions whose coordinates only differ in the z-axis, i.e., their height). In this paper, we extend this approach of reduced complexity with very low additional computing time: no super anchors having two antennas are needed because our three-dimensional calibration is performed with a mobile node, just beneath the two-dimensional selected anchors. Very few extra calibrations are needed.

3. THE HARDWARE

Our hardware consists of the iMinds (formerly IBBT) iLab.t Wireless Lab (W-iLab.t) test bed. More about this test bed can be found in [5]. Only the third floor is used in this paper. On this $16.8 \times 90 \times 2.65$ m floor, there are 41 active nodes, represented by small green circles in Figure 1.

Figure 1 shows a floor plan of the third floor. The Tmote Sky nodes are equipped with a CC2420 radio chip operating in the 2.4 GHz frequency band. The nodes are mounted with the top of the printed circuit board down in a plane, approximately 0.12 m below the highly conductive ceiling. The integrated inverted-F antenna is mounted horizontally with its longest leg parallel to the longest wall. An anchor selection results in the best available quality (this will be explained in section 4.2.1) anchors at a given time. These are represented by the black squares. In this environment, there is a lot of constructive multipath fading, because of the presence of long corridors. Because the walls mainly consist of plywood (these walls are represented by blue solid lines), the presence of wall attenuation is rather limited. In this typical office environment, there is also furniture: for example, bookshelves that are about 2 m tall. There are also two microwave ovens, situated in the right-hand side and left-hand side concrete zone of the building. A previous study revealed that multipath fading is by far the most annoying factor in this indoor environment.

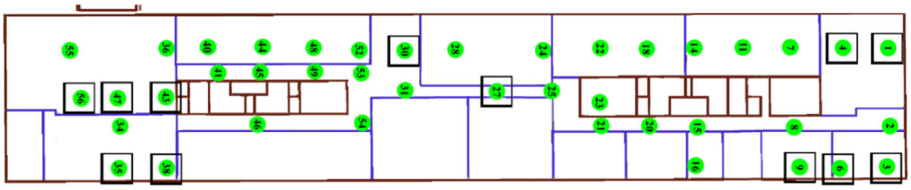

Figure 1. Position of sensor nodes on the third floor of the iMinds office building. Plywood walls are presented by blue solid lines. The solid brown lines are concrete wall. The black squares are the selected anchors. After the linear regression based selection between the already deployed nodes, these anchors are calibrated (section 4.2.1).

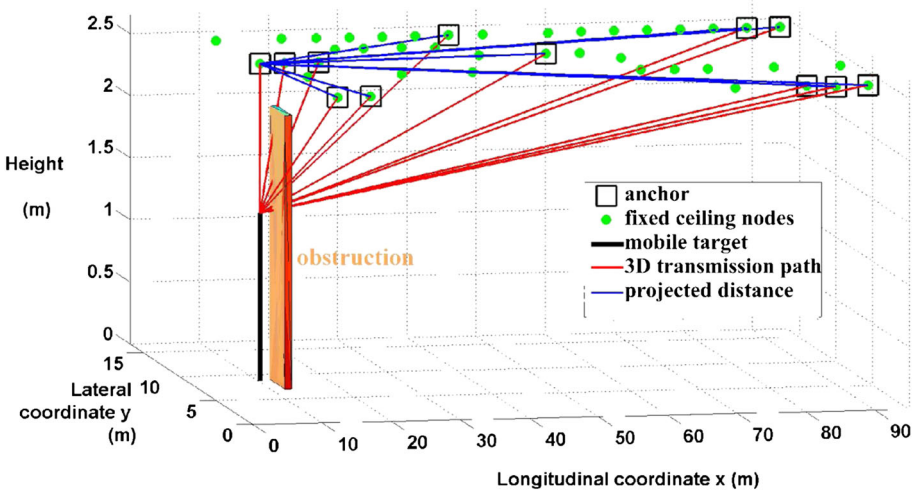

Figure 2. Schematic overview of the pseudo-3D environment. For comparison reasons, two tests are performed: the two-dimensional test uses the plane of the fixed nodes and the pseudo-3D test uses a mobile node in the plane 1.2 m below this plane of the fixed nodes.

4. TEST SETUP

4.1. Changing from two-dimensional distances to pseudo-3D distances

Figure 2 presents the schematic overview of the pseudo-3D environment. Again, the fixed ceiling nodes are shown as green circles. The remainder of this section compares a two-dimensional algorithm with a pseudo-3D algorithm. Therefore, two tests are performed. The first one is two-dimensional: every fixed node sequentially broadcasts 240 packets to all other nodes. Every packet consists of 100 bytes, transmitted at a symbol rate of 62.5 ksymbols/s. The inter packet delay is 25 ms. Transmission is at channel 26 in order to avoid Wi-Fi interference. Upon swapping sending nodes, the test bed is idle for at least 3.5 s. The test is performed at transmit power levels of 0 dBm. Assuming a receiver sensitivity of −92 dBm, this power level corresponds to a distance range of approximately 82 m [5], or almost the complete building. The test bed only counts valid RSSI-measurements: our software detects corrupted and lost packets; these packets are excluded in the averaging process. Fast fading fluctuations can be averaged out by considering large number of RSSI-readings levels. Microwave ovens with a traditional power supply only radiate at the positive peaks of the mains supply. Our software is able to distinguish valid and invalid packets, therefore RSSI-measurements from packets sent at the negative peaks of the main supply are recorded correctly. For this two-dimensional test, the Euclidean distances are taken into account. In Figure 2, these distances are represented by blue solid lines. The second test is pseudo-3D. Now, a test person activates a mobile node at a constant height from the ceiling. This test was performed just 1.2 m beneath each fixed node. The mobile node is hand carried with the antenna parallel to the antenna of the fixed node above it. The swapping time of the sending nodes is different, but the other conditions remain the same. All but one pseudo-3D distance (represented by solid red lines in Figure 2) are approximated by their projection in the plane of the fixed nodes (represented by solid blue lines). Only for the closest distance, the exact distance is used (1.2 m). Although this introduces an error (of maximum 3.5%) in the calibration of the distances, this error is acceptable because the nodes are on a relatively coarse grid of approximately 4.5 m. This error is an order of magnitude smaller than the error on the multipath faded RSSI-measurements [5].

4.2. The algorithm

Figure 3 gives a flowchart of the positioning algorithm. It contains both the two-dimensional and the new pseudo-3D steps. For a full understanding, the principles of the two-dimensional algorithm are explained in a first subsection. In a next subsection, the pseudo-3D part is presented. More details about the two-dimensional algorithm can be found in [5].

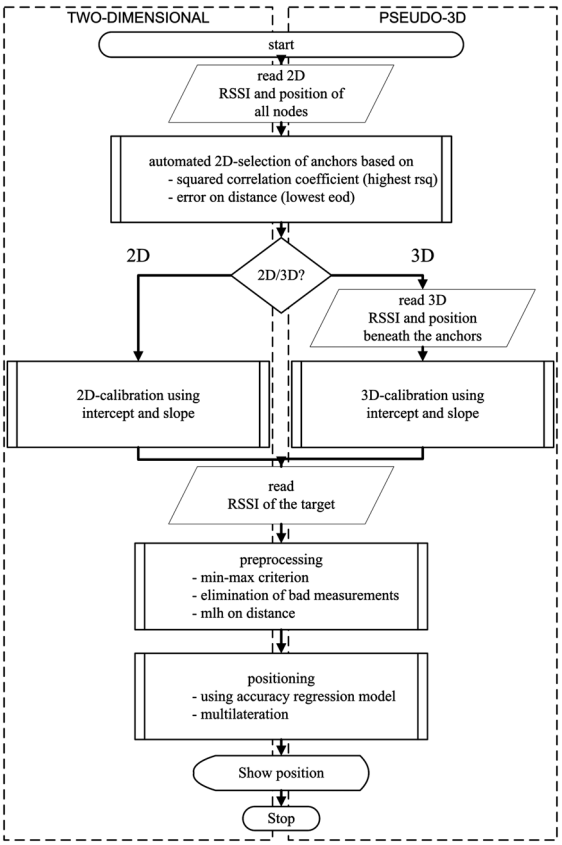

Figure 3. This flowchart illustrates the similarities and the differences between the two-dimensional and three-dimensional algorithm.

4.2.1. The two-dimensional algorithm.

The two-dimensional algorithm, called LiReFLoA, is based on linear regression tools [33] to select and calibrate anchors, preprocess the measurements, and locate the target.

It is a simple RSSI-based localization algorithm, assuming an already deployed two-dimensional wireless sensor network, which is a realistic scenario for future dynamic wireless indoor environments.

In Figure 3, the leftmost procedures are followed. This path leads through the following steps: RSSI-measurements for all nodes, a 2D-selection of anchors, a 2D-calibration, the target RSSI-measurement, a 2D-preprocessing, and a 2D-positioning. With the two-dimensional test results, a linear regression is performed for each sending node and an error_on_distance (eod) parameter is computed. Eod is defined as twice the estimated standard deviation on the (logarithmic) distance (after an axis swap) [5]; it is a measure of how close the measurements are to the regression line. This parameter and the square of the correlation coefficient (rsq) are used to select the best anchors. An rsq of zero indicates that there is no linear fit between the RSSI and the logarithm of the

distance, whereas an rsq of 1 implies that all points lie on a straight line.

Equations (1) and (2) give an expression of the rsq and eod, respectively:

$$rsq(i) = \frac{slope^2(i) \times \sum_j \left(log_{10}d_{i,j} - mean_j\left(log_{10}d_{i,j}\right)\right)^2}{\sum_j \left(RSSI_{i,j} - mean_j\left(RSSI_{i,j}\right)\right)^2} \quad (1)$$

$$eod(i) = 2\sqrt{\frac{\sum_j \left(log_{10}d_{i,j} - mean_j\left(log_{10}d_{i,j}\right)\right)^2 - slope_{sw}^2(i)\sum_j \left(RSSI_{i,j} - mean_j\left(RSSI_{i,j}\right)\right)^2}{\#\text{ (received RSSI measurements from node i above noise floor)} - 2}} \quad (2)$$

Slope$(i)_{sw}$ is the slope of the regression line of node i after the axis swap, and $d_{i,j}$ is the exact distance between node i and node j. In our building, a choice of ten best correlated and two nodes having the lowest eod, results in a selection of twelve anchors. This corresponds to an anchor density of less than 0.008 anchors per square meter. The nodes with the highest rsq are found in the extremities of the building, not in the corridors, as obtained with a previous anchor selection method (where we selected the best anchor nodes from a set of active nodes based on the linearity of their calibrated path loss model and on the error_on_distance) [5]. In Figure 1 and Figure 2, the anchors are marked with a small black square. Next, the intercept point (defined as the received power at a distance of 1 m) and slope (defined as the rise of the RSSI over the run of the (logarithmic) distance) of the regression lines are used to individually calibrate these anchors. The two-dimensional selection of anchors is a fast procedure. All nodes are already deployed just below the ceiling. Therefore, transmitting and receiving the packets, finding the nodes with the highest correlation coefficient, and calculating the estimated standard deviation on the regression quickly result in the best available anchors at a given time. After the calibration, all location information of the unselected nodes is discarded. The anchors report the respective (valid) RSSI-measurements. Next, the position of the target is estimated with the RSSI-measurements (from packets transmitted by the target and received by the anchors), and the position of the calibrated anchors. The preprocessing includes a min-max algorithm, an elimination of bad measurements (based on the accuracy of the regression model) and a MLH algorithm on the distance. After the preprocessing, the position is calculated.

4.2.2. The pseudo-3D algorithm.

The pseudo-3D algorithm is based on this two-dimensional algorithm. In Figure 3, the rightmost path is taken. The difference between the two-dimensional and pseudo-3D algorithm is the transmission of the RSSI-beacons (triggered by the person to be located) beneath the two-dimensionally selected anchors and the calibration procedure. The remainder of this section explains this approach. A full three-dimensional counterpart of the fast two-dimensional anchor selection procedure requires either the automatic height variation of the nodes or, alternatively a physical walk in the building with a mobile target transmitting at many known places. Because this full three-dimensional selection is neither cheap nor simple, the two-dimensional selection is kept in our pseudo-3D algorithm.

The two-dimensional calibration of the anchors requires linear regression between the measured RSSI and the logarithm of the distances. The (unknown) distances can be calculated fast by using the intercept and the slope of this regression line. The regression lines change when the target is moved away from the two-dimensional plane of the anchors, basically because there is a larger attenuation because of the presence of furniture. A human body in the vicinity of the transmitter has a comparable attenuating effect on the communication link. Therefore, an easy pseudo-3D calibration is performed. It is not necessary to use the complete test data of the second test. Only the test bed RSSI values of sending mobile node 1.2 m beneath the twelve anchors are needed. A linear regression between the RSSI-values and the logarithmic distance is executed for each of the twelve mobile node positions. The new slopes and intercepts are used for pseudo-3D calibration of the anchors and the distance calculations.

The pseudo-3D selection and calibration has several advantages:

(1) A good two-dimensional anchor is also a good pseudo-3D anchor. This is illustrated in Figure 4, where the rsq-values of the two-dimensional regression (in the first test, i.e., on the ceiling) are plotted versus the pseudo-3D regression (in the second test, i.e., with one mobile node for regression). The large circles are the ten best pseudo-3D correlated ones, and the smaller circles represent the others. The large '+'-signs denote the ten best two-dimensional correlated ones and the small '+'-signs are the others. The blue dash-dotted horizontal line corresponds with a two-dimensional rsq of 0.76 and the red vertical dash-dotted line corresponds with a pseudo-3D rsq of 0.76. Above this 0.76 value, the best correlated nodes are found. The diagonal dashed line represents the bisector of the pseudo-3D

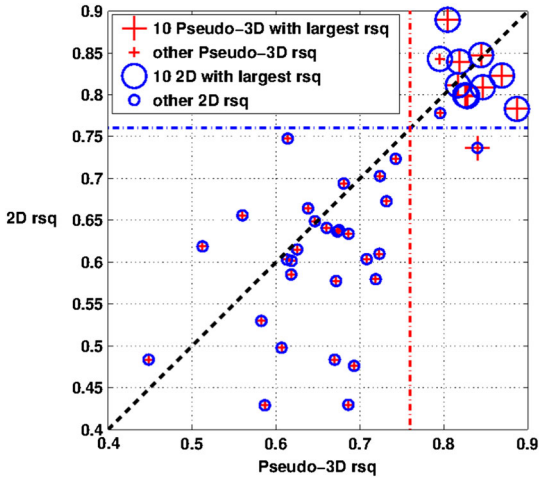

Figure 4. Comparison of the correlation coefficient of the two-dimensional and pseudo-3D tests. Two-dimensional nodes with high rsq are also pseudo-3D nodes with high rsq.

rsq and two-dimensional rsq axes. The nodes with largest rsq are in the upper right hand-side part of the figure. A pseudo-3D selection instead of a two-dimensional one would have resulted in only one different anchor at a cost of a higher computational time (because measurements beneath each node are needed). Therefore, the two-dimensional selection of the anchors is kept in P3DLiRefLoA, and there is no extra time needed for a dedicated pseudo-3D anchor selection. Please note that a pseudo-3D correlation coefficient of 0.795 (instead of 0.840) is still a very good value. We further observe that the pseudo-3D rsq-values are somewhat higher than the two-dimensional rsq-values: there are 12 anchors at the right-hand side of the vertical line at 0.76 and only 11 anchors above the horizontal line at 0.76. Furthermore, there are 11 points above the bisector (positive ranks), none of the points are on the bisector (no ties) and 30 points below the bisector (negative ranks), therefore a Wilcoxon signed-rank test rejects the null-hypothesis – that the pseudo-3D rsq-distribution equals the two-dimensional rsq-distribution.

(2) Thanks to the automated selection and two-dimensional calibration of nodes, it can dynamically select the best available two-dimensional anchors at a given time. When anchors fail, the algorithm can quickly reselect and calibrate other anchors. In P3DLiReFLoA, this advantage is kept. Selecting uniformly distributed anchors could have resulted in an anchor that is down. In our sparse anchor density environment, each selected (and high quality) anchor is needed.

(3) The measurements beneath an anchor result in a RSSI at a very short distance. This nearby information is very useful, because it is at the beginning of the regression line. Without this measurement, the regression line would have been extrapolated, resulting in large errors [33].

(4) The complete extension to the pseudo-third dimension requires minimal effort and computation time: only on twelve anchors a linear regression between the logarithm of the distance and the RSSI is needed to obtain the new propagation constants for the calibration. This can easily be implemented: the person to be located first takes a walk through the building. Under the two-dimensionally selected anchors, he pushes a button, and triggers the pseudo-3D calibration. Each time, the algorithm calibrates the corresponding anchor. No fixed route can be postulated; this is dependent on the dynamic anchor selection and not always the same anchors are selected. The order in which the anchors are calibrated is not important, only the location where the device is triggered (below the correct anchor) matters. Therefore, we recommend that the anchors are calibrated in numerical order. After this calibration, the target can be located anywhere in the building.

The algorithm can calculate the (x,y)-position for any point in the plane 1.2 m beneath the anchors, starting from the position of the anchors and their pseudo-3D propagation parameters. In this paper, the moving target sends packets beneath the deployed two-dimensional nodes, enabling a good comparison with the two-dimensional parameters. The remainder of the two-dimensional algorithm in Figure 3 is unchanged in the pseudo-3D algorithm.

5. RESULTS

5.1. Comparison of the two-dimensional (first test) and pseudo-3D calibration (second test)

Figure 5 compares the two-dimensional calibration with the pseudo-3D calibration for a selected anchor (node 56) at the left hand side of the building (Figure 2). The small circles are the recorded RSSI-distance pairs for all nodes when the selected anchor is sending (two-dimensional values). The '+'-values are the recorded RSSI-distance pairs for all nodes when sending with the mobile node beneath that anchor (pseudo-3D values). The solid and the dashed line show the corresponding regression lines. The pseudo-3D regression line has a higher attenuation at low distance levels and a flatter slope. At high distances, the difference in RSSI decreases. This can be explained by the fact that the attenuation of furniture (or other obstacles) is more pronounced at low to medium distances. This is in agreement with the angle dependency of the attenuation factors found in [34]: When electromagnetic radiation is obliquely incident on a wall or floor, less power will be transmitted through the wall than would occur at normal incidence. Nodes in the neighborhood of the target incident at larger angles than nodes that are further away. Although only

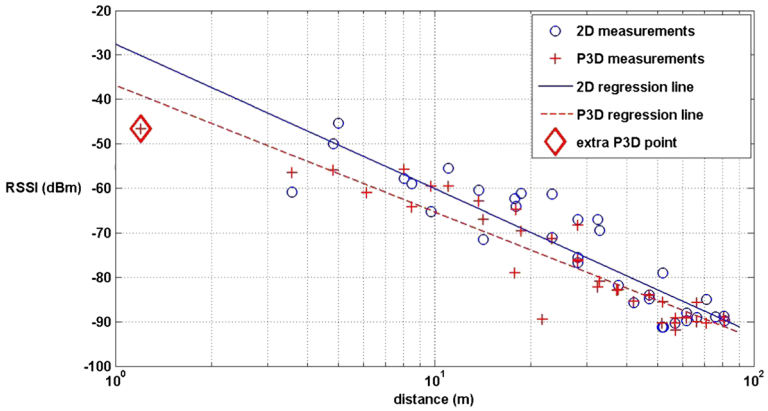

Figure 5. Pseudo-3D regression lines not only have higher attenuation at low distances, but are also less steep than two-dimensional ones. The pseudo-3D calibration procedure has an extra measurement at the beginning of the regression line.

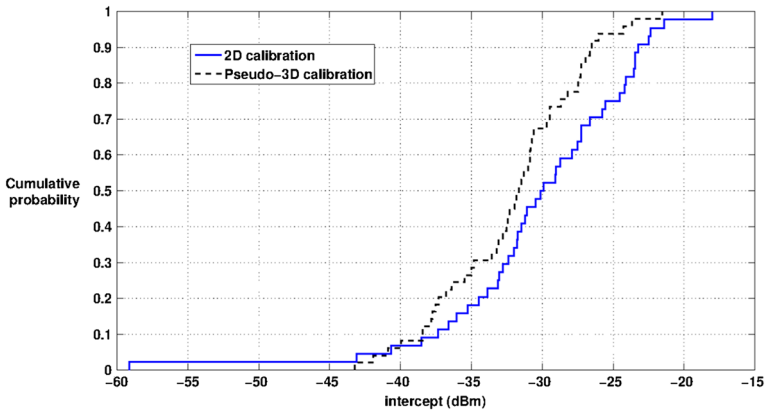

Figure 6. The pseudo-3D calibration results in higher attenuation at small distances (compared to two-dimensional calibration).

one typical regression comparison is shown here, this conclusion holds true for the vast majority of the nodes, as illustrated in figures 6 and 7. Figure 6 represents the cumulative distribution plot for the intercept for two-dimensional and pseudo-3D calibrations, respectively. Except for a few nodes (not being anchor nodes), the two-dimensional plot is at the right hand side of the pseudo-3D plot. Hence, this confirms the higher attenuation at low distances of the pseudo-3D calibration compared to the two-dimensional calibration. Figure 7 represents the cumulative distribution plot of the slope of the two-dimensional and pseudo-3D calibration of all nodes. Now, the two-dimensional plot is at the left hand side of the pseudo-3D plot for the vast majority of the nodes. Therefore, the two-dimensional calibration results in higher slopes.

Figure 8 represents a cdf plot of the error on distance for the two-dimensional and pseudo-3D calibration of the anchors. Only the restricted data set limited to the measurements of the anchors (two-dimensional) and the mobile node 1.2 m below the anchors (pseudo-3D) is considered here. Being defined as twice the estimated standard deviation of the (logarithmic) distance frequency distribution [5], the error on distance is a logarithmic value on the tolerances of the distances. For example, a value of 0.3 means that the tolerances on a distance are $10^{-0.3}$ or minus 50% and $10^{+0.3}$ or plus 200%. This figure illustrates that the anchors have a higher error on distance when pseudo-3D calibrated than when two-dimensionally calibrated. This difference in error on distance can be explained by the higher attenuation in presence of furniture. The tolerances are not only used in the preprocessing step (both elimination of bad measurements and MLH on the distance), but also in the positioning step for a decision on the amount of constructive multipath fading present. It therefore is important to offer the right empirical error on distance to the algorithm: the two-dimensional eod for two-dimensional localization and the pseudo-3D eod for the pseudo-3D localization.

5.2. Comparison of our algorithm with a more conventional algorithm

The three-dimensional algorithm is tested: the algorithm is executed, and the results are compared with the exact

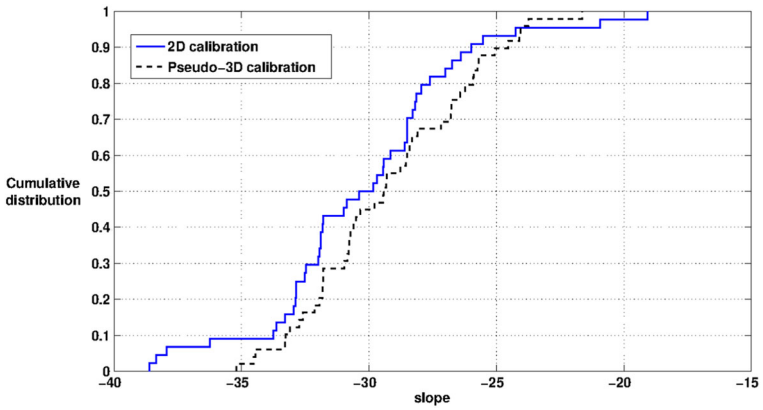

Figure 7. The two-dimensional calibration results in higher slopes (compared to pseudo-3D calibration).

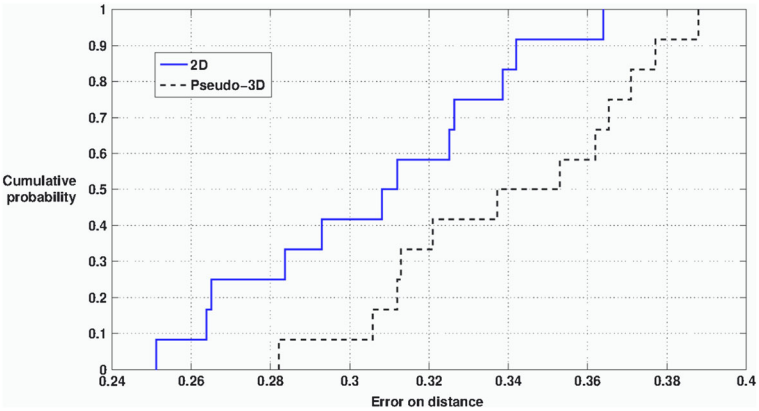

Figure 8. The pseudo-3D errors on distance of the anchors are larger than their two-dimensional ones.

position. For comparison, a MLH algorithm on the position is implemented using the same anchors and the MMSE cost function:

$$\arg\min_{x,y} \sum_{j \,\epsilon\, anchor\,(i)} \left(\tilde{d}_{i,j} - d_{i,j}\right)^2 \qquad (3)$$

where $d_{i,j}$ is the exact distance between anchor i and point j. The tilded $\mathrm{d}_{i,j}$ is the calculated distance (with the measured RSSI and the propagation constants). Like in our previous work, the grid method is used. Unlike the conjugate gradient algorithm [35], the grid method always finds the exact minimum and does not get stuck in local minima (sacrificing computational time). The grid plane is now formed 1.2 m below the vertical position of the anchors. The point in this plane with coordinates (x,y) that minimizes the sum of squared differences between the calculated and exact distances is the estimated position. With both algorithms, the position is calculated and compared with the exact distance. This is performed for both cases: with and without our preprocessing steps, which are based on a min-max criterion, elimination of bad measurements and a MLH on the position [5]. The results of the position errors of both algorithms are outlined in Table I and

Table I. Position error comparison between P3DLiReFLoA and pseudo-3D maximum livelihood algorithm with a mean square error.

	P3DLiReFLoA w/o (m)	P3DLiReFLoA with (m)	P3D MMSE w/o (m)	P3D MMSE with (m)
Upper outlier	-	-	22.47	17.56
Upper adjacent	19.00	13.37	17.67	12.36
Third quartile	9.80	8.48	10.33	7.68
Median	7.22	5.11	7.79	4.90
First quartile	3.49	2.91	4.70	3.50
Lower adjacent	1.19	0.46	1.79	1.22
Under outlier	-	-	-	-

Figure 9. Table I lists the outliers, adjacent (most extreme data points not considered outliers), and the first (25^{th} percentile), second (50^{th} percentile, median) and third (75^{th} percentile) quartile. The same information is given a graphical form with boxplots in Figure 9. On each box, the central mark is the median, the edges of the box are the first and third quartile, the whiskers extend to the upper and lower adjacent. All position errors that are 1.5 times the box size above the third quartile are deemed upper outliers, and all position errors that are 1.5 times the box size below the first quartile are considered under outliers. The outliers are plotted individually with a '+'-sign [33]. This Table I and Figure 9 reveal that

- there are no outliers with P3DLiReFLoA, in contract with MMSE.
- a MLH algorithm with a mean square error cost function has a higher position error median when our preprocessing is not applied.
- when our preprocessing is applied not only this median, but also the high percentiles of this algorithm are improved.
- none of the distributions are normal distributions:
  - all of them have larger upper tails than lower tails, introducing skewness in the distribution.
  - for MMSE, there are outliers, which are absent in normal distributions.
  - except for MMSE without the processing, the medians (the large horizontal lines in the box) are not in the center of the box. This means that the median of the distribution is different to the average.

Further interpretation of the test results is based on statistical inference. Because the position error is not normally distributed, Student's *t*-tests could lead researchers to draw incorrect conclusions. Non-parametric tests make no assumption on the distribution, and are a better option here. Thanks to the increased availability of software, these non-parametric statistical analyses are often found in medical research [36]. The Wilcoxon signed-rank test [37] is a non-parametric statistical hypothesis test for comparing two related samples, for example, before preprocessing and after. The null-hypothesis of a first Wilcoxon test – that the

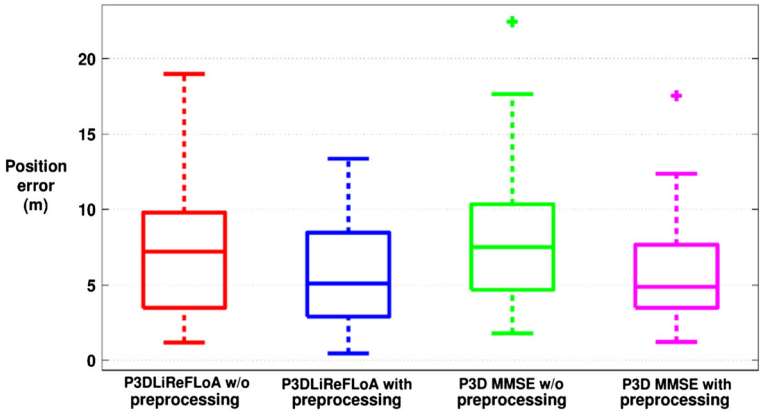

Figure 9. Boxplot of the position errors with and without (w/o) the preprocessing steps: at the left hand side for the position errors of the algorithm proposed in this paper and at the right hand side for the position errors of the minimum mean square error maximum likelihood algorithm. A maximum likelihood algorithm with a mean square error cost function has a higher position error median when our preprocessing is not applied.

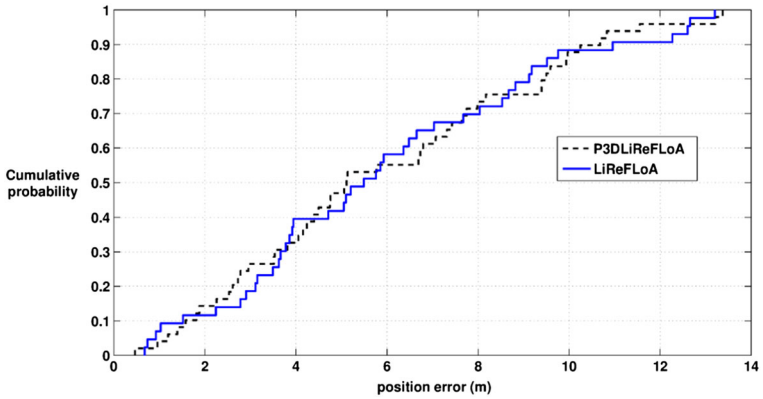

Figure 10. Comparison of the cdfplot of LiReFLoA and P3DLiReFLoA.

Table II. Comparison of the two-dimensional and pseudo-3D algorithm.

	LiReFLoA	P3DLiReFLoA
Localization	Two-dimensional	Pseudo-3D
Anchor selection	Two-dimensional	
Calibration	Two-dimensional	Three-dimensional on selected anchors
Position engine	The same	
Computation time	Fast	Slightly less fast: only quick additional calibration is required for a limited number of anchors
Accuracy		
Without preprocessing	Median 5.79m	Median 7.22m
With preprocessing	Median 5.49m	Median 5.11m
Cdfplot	Figure 10	
Wilcoxon p-values	Two-tailed p-value = 98%	

P3DLiReFLoA position error distribution after the preprocessing equals the P3DLiReFLoA position error distribution before the preprocessing – is rejected. The two-tailed *p*-value (defined as the probability that the test statistic is equal to or more extreme than the one observed under the null hypothesis) equals 0.2%. A second test is performed on MMSE with and without the preprocessing steps and results in a two-tailed *p*-value of less than 0.1%. The third Wilcoxon test follows, comparing P3DLiReFLoA with the preprocessing steps and MMSE without them, results in a two-tailed *p*-value of 2%. The first test concludes that there is a difference in position error for P3DLiReFLoA with and without the preprocessing. Test 2 shows that this is also the case for MMSE. This proves that our preprocessing has a positive effect on the positioning error. Furthermore, the third Wilcoxon test shows that the position errors of P3DLiReFLoA are significantly lower than those of the MMSE without our preprocessing.

5.3. Comparison of LiReFLoA and P3DLiReFLoA

This section compares the two-dimensional algorithm LiReFLoA of our previous work (first test, left-hand side of Figure 3) with the P3DLiReFLoA (second test, right-hand side of Figure 3) described in this paper. Testing the algorithm in their respective environment (2D for LiReFLoA and pseudo-3D for P3DLiReFLoA) gives comparable medians. Figure 10 further illustrates that also the distributions are very similar. Table II gives an overview of the comparison of LiReFLoA and P3DLiReFLoA. This table reveals that both algorithms use the same position engine.

Both the preprocessing and the positioning procedure are the same (as already outlined in Figure 3). Also the anchor selection remains unchanged. Although LiReFLoA calibrates these anchors in the two-dimensional plane of these anchors, a pseudo-3D calibration with a mobile node beneath these anchors is needed for P3DRiReFLoA. With an anchor density of 0.008 anchors per square meter, very little additional time is needed. A paired Wilcoxon test results in a two-tailed p-value of 98%. Therefore the null-hypothesis that P3DLiReFLoA performs equally well in a pseudo-3D environment as LiReFLoA in a two-dimensional environment is accepted.

6. CONCLUSIONS

This paper presents a new pseudo-3D localization algorithm, based on a fast two-dimensional algorithm. Only a quick recalibration is required for the limited number of anchors. Our empirical tests show that the position errors are lower than with a MLH algorithm with a mean square error cost function. Preprocessing of the data also reduces the position errors for the MLH algorithm in a statistically significant way.

REFERENCES

1. Tran-Xuan C, Vu VH, Koo I. Calibration Mechanism for RSS Based Localization Method in Wireless Sensor Networks, In *Proceedings of the 11th International Conference on Advanced Communication Technology, Vols I-III*, Gangwon-Do, South Korea, 2009; 560–563.
2. Awad A, Frunzke T, Dressler F. Adaptive Distance Estimation and Localization in WSN using RSSI Measures, In *DSD 2007: Proceedings of the 10th Euromicro Conference on Digital System Design Architectures, Methods and Tools*, Lubeck, Germany, 2007; 471–478, DOI: 10.1109/DSD.2007.4341511.
3. Li S, Meng MQH, Liang HW, You ZH, Zhou YJ, Chen WM. A Localization Error Estimation Method based on Maximum Likelihood for Wireless Sensor Networks, In *2007 IEEE International Conference on Mechatronics and Automation* Vols I-V, *Conference Proceedings*, Lubeck, Germany, 2007; 348–353, DOI: 10.1109/ICMA.2007.4303567.
4. Fontanella D, Nicoli M, Vandendorpe L. Bayesian Localization in Sensor Networks: Distributed

Algorithm and Fundamental Limits, In *2010 IEEE International Conference on Communication*, Cape Town, South Africa, 2010; 1–5, DOI: 10.1109/ICC.2010.5502618.

5. Vanheel F, Verhaevert J, Laermans E, Moerman I, Demeester P. Automated linear regression tools improve RSSI WSN localization in multipath indoor environment. *Eurasip Journal on Wireless Communications and Networking* 2011. DOI: 10.1186/1687-1499-2011-38.
6. Shih CY, Marrón PJ. COLA: Complexity-reduced Trilateration Approach for 3D Localization in Wireless Sensor Networks, In *2010 Fourth International Conference on Sensor Technologies and Applications (SENSORCOM)*, Venice, Italy, 2010; 24–32, DOI: 10.1109/SENSORCOM.2010.11.
7. Rantakokko J, Händel P, Marsten-Eklöf F. User Requirements for Localization and Tracking Technology: A Survey of Mission-specific Needs and Constraints, In *IPIN 2010: International Conference on Indoor Positioning and Indoor Navigation*, Zurich, Switzerland, 2010; 1–9, DOI: 10.1109/IPIN.2010.5646765.
8. Bal M, Shen WM, Ghenniwa H. Collaborative Signal and Information Processing in Wireless Sensor Networks: a Review, In *2009 IEEE International Conference on Systems, Man and Cybernetics (SMC 2009),* Vols 1-9, San Antonio, TX, USA, 2009; 3151–3156, DOI: 10.1109/ICSMS.2009.5346152.
9. Liu H, Darabi H, Banerjee P, Liu J. Survey of wireless indoor positioning techniques and systems. *IEEE Transactions on Systems, Man and Cybernetics, —Part C: Applications and Reviews* 2007; **37**: 1067–1080. DOI: 10.1109/TSMCC.2007.905750.
10. Koyuncu H, Yang SH. A Survey of indoor positioning and object locating systems. *International Journal of Computer Science and Network Security (IJCSNS)* 2010; **10**(5): 121–128.
11. Gu Y, Lo A, Niemegeers I. A survey of indoor positioning systems for wireless personal networks. *IEEE Communications Surveys & Tutorials* 2009; **11**: 13–32.
12. Pandey S, Agrawal P. A survey on localization techniques for wireless networks. *Journal of the Chinese Institute of Engineers* 2006; **29**(7): 1125–1148.
13. Pivato P, Palopoli L, Petri D. Accuracy of RSS-Based centroid localization algorithms in an indoor environment. *IEEE Transactions on Instrumentation and Measurement* 2011; **60**: 3451–3460. DOI: 10.1109/TIM.2011.2134890.
14. Zanca G, Zorzi F, Zanella A, Zorzi M. Experimental comparison of RSSI-based localization algorithms for indoor wireless sensor networks, In *REALWSN 2008: Proceedings of the Workshop on Real-World Wireless Sensor Networks*, Glasgow, Scotland, UK, 2008; 1–5, DOI: 10.1145/1435473.1435475.
15. Luo X, O'Brien WJ, Julien CL. Comparative evaluation of received signal-strength index (RSSI) based indoor localization techniques for construction jobsites. *Journal of Advanced Engineering Informatics* 2010; **25**(2): 355–363.
16. Gao CW, Yu Z, Wei YW, Russell S, Guan Y. A statistical indoor localization method for supporting location-based access control. *Mobile Networks & Applications* 2009; **14**: 253–263. DOI: 10.1007/S11036-008-0143-4.
17. Cleveland WS, Fitting DSJ. L-W. An approach to fitting analysis by local fitting. *Journal of the American Statistical Association* 1988; **83**: 596–610.
18. Hashemi H. The indoor radio propagation channel. *Proceedings of the IEEE 1993*; **81**: 943–968. DOI: 10.1109/5.231342.
19. Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs), 2006. *IEEE Std. 802.15.4-2006*.
20. Shirahama J, Ohtsuki T. RSS-based Localization in Environments with Different Path Loss Exponent for Each Link, In *2008 IEEE 67th Vehicular Technology Conference-Spring, Vols 1-7*, 2008; 1509–1513, DOI: 10.1109/VETECS.2008.353.
21. Patwari N, Hero AO, Perkins M, Correal NS, O'Dea RJ. Relative location estimation in wireless sensor networks. *IEEE Transactions on Signal Processing* 2003; **51**: 2137–2148.
22. Takashima M, Zhao D, Yanagihara K, Fukui K, Fukunaga S, Hara S, Kitayama KI. Location estimation using received signal power and maximum likelihood estimation in wireless sensor networks. *Electronics and Communications in Japan Part I-Communications* 2007; **90**: 62–72. DOI: 10.1002/ECJA.20359.
23. Yamada I, Ohtsuki T, Hisanaga T, Zheng L. An Indoor Position Estimation Method by Maximum Likelihood Algorithm using RSS, In *Proceedings of SICE Annual Conference*, 2007, Vols 1-8; 2918–2921.
24. Cho H, Kang M, Park J, Park B, Kim H. Performance Analysis of Location Estimation Algorithm in ZigBee, In *Proceedings of the 21st International Conference on Advanced Networking and Applications Workshops/Symposia, Vol 2*, Niagara Falls, Ontario, Canada, 2007; 302–306.
25. Vanheel F, Verhaevert J, Laermans E, Moerman I, Demeester P. A Linear Regression Based Cost Function for WSN Localization, In *2011 19th Interna-*

tional Conference on Software, Telecommunications and Computer Networks (SoftCOM), Split, Croatia, 2011; 1–5.
26. Lewandowski A, Wietfeld C. A Comprehensive Approach for Optimizing ToA-Localization in Harsh Industrial Environments, In *2010 IEEE-Ion Position Location and Navigation Symposium Plans*, Indian Wells, CA, USA, 2010; 759–768, DOI: 10.1109/PLANS.2010.55077255.
27. Sarkar TK, Ji Z, Kim KJ, Medouri A, Salazar-Palma M. A survey of various propagation models for mobile communication. *IEEE Antennas and Propagation Magazine* 2003; **45**: 51–82. DOI: 10.1109/MAP.2003.1232163.
28. Plets D, Joseph W, Vanhecke K, Tanghe E, Martens L, Bouckaert S, Moerman I, Demeester P. Validation of Path Loss by Heuristic Prediction Tool with Path Loss and RSSI Measurements, In *IEEE Antennas and Propagation International Symposium*, 2010; 1–4, DOI: 10.1109/APS.2010.5562014.
29. Bahl P, Padmanabhan V. RADAR: An In-building RF-based User Location and Tracking System, In *Proceedings IEEE INFOCOM*, Tel-Aviv, Israel, 2000; 775–784, DOI: 10.1109/INFCOM.2000.832252.
30. Hamza L, Nerguizian C. Neural Network and Fingerprinting-based Localization in Dynamic Channels, In *WISP 2009: Proceedings of the 6th IEEE International Symposium on Intelligent Signal Processing*, Budapest, Hungary, 2009; 253–258, DOI: 10.1109/WISP.2009.5286554.
31. Robles JJ, Deicke M, Lehnert R. 3D Fingerprint-based Localization for Wireless Sensor Networks, In *WPNC 2010: Proceeding of the 7^{th} Workshop on Positioning, Navigation and Communication*, Dresden, Germany, 2010; 77–85.
32. Gansemer S, Hakobyan S, Puschel S, Grossmann U. 3D WLAN Indoor Positioning in Multi-Storey Buildings, In *2009 IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*, Rende, Italy, 2009; 669-672, DOI: 10.1109/IDAACS.2009.5342893.
33. Neter J, Wasserman W, Kutner M. *Applied Linear Statistical Models*, Fifth Edition. MacGraw Hill: New York, 2005. ISBN-10: 0072386886.
34. Cheung KW, Sau JHM, Murch RD. A New Empirical Model for Indoor Propagation Prediction. *IEEE Transactions on Vehicular Technology* 1998; **47**(3): 996–1002.
35. Nazareth J. Conjugate gradient method. *WIREs Comp Stat* 2009; **1**: 348–353. DOI: 10.1002/WICS.13.
36. McElduff F, Cortina-Borja M, Chan SK, Wade A. When t-tests or Wilcoxon-Mann-Whitney tests won't do. *Advances in Physiology Education* 2010; **34**: 128–133. DOI: 10.1152/ADVAN00017.2010.
37. Higgens J. *Introduction to Modern Nonparametric Statistics*. Duxbury Press: Belmont, CA, 2004. ISBN-10 0534387756.

AUTHORS' BIOGRAPHIES

Frank Vanheel received his degree in electrical engineering (1979) and higher education faculty qualification (1992) from the Ghent University. He started his career in the private sector as an RF design engineer. He currently teaches courses on signals and systems at the University College Ghent, Belgium. He is also a part-time researcher, affiliated with the Information Technology Department of Ghent University. His interests include Wireless Sensor Networks, applied statistics, and hardware design.

Jo Verhaevert received an engineering degree and a PhD degree in electronic engineering from the Katholieke Universiteit Leuven, Belgium, in 1999 and 2005, respectively. He currently teaches courses on telecommunication at the Department of Applied Engineering Sciences, University College Ghent, Ghent, Belgium, where he also performs research. His research interests include indoor wireless applications (such as Wireless Sensor Networks), indoor propagation mechanisms, and smart antenna systems for wireless systems.

Eric Laermans received his Masters degree in engineering physics and a PhD degree in electrical engineering from Ghent University in 1994 and 1999, respectively. He has been an assistant professor in the Internet Based Communication Networks research group at Ghent University since 2003. His research domain has evolved from the electromagnetic modeling of high-speed interconnection structures (with special attention to via holes) and reverberation chambers to data analysis and machine learning, more specifically, surrogate modeling and experimental design.

Ingrid Moerman received her degree in electrical engineering (1987) and a PhD degree (1992) from Ghent University, where she became a part-time professor in 2000. She is a staff member of the research group on Internet Based Communication Networks and Services (www.ibcn.intec.ugent.

be), where she is leading the research on mobile and wireless communication networks. Since 2006, she joined iMinds (formerly IBBT), where she is coordinating several interdisciplinary research projects. Her main research interests include sensor networks, cooperative and cognitive networks, wireless access, self-organizing distributed networks, and experimentally-supported research. Ingrid Moerman is author or co-author of more than 400 publications in international journals or conference proceedings.

Piet Demeester received his Masters degree in electro-technical engineering and a PhD degree from Ghent University, Ghent, Belgium in 1984 and 1988, respectively. He is a full-time professor at Ghent University where he teaches courses in communication networks. He is the head of the research group on Internet Based Communication Networks and Services (http://www.ibcn.intec.ugent.be). His research interests include multilayer internet protocol optical networks, mobile networks, end-to-end quality of service, grid computing, network and service management, distributed software, and multimedia applications. He has published over 1250 papers in these areas in international journals and conference proceedings. In this research domain he was and is a member of several program committees of international conferences, such as Optical Fiber Communication, European Conference on Optical Communications, International Conference on Communications, Globecom, Infocom and Design of Reliable Communication Networks. He is a fellow of the Institute of Electrical and Electronics Engineers.