

# Situation goodness method for weighted centroid-based Wi-Fi APs localization

Germán M. Mendoza-Silva, Joaquín Torres-Sospedra, Joaquín Huerta, Raul Montoliu, Fernando Benítez, Oscar Belmonte ([gmendoza,jtorres,huerta,montoliu,benitez,m,belfern]@uji.es)

Institute of New Imaging Technologies, Universitat Jaume I, Avda. Vicente Sos Baynat S/N, 12071, Castellón, Spain

**Abstract.** Knowing the location of Wi-Fi antennas may be critical for indoor localization. However, in a real environment, their positions may be unknown since they can be managed by external entities. This paper introduces a new method for evaluating the suitability of using the weighted centroid method for the 2D localization of a Wi-Fi AP. The method is based on the idea that the weighted centroid method provides its best results when there are fingerprints taken around the AP. In order to find the probability of being in the presence of such situations, a natural neighbor interpolation method is used to find the regions with the highest signal strengths. A geometrical method is then used to characterize that probability based on the distribution of those regions in relation to the AP position estimation given by the weighted centroid method. The paper describes the testing location and the used Wi-Fi fingerprints database. That database is used to create new databases that recreate different sampling possibilities through a samples deletion strategy. The original database and the newly created ones are then used to evaluate the localization results of several AP localization methods and the new method proposed in this paper. The evaluation results have shown that the proposed method is able to provide a proper probability for the suitability of using the weighted centroid method for localizing a Wi-Fi AP.

**Keywords.** Indoor localization, Wi-Fi APs Localization, Weighted Centroid, Interpolation, LBS

## 1. Introduction

With the widespread presence of mobile devices able to consume online services and provide them the user position in some scenarios, location-based

services (LBS) have gained remarkable importance in the recent years. Applications that use LBS can shape the content they provide to the users according to the determined (or estimated) positions of their mobile devices (Werner, 2014). While GPS-based techniques provide a reasonably good solution for outdoor localization, they are not suitable for indoor environments. GPS signal strengths inside a building are too low and fluctuate too much to be reliable (Chen & Kotz, 2000). Furthermore, the use of the device's GPS sensor is sometimes avoided by the users because of its high power consumption.

Along the years, several methods have been created to provide indoor localization (Al-Ammar et al., 2014; Torres-Solis et al., 2010), based on device sensors other than the GPS one. Though there are device-centric techniques like dead-reckoning (Gutmann et al., 1998), most methods measure a physical quantity from the device surroundings. These physical quantities include sound, light, radio-frequency (RF), magnetic field and others. The methods that focus on RF, especially Wi-Fi, have been very popular. Reasons for this popularity include: (i) the knowledge already existing for the outdoor case and for wireless sensor networks, (ii) the pervasive presence of indoor wireless antennas (IEEE 802.11 and 802.15 standards, i.e., Wi-Fi and others like Bluetooth, respectively) and (iii) the widespread use of smartphones able to connect to those antennas. It is attractive to use already existing building's Wi-Fi infrastructure.

The Wi-Fi based indoor positioning techniques determine the device location based on antennas' signal intensity values (RSSI) that the device receives. Most of these techniques can be grouped according to positioning algorithm and measured property into three main positioning principles: proximity, trilateration, and scene analysis (Farid et al., 2013; H. Liu et al., 2007; H.-H. Liu & Yang, 2012). For proximity and trilateration techniques, the location of the Wi-Fi emitting antenna or access point (AP) is fundamental. Discovering the AP location is also important for management tasks, such as optimizing AP placement and detecting rogue APs.

For entities whose deployed Wi-Fi networks are unmanaged, or managed by a third party entity, it is usually necessary to estimate AP locations based on the Wi-Fi signals. The process of measuring those signals and capturing other characteristics of an AP (including its position or coverage) is often termed as war-driving (Berghel, 2004). Some studies have addressed the creation of such databases (Ledlie et al., 2011; Moreira & Meneses, 2015) and currently several global AP databases exist, such as Wigle.net<sup>1</sup>. With those

---

<sup>1</sup> <https://wile.net/>, visited on 10/06/16

databases, it is possible to obtain a gross location based only on the MAC address of the strongest AP signal received (nearest neighbor).

Several studies have developed methods for estimating an AP location based on its signal's intensities. The most known of these methods is the weighted centroid. It has been used (i) in user position localization methods (Knauth et al., 2015; Kosović & Jaguš, 2014; Lohan et al., 2015; Wang et al., 2011), (ii) as an AP (and other emitters) localization method or part of it (Blumenthal et al., 2007; Cheng et al., 2005; Y Cho et al., 2012), or (iii) as a baseline method for new AP localization methods (Han et al., 2009; Ji et al., 2013; Koo & Cha, 2011b; Zhao et al., 2014).

Other AP localization methods based on signal strength data have been developed, including the ones presented in Ji et al. (2013); Koo & Cha (2011b); Zhao et al. (2014). According to their published results, their AP localization accuracy ranges from 15 m to less than 2 m. These methods used relatively dense signal measurements, and their robustness to distribution and different numbers of signal measurements was only evaluated in a few cases. Although relatively dense mappings can be practical and they are common in laboratory settings, due to scalability issues, realistic indoor positioning systems avoid dense mappings mainly due to the large time required to generate the reference database or radio map.

This paper's main contribution is a method that, from a set of measured RSSI values of a Wi-Fi antenna, estimates the likelihood of whether the weighted centroid can provide an accurate 2D location estimation for that antenna. We have called this method as the Situation Goodness method. This paper also introduces a new AP localization method based on the natural neighbor interpolation. Furthermore, the paper provides an evaluation of AP localization methods, including the weighted centroid and the new method based on natural neighbor interpolation. The evaluation considers the localization error and considers different signal measurements situations. The Situation Goodness method is evaluated regarding its accuracy to calculate the aforementioned likelihood. This evaluation also considers different signal measurements situations.

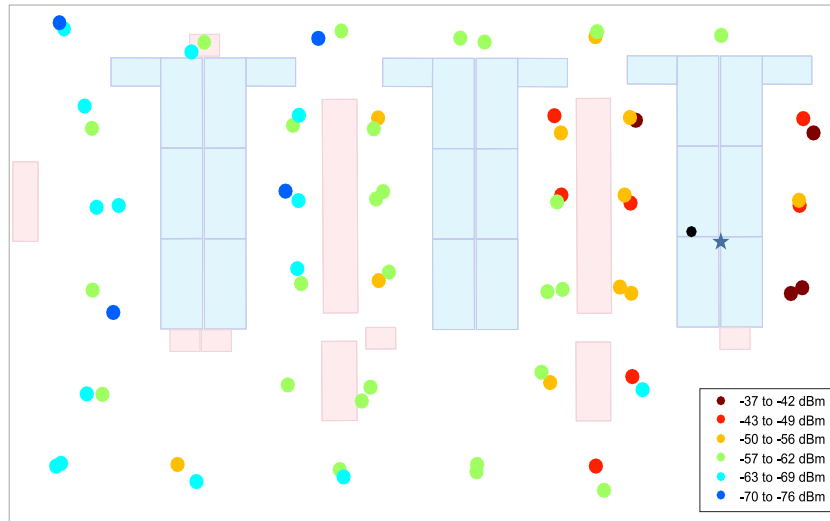
The rest of the paper is organized as follows: *Section 2* addresses a new AP localization method and others found in the literature. *Section 3* describes the proposed method for assessing the situation goodness for the weighted centroid method. *Section 4* presents experiments done to evaluate the addressed AP location methods and the new Situation Goodness method assessment for the weighted centroid. Finally, *Section 5* resumes the results obtained in this paper.

## 2. Wi-Fi AP localization methods

The problem of Wi-Fi AP localization from collected signal strength measurements (RSSI) have been addressed in the literature for more than 10 years, and several methods have been proposed. The following sections describe some of these methods applied to an AP 2D localization. *Section 2.1* describes the weighted centroid method, a building block of this paper's main contribution. *Section 2.2* introduces a new AP localization method based on natural neighbor interpolation. *Section 2.3* addresses another three AP localization methods we have considered to be relevant. All these methods are the ones used in the evaluation that *Section 4.2* presents.

### 2.1. Weighted centroid method

The weighted centroid method has been used for Wi-Fi APs in many research works: in user position localization methods, in AP localization methods and as a baseline method for new AP localization methods (Blumenthal et al., 2007; Cheng et al., 2005; Y Cho et al., 2012; Han et al., 2009; Ji et al., 2013; Knauth et al., 2015; Koo & Cha, 2011b; Kosović & Jaguš, 2014; Lohan et al., 2015; Wang et al., 2011; Zhao et al., 2014). The reasons behind its popularity include its simplicity, its low computational complexity, and its relatively low AP localization error in some known situations.



**Figure 1.** Weighted centroid method's estimation when there are measurement taken around the AP. The colored circles represent the input measurement positions and strengths. The black point and the blue star represent the AP location estimation and its actual location, respectively. The pale blue and pink rectangles are the furniture found in the sampled area.

Given a ground truth of  $n$  signal strengths ( $R_i$ ) of an AP, measured at known positions  $\mathbf{P}_F=(x_i, y_i)$ , the weighted centroid method estimates the AP's location using formula (1).

$$\mathbf{P}_a = \frac{\sum_{i=1}^n \mathbf{P}_i W_i}{\sum_{i=1}^n W_i}, \quad (1)$$

where  $W_i$  represent the calculated weight of the  $i$ th measurement and  $\mathbf{P}_a=(x_a,y_a)$  is the estimated AP's location. The weights represent the importance given to the position of each measurement and are, ideally, inversely proportional to the distance between the measurement point and the AP. As the real distance is usually unknown, it can be estimated using a power-based propagation model like, e.g., the one presented in Bahl & Padmanabhan (2000). However, the weights are commonly calculated directly from the RSSI values, i.e., without estimating a distance. The calculation is based on the idea that the shorter the distance, the more intense the RSSI. Therefore, the more intense the RSSI, the greater must be its associated weight. The weight calculation used in this paper is presented in Lohan et al. (2015) and is shown in (2).

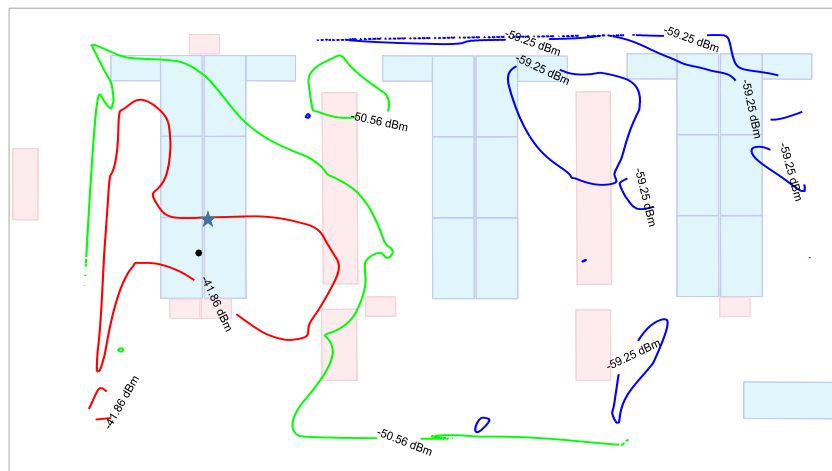
$$W_i = 10^{\frac{R_i}{10}} \quad (2)$$

The way this method calculates the AP position determines the situations in which the method provides good location estimations: If (i) there are measurements taken around the real AP position or (ii) the measurements are close to the AP. *Figure 1* shows an example of case (i). This figure present measured signal intensities and the location estimation of the AP. The data is shown over a furniture layout of the sampled area composed by shelves and cabinets (pink) and tables (blue). In this example, the estimated location is closest to the locations of the measurements with the strongest RSSI values, i.e., the measurements with RSSI values stronger than  $56 \text{ dBm}$ . In the case (ii) the measurements closest to the AP are more likely to be the strongest and to have the highest weights. However, if the measurements were taken only towards one side of the AP, and none of them is close to the AP, the location estimation is poor. Despite those facts, many studies have used the weighted centroid as a baseline comparison method while considering situations where this method does not produce good localization results.

## 2.2. Interpolation contours centroid method

Interpolation and extrapolation methods have already been used in RSSI-based indoor localization, mainly in the Wi-Fi radio map's database enrichment (Arai & Tolle, 2013; Ezpeleta et al., 2015; Lee & Han, 2012). Although some research works have also used extrapolation methods in indoor localization-related contexts, like in Talvitie et al. (2015), the interpolation methods are far more accurate than the extrapolation ones.

Interpolation methods can be used to spot regions of the sampled area where the signal intensities are the highest, and use these regions to estimate the AP's location. This approach shares a drawback with the weighted centroid: It is able to locate the AP only inside the sampled area. The drawback results from the fact that interpolation methods provide function values estimations for within the convex hull of the original set of points. However, as the selected regions are the ones closest to AP's location, they can be used to get AP localization results similar to those obtained by the weighted centroid in the cases where the weighted centroid performs the best; and even better results in the other cases.



**Figure 2.** AP location estimation using the interpolation contours centroid method. The AP location estimation (red contours' centroid) is shown by the black point and the real AP location by a blue star.

In particular, the implementation we have used employs the natural neighbor interpolation (Sibson & others, 1981). The steps we propose are the following:

- 1- Calculate an interpolated grid of points using natural neighbor interpolation.
- 2- Determine the goal intensity level. To find the goal level, three intensity levels (“high”, “medium” and “low”) are determined so that they split the input RSSI values into 4 equally spaced intervals. The “high” level is the goal intensity level.
- 3- Compute the contours corresponding to the goal (highest) intensity level. Contours are calculated using the Marching Squares algorithm (Maple, 2003)<sup>2</sup>.

---

<sup>2</sup> As implemented in Matlab.

- 4- Obtain the points that represent the contours computed in step 3. If needed, assure two requirements: (i) the point densities along the edges are similar and (ii) the edges' curvature are well described by their points.
- 5- Using the points from step 4, calculate a centroid by averaging the points coordinates. The position of this centroid is the AP location estimation.

*Figure 2* shows the result of estimating an AP's position using this method on a set of RSSI values. The blue, green and red lines depict the contours corresponding to the "low", "medium" and "high" intensity levels, respectively. The number over a contour shows the corresponding intensity level value. Notice that only the red contours, which are the ones that bound the highest signal strengths regions, are used to estimate the AP location. This method is more computationally complex than the weighted centroid, and its complexity depends on the number of interpolated grid's points. Also, the grid interpolation density influences in the actions needed in step 4. Step 3 already provide the contours as a discretized set of points. If the grid is interpolated with high density, e.g., 0.05 m of separation between points (which is the one used in the experiments presented in *Section 4*), then the interpolation is smooth, and the contours' points satisfy the requirements of step 4. However, if the grid's interpolation density is low, a further interpolation procedure on the contours' points is needed to meet the above requirements.

### **2.3. Other methods from literature**

In the literature, many research works have addressed the AP localization (Blumenthal et al., 2007; Cheng et al., 2005; Y Cho et al., 2012; Youngsu Cho et al., 2012; Han et al., 2009; Ji et al., 2013; Koo & Cha, 2011a, 2011b, 2012; Nam, 2014; Varzandian et al., 2013; Zhang et al., 2011; Zhao et al., 2014). In this section, we describe three methods we have considered to be relevant.

The method presented in Koo & Cha (2011b) linearly approximates the exponential relation between distance and signal strength that establishes the well-known distance-power law. They introduce this linear approximation in an equation system in which each equation corresponds to a measurement point. Then, they perform the proper transformations - based on Savvides et al. (2001) - and obtain a system of linear equations that is only dependent on the measurements' positions and intensities, and on one linearization coefficient that does not affect the final AP position estimation. The authors tested their method through simulations and the errors estimation they presented were higher than 5 m.

In Ji et al. (2013), a simulation-like (Monte Carlo test) approach is used to test a set of possible locations of an AP to find the one that better fits the measured RSSI values for the AP. For each possible AP position, and based

on the measured RSSI values, the path-loss model parameters are estimated by solving a system of linear equations. Those parameters are in turn used to obtain an estimated RSSI value at each measurement point. The difference between the expected RSSI and the measured RSSI values is then used as a metric to evaluate the goodness of the tested location for the AP. With simulations, their method achieved an accuracy below 10 m in 95% of cases. With the real samples, the method achieved an accuracy below 12 m in 90% of cases.

The authors in Zhao et al. (2014) proposed a method for AP localization based on RSSI gradients calculation. The gradient (intensity and direction) for each point on a rectangular, uniformly distributed, grid is calculated using 1D centered, point discrete derivative masks, i.e., taken into account the signal intensity variation in the  $y$ -axis - between the two (up and down) closest neighbors - and in the  $x$ -axis - between the two (right and left) closest neighbors. Then, a  $k$ -means clustering is applied to (i) identify direction outliers and (ii) find the cluster with the highest number of sampling points for using its head's position as the AP estimated position. The authors compared their method with the weighted centroid algorithm and the gradient approach without clustering, using experimentally collected data. Their method outperformed the other two methods and had a mean localization error of 1.5 m on the experimentation setup.

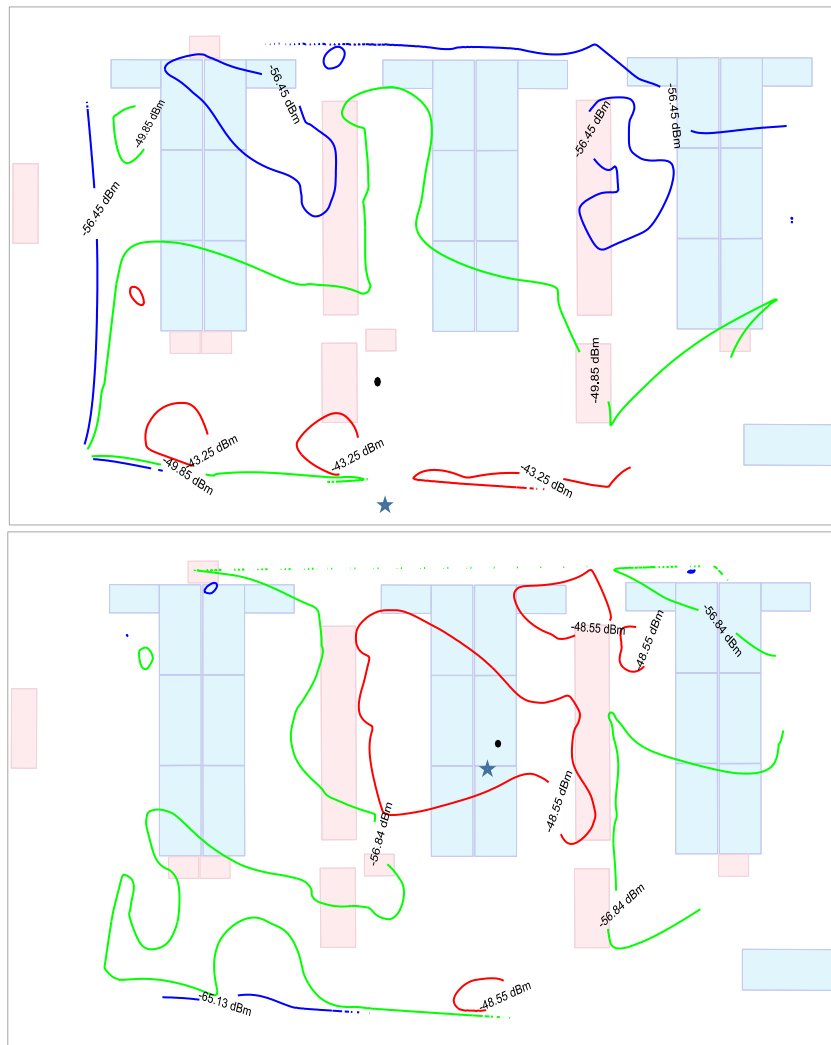
### **3. Situation Goodness method**

This section presents a method that, given a set of Wi-Fi intensity measurements for an AP, provides a probability that describes whether the weighted centroid method is likely to provide a solution that is close to the actual AP location. The weighted centroid method provides its best solution when some of the intensity measurements are around the AP, and to a lesser extent, when they are close to the AP.

The idea behind the method combines part of the interpolation method presented in *Section 2.2* and the weighted centroid method. Specifically, the new method explores how the highest intensity level contours explained in *Section 2.2* are distributed regarding the AP location estimation provided by the weighted centroid.



Figure 3 presents the idea in two typical situations. The top image presents the case where the measurements do not surround the AP (which is located at the bottom, outside the measurements convex hull). For the top image case, the highest level regions are not balanced around the weighted centroid estimation. In the bottom image, on the other hand, the highest level regions are distributed around the weighted centroid estimation. For this second case, the weighted centroid estimation is very close to the actual AP location.



**Figure 3.** Two cases of the distribution of highest intensity level contours (drawn in red) regarding the weighted centroid AP location estimation (the black dot). The real AP location is shown by a blue star.

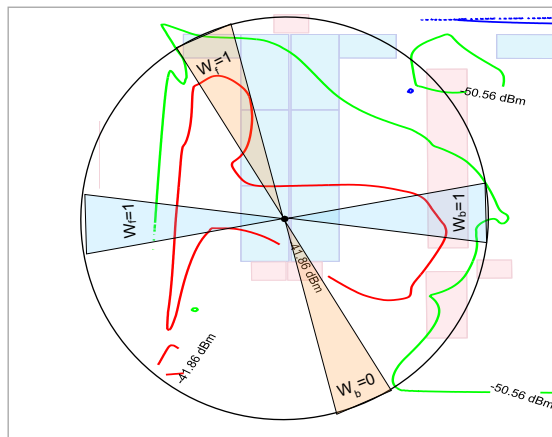
The new method explores the geometrical relation between the weighted centroid solution and the aforementioned highest level regions. That relation is

then used to calculate the probability for the weighted centroid method to provide a solution that is close to the actual AP location. The method performs the following steps:

- 1- Calculate the weighted centroid (WC) AP location estimation  $p_{wc} = (x_{wc}, y_{wc})$ .
- 2- Calculate the highest level regions' contours, as set of points  $C = \{C_1, \dots, C_n\}$ , where  $C_i = \{c_{i1}=(x_{i1}, y_{i1}), \dots, c_{im}=(x_{im}, y_{im})\}$ .
- 3- Calculate the directions (angles) of all contour points in  $C$ , taking  $p_{wc}$  as origin, as well as the distances to  $p_{wc}$ . The direction of the contour point  $c_{ij} = (x_{ij}, y_{ij})$ , which is the  $j$ th point of  $i$ th contour, is calculated as  $v_{ij} = \tan^{-1}((y_{wc} - y_{ij})/(x_{wc} - x_{ij}))$ . The distance associated to  $c_{ij}$  is calculated as the Euclidean distance between  $c_{ij}$  and  $p_{wc}$ .
- 4- Group the contours' points according to their directions in windows with a small size ( $\pi/60$ ).
- 5- For each pair of opposing windows ( $W_f, W_b$ ), e.g.,  $W_f = [0, \pi/60]$  and  $W_b = [\pi, \pi + \pi/60]$ , do:
  - a. Calculate  $s_{wf}$  and  $s_{wb}$  as the sum of all distances (calculated in step 3) of points in  $W_f$  and  $W_b$ , respectively.
  - b. Determine the dominance  $d_{wf}$  and  $d_{wb}$  of windows  $W_f$  and  $W_b$ , respectively, as follows:
    - i. When  $(s_{wf}/s_{wb}) > 2$ ,  $d_{wf} = 1$  and  $d_{wb} = 0$ .
    - ii. When  $(s_{wf}/s_{wb}) < 0.5$ ,  $d_{wf} = 0$  and  $d_{wb} = 1$ .
    - iii. When  $0.5 \leq (s_{wf}/s_{wb}) \leq 2$ ,  $d_{wf} = 1$  and  $d_{wb} = 1$ .
- 6- Fill the dominance gaps. When two windows whose dominance is 1 are separated by no more than 3 (grace gap) windows whose dominance is 0 (gap windows), set the dominance of the separating windows to 1.
- 7- Set  $u$  as the number of opposing windows whose dominance is different and  $t$  as the total number of windows whose dominance is 1.
- 8- Provide the output probability as:  $prob = 1 - u/t$ .

The calculation methods for steps 1 and 2 were already explained in *Section 2.1* and *Section 2.2*, respectively. Step 2 takes into account how the contours are described as a set of points in the same way that the *Section 2.2* presents it. The directions and distances obtained in the step 3 help in finding whether the highest level contours are located in a balanced or unbalanced way regarding the weighted centroid estimation. To ease the required comparisons, the contour points are grouped into windows in step 4. The window's size presented in that step, ( $\pi/60$ ), was chosen because it is relatively small and round, but smaller values were also tested and provided equally good results.

The step 5 associates values called dominances to each window. Each of these values indicates, for a particular pair of opposing windows (opposing sets of directions), towards where the highest level areas are located. To determine the winning window(s), the sum of distances from the window's points to the centroid position is calculated. If the sum of a windows is twice or more the opposing window's sum, the first window is the dominant one. *Figure 4* shows the dominances of two sample pairs of opposing windows for a given highest level contour. Notice how the number of points found inside the blue triangles and the distances between these and the weighted centroid's AP location estimation are similar. Thus, the windows corresponded to these triangles are dominant. In the case of the windows represented by the orange triangles, it can be noticed that only one is dominant.



**Figure 4.** The concepts of opposing windows and dominance. The orange triangles represent two opposing windows, as well as the blue ones. The blue windows are equally dominant, but only one of the orange ones is.

In step 6, as the contours can be very irregular, the gaps (windows) between dominant windows that are close to each other are filled (made dominant). The value determining the gap maximum size is called grace gap. This value is directly related to the window size. In our tests, we found that a grace gap value of 3 worked well for a window size of  $(\pi/60)$ .

Step 7 uses the dominance values to find unbalanced situations, i.e., where one window is dominant and the opposing one is not. This unbalanced situation count is finally used in step 8 to calculate the intended probability. *Section 4.3* presents the evaluation of the above method.

## 4. Experiments

This section presents an evaluation of the AP localization methods described in *Section 2*, as well as an evaluation of the method presented in *Section 3*. Before the evaluations, the Wi-Fi measurements datasets used in the evaluations are presented.

### 4.1. Test locations and Wi-Fi samples databases

The GeotecLab Wi-Fi database<sup>3</sup> was used to perform the experiments. That database and its related test location are already described in Torres-Sospedra et al. (2016). In short, the samples in the database were taken in the headquarters' lab of the Geospatial Technologies (GEOTEC) research group, in two separate moments, by two persons using two different mobile phones, respectively. Although plenty of details can be found in Torres-Sospedra et al. (2016), it is worth mentioning that the lab is a typical office working environment - wooden and metallic furniture, working computers, typical office equipment, people working and moving, and was not intentionally prepared in any way for Wi-Fi related experiments. Also, the lab occupies 260 m<sup>2</sup> and thus can be considered as a medium-sized facility.

The GeotecLab Wi-Fi database is divided into two datasets, one containing samples that are meant to be used for training, and the other for validation purposes. This configuration is typically used in the machine learning techniques applied in the most popular method for Wi-Fi based indoor localization: Wi-Fi fingerprinting. Both datasets contain data regarding the intensities of the APs detected by the used phones.

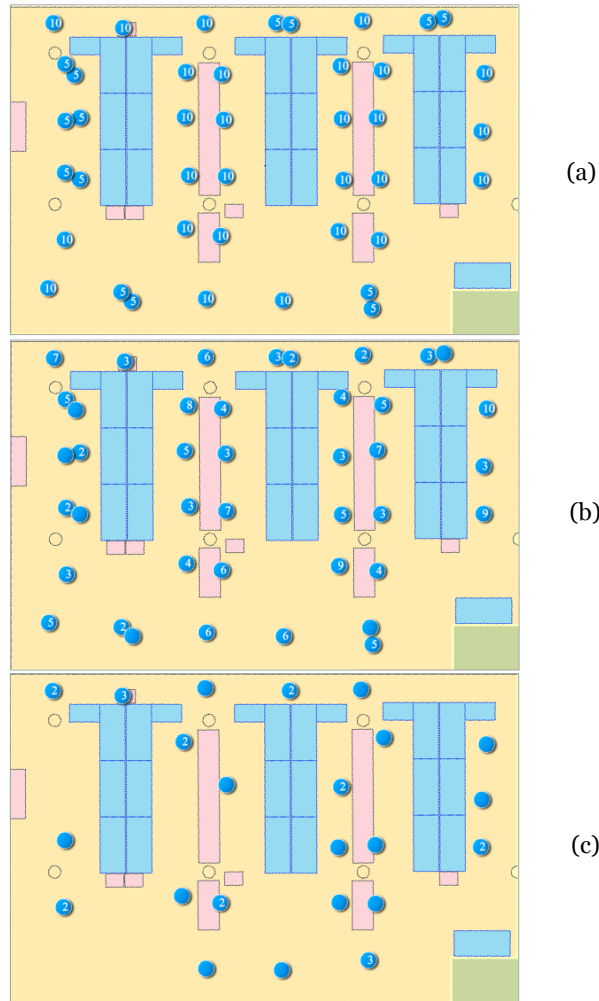
The two mentioned datasets were used to create the one dataset used in work described in our paper. The new dataset was created by selecting samples using the following three criteria:

- 1- Combine the training and validation datasets. As this work does not use machine learning approaches, the datasets were combined in order to have a large number of samples to work with.
- 2- Select the samples corresponding to APs whose positions are known. The selected antennas are the ones located inside the GEOTEC's lab and knowing their locations was mandatory for the evaluation of AP localization methods.

---

<sup>3</sup> Available at <http://indoorloc.uji.es>, visited on 10/06/16

- 3- Use only the samples taken using the Samsung S3 (Android 4.3) phone. The Samsung S3 phone is able to detect 2.4 GHz and 5.0 GHz networks, while the LG Spirit (Android 5.0.1) is only able to detect 2.4 GHz networks. By using only these samples, it was possible to test the AP localization methods for 2.4 GHz and 5.0 GHz emitting antennas and also to have a similar number of samples for each AP.



**Figure 5.** The original database samples (a), and the results of eliminating 50% (b) and 90% (c) of measurements. The numbered circles represent the amount of samples that are located at that position (or very close to it).

The new dataset, from now on called the root dataset and shown in *Figure 5* (a), was in turn used to create new datasets by applying a sample elimination

strategy in order to recreate different sampling alternatives. The used elimination strategy is a uniform random elimination. This strategy led to obtaining 10 datasets, which contain 10, 20, 30, ..., 100 percent of the root dataset samples, respectively. *Figure 5* (b) and (c) present sample datasets with 50% and 10% of the root dataset samples, respectively.

#### 4.2. AP localization methods evaluation

The methods evaluated in this section were already described in previous sections:

- Weighted centroid, as described in *Section 2.1*,
- Interpolation contours centroid, as explained in *Section 2.2*,
- Linearization, Monte Carlo, and Gradients as shown in *Section 2.3*.

<i>Method</i> <i>Dataset</i>	<i>Mean Error (m)</i>				
	<i>Centroid</i>	<i>Monte Carlo</i>	<i>Linear</i>	<i>Interpolation</i>	<i>Gradient</i>
<i>D10</i>	3.8	8.2	7.5	3.5	11.9
<i>D20</i>	4.7	8.1	7.3	4.0	13.0
<i>D30</i>	4.4	8.1	7.3	4.3	13.6
<i>D40</i>	4.2	8.0	7.2	2.9	14.0
<i>D50</i>	3.7	8.0	7.2	3.6	14.3
<i>D60</i>	3.8	8.0	7.1	3.9	14.4
<i>D70</i>	4.0	7.9	7.1	3.4	14.6
<i>D80</i>	4.3	7.8	7.1	3.3	14.4
<i>D90</i>	4.1	7.8	7.1	3.4	14.4
<i>D100</i>	4.0	7.7	7.1	2.8	14.6
<i>Overall Mean Error</i>	4.1	8.0	7.2	3.5	13.9
<i>Overall Variance</i>	0.1	0.0	0.0	0.2	0.7

**Table 1.** Mean estimation error of the selected AP location methods. Numbers are rounded to decimeter level.

Each localization experiment consisted of applying one of the five AP localization methods to one of the datasets created according to the elimination strategy described in *Section 4.1*. *Table 1* describes the results obtained for each of the above experiments. As that strategy involves random elimination, the evaluation for a method proceeded as follows:

- 1- Repeat 50 times:
  - a. Create the datasets  $D_{10}$ ,  $D_{20}$ , ...,  $D_{100}$  according to the elimination strategy.
  - b. Run the chosen method against each dataset and store the localization error results (distance from the estimated location to the real one for each antenna) in  $E_{10}$ ,  $E_{20}$ , ...,  $E_{100}$ , respectively.
  - c. Accumulate the results from the previous step into  $Ac_{10}$ ,  $Ac_{20}$ , ...,  $Ac_{100}$ , respectively.
- 2- Calculate the mean values  $M_{10}$ ,  $M_{20}$ , ...,  $M_{100}$  from values accumulated in the previous step.

<i>Method</i> <i>Dataset</i>	<i>Mean Error (m)</i>	
	<i>Centroid</i> <i>(surrounded)</i>	<i>Centroid</i> <i>(non-surrounded)</i>
<i>D10</i>	2.2	5.2
<i>D20</i>	2.8	6.3
<i>D30</i>	2.0	6.4
<i>D40</i>	2.6	5.6
<i>D50</i>	1.3	5.7
<i>D60</i>	2.1	5.2
<i>D70</i>	2.3	5.3
<i>D80</i>	2.0	6.0
<i>D90</i>	1.9	5.9
<i>D100</i>	1.9	5.6
<i>Overall Mean Error</i>	2.1	5.7

**Table 2.** Differences in mean error localization when differencing between cases where there are samples taken around the AP, and those that do where no samples are taken around the AP.

The figures presented in *Table 1* are the ones obtained by applying the above steps. Despite the evaluation process considered to repeat 50 times the same experiment to take into account random value selections - like in the elimination strategy or the Monte Carlo method, the mean error presented in *Table 1* does not steadily increase as the number of measurement points decreases. Although experiments with a higher number of repetitions can be

performed, the above fact means that none of these methods is heavily affected by the sampling density, as long as the point's distribution remains practically the same. However, considering the overall mean error, the best two are the weighted centroid and interpolation methods. If we consider the overall variance for these two methods, the weighted centroid is the best one.

*Table 2* deepens into the fact that the weighted centroid method provide better localization results when fingerprints have been collected around the target AP. The column “Centroid (surrounded)” presents the mean localization error for such cases, i.e., considering only antennas for which there are fingerprints taken around them. The overall localization error, which consider all datasets, for the “surrounded” cases is 3.7 m better than in the “non-surrounded” cases.

### 4.3. Situation Goodness method evaluation

In this section, we describe the results of evaluating the Situation Goodness method. For the evaluation, we used similar approaches to the ones used in *Section 4.2*. In particular, for the evaluation that used the datasets created by applying the elimination strategy, the following steps were used:

- 1- Repeat 50 times:
  - a. Create the datasets D10, D20, ..., D100 according to the elimination strategy.
  - b. Run the method against each dataset and store the probability results in P10, P20, ..., P100, respectively.
  - c. Accumulate the results from the previous step into Ac10, Ac20, ..., Ac100, respectively.
- 2- Calculate the mean values M10, M20, ..., M100, and variance values V10, V20, ..., V100 from values accumulated in the previous step.

The evaluation results are presented in *Table 3* and *Table 4*. Hereinafter, we call “surrounded cases” the cases in which the weighted centroid method provides a good AP location estimation (as a result of the existence of measurements taken around the AP). We have called the other cases as “non-surrounded cases”. In *Table 3* and *Table 4*, the column “WC good solution” indicates the surrounded cases with a “yes” value and non-surrounded cases with a “no” value.

The figures in *Table 3* show how the Situation Goodness method provides low probability values for non-surrounded cases. For these cases, the probability values are always low (below or equal to 0.3). As can be expected, the most confident results correspond to the datasets with the highest number of measurements.

For the surrounded cases, the probability values are above 0.3 for datasets from D20 to D100, i.e., if the dataset is not much degraded, the method does



identify the surrounded cases taken into account the mean value. The most confident results correspond to the datasets with the highest number of measurements.

Dataset Antenna	Probability Mean										WC good solution
	D10	D20	D30	D40	D50	D60	D70	D80	D90	D100	
67	0.19	0.29	0.27	0.28	0.30	0.24	0.22	0.28	0.26	0.13	no
68	0.18	0.36	0.42	0.43	0.42	0.44	0.50	0.53	0.52	0.52	yes
70	0.14	0.22	0.15	0.19	0.09	0.13	0.13	0.09	0.08	0.00	no
74	0.14	0.17	0.21	0.21	0.15	0.20	0.19	0.11	0.19	0.00	no
80	0.15	0.20	0.17	0.17	0.12	0.13	0.14	0.10	0.10	0.00	no
84	0.16	0.17	0.18	0.15	0.11	0.11	0.14	0.10	0.12	0.00	no
95	0.42	0.54	0.59	0.53	0.62	0.63	0.70	0.71	0.66	0.85	yes
96	0.17	0.31	0.36	0.40	0.43	0.48	0.42	0.46	0.51	0.66	yes
97	0.46	0.65	0.71	0.71	0.79	0.82	0.82	0.86	0.83	0.91	yes

**Table 3.** Mean values of the probability provided by the Situation Goodness method.

The mean probability values of antennas 70, 74, 80 and 84 are similar. This is a likely result of the fact that those are actually one antenna that transmits at 2.4 GHz and 5.0 GHz, providing 2 networks at each frequency. Considering the surrounded cases, the antenna 68 is the one with lower probability values. Antennas 95, 96 and 97 are located on top of desks, while antenna 68 is located in the ceiling. The antenna 67 is located inside of a room adjacent to the sampled area.

*Table 4* presents the variability of the probability values across the 50 experiment repetitions. As expected, there is no variation for the dataset D100, due to no samples elimination was performed on it. For datasets D50 to D90, the variance value is no higher than 0.07, which, paired with the mean probability values for those datasets (presented in *Table 3*), indicates that the probability value will likely be above 0.3 for the surrounded cases, and below that value for the non-surrounded cases.

Dataset Antenna	Probability Variance										WC good solution
	D10	D20	D30	D40	D50	D60	D70	D80	D90	D100	
67	0.08	0.08	0.08	0.06	0.07	0.06	0.04	0.05	0.04	0.00	no
68	0.05	0.08	0.07	0.06	0.07	0.05	0.02	0.02	0.02	0.00	yes
70	0.04	0.07	0.05	0.06	0.01	0.02	0.02	0.01	0.01	0.00	no
74	0.06	0.07	0.08	0.10	0.04	0.07	0.07	0.04	0.06	0.00	no
80	0.05	0.06	0.06	0.06	0.02	0.02	0.03	0.01	0.01	0.00	no
84	0.06	0.06	0.07	0.06	0.03	0.03	0.06	0.03	0.04	0.00	no
95	0.11	0.11	0.08	0.08	0.07	0.05	0.04	0.04	0.06	0.00	yes
96	0.05	0.05	0.04	0.05	0.05	0.04	0.04	0.02	0.03	0.00	yes
97	0.12	0.08	0.06	0.07	0.02	0.02	0.01	0.01	0.02	0.00	yes

**Table 4.** Variance values of the probability provided by the Situation Goodness method.

Dataset Antenna	Mean Error (m) for D100		WC good solution
	Linear	W.C.-Linear	
67	15.0	15.0	no
68	7.8	5.2	yes
70	5.7	5.7	no
74	5.8	5.8	no
80	5.6	5.6	no
84	5.8	5.8	no
95	3.5	0.6	yes
96	10.2	0.8	yes
97	4.1	1.0	yes

**Table 5.** Localization mean errors on dataset D100 when the Situation Goodness method is applied to combine the Linear and Weighed Centroid methods.

Based on the previous analyses, the Situation Goodness method can be used to combine the weighted centroid method with another AP localization method. We tested the following way of combination: Given a set of RSSI

measurements for an AP, if the probability given by the Situation Goodness method is lower than 0.31, use the main method (in our case, one of the tested ones); otherwise, use the weighted centroid method. *Table 5* presents, as an example, the localization errors when the above methods combination is applied to dataset D100. The weighted centroid method was combined with the linear approximation method. The probability provided by the Situation Goodness method was correctly used to identify the surrounded cases, and for the AP associated to those cases, the localization solution provided by the weighted centroid was used instead of the one provided by the linear approximation method.

		<i>Mean Error (m)</i>				
<i>Dataset</i>	<i>Case</i>	<i>W. C.- W. C.</i>	<i>W. C.-Monte Carlo</i>	<i>W. C.- Linear</i>	<i>W. C.- Interpolation</i>	<i>W. C.- Gradient</i>
<i>D10</i>		4.7	6.5	6.2	4.5	8.3
<i>D20</i>		4.4	5.9	5.8	3.9	9.2
<i>D30</i>		4.2	5.2	5.6	3.8	9.1
<i>D40</i>		4.2	5.3	5.4	3.6	9.9
<i>D50</i>		4.0	5.0	5.5	3.5	9.9
<i>D60</i>		4.0	4.8	5.2	3.5	10.2
<i>D70</i>		4.1	4.8	5.1	3.4	9.8
<i>D80</i>		4.1	4.7	5.4	3.3	10.3
<i>D90</i>		4.1	4.8	5.2	3.2	10.4
<i>D100</i>		4.0	4.3	5.1	2.8	11.1
<i>Overall Mean Error</i>		4.2	5.1	5.4	3.5	9.8
<i>Overall Variance</i>		0.0	0.4	0.1	0.2	0.6

**Table 6.** Mean localization error for combinations of weighted centroid and other methods.

*Table 6* presents the localization errors when the AP localization evaluation is performed over the methods combination as explained above. When compared with the data presented by *Table 1*, *Table 6* shows how the mean localization error is mostly similar for the interpolation method, and is greatly reduced for the Monte Carlo, Linear and Gradient methods (which have poor performances on our datasets).

According to the results of the experimentation presented in this paper, if the probability calculated with our Situation Goodness method is reasonably

high, the weighted centroid's estimation can be considered as a proper AP location solution for other RSSI measurement datasets. Thus, we recommend to use the Situation Goodness method to combine the weighted centroid method with another AP localization method for determining several AP s' locations in other datasets. Despite that the probability threshold value we used in our experimentation is 0.31, we recommend to use at least a value of 0.5, because further experimentation in other datasets is required to provide a more accurate threshold value.

## 5. Conclusion

This paper has presented a new method that, given a set of Wi-Fi intensity measurements for an AP, provides a probability that describes whether the weighted centroid method is likely to provide a solution that is close to the actual AP location. We have called this method as the Situation Goodness method. To spur the interest in this method, we have provided facts from literature and an AP localization methods evaluation to show the goodness of using the weighted centroid method. In the AP localization evaluation, we included a new method that we have developed and which is based on natural neighbor interpolation. Through evaluation, we have shown that the Situation Goodness method we propose can effectively achieve its intended goal in a range of Wi-Fi measurement collection alternatives.

Using the Situation Goodness method, it is possible to choose when to use the weighted centroid method for an actual AP localization or as a baseline comparison method. For those cases where the Situation Goodness method indicates that the weighted centroid method is not the best choice, a different AP localization method should be used. Among the alternatives, the AP localization method we have proposed can be used.

## Acknowledgements

The authors gratefully acknowledge funding from the European Union through the GEO-C project (H2020-MSCA-ITN- 2014, Grant Agreement Number 642332, <http://www.geo-c.eu/>).

## References

- Al-Ammar, M. A., Alhadhrami, S., Al-Salman, A., Alarifi, A., Al-Khalifa, H. S., Alnafessah, A., & Alsaleh, M. (2014). Comparative Survey of Indoor Positioning Technologies, Techniques, and Algorithms. In *Cyberworlds (CW), 2014 International Conference on* (pp. 245–252). <http://doi.org/10.1109/CW.2014.41>

- Arai, K., & Tolle, H. (2013). Color radiomap interpolation for efficient fingerprint wifi-based indoor location estimation. *International Journal of Advanced Research in Artificial Intelligence*, 2(3), 10–15.
- Bahl, P., & Padmanabhan, V. N. (2000). RADAR: An in-building RF-based user location and tracking system. In *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE (Vol. 2, pp. 775–784)*.
- Berghel, H. (2004). Wireless infidelity I: War driving. *Communications of the ACM*, 47(9), 21–26.
- Blumenthal, J., Grossmann, R., Golatowski, F., & Timmermann, D. (2007). Weighted Centroid Localization in Zigbee-based Sensor Networks. In *Intelligent Signal Processing, 2007. WISP 2007. IEEE International Symposium on (pp. 1–6)*. <http://doi.org/10.1109/WISP.2007.4447528>
- Chen, G., & Kotz, D. (2000). *A Survey of Context-Aware Mobile Computing Research*. Hanover, NH, USA: Dartmouth College.
- Cheng, Y.-C., Chawathe, Y., LaMarca, A., & Krumm, J. (2005). Accuracy Characterization for Metropolitan-scale Wi-Fi Localization. In *Proceedings of the 3rd International Conference on Mobile Systems, Applications, and Services (pp. 233–245)*. New York, NY, USA: ACM. <http://doi.org/10.1145/1067170.1067195>
- Cho, Y., Ji, M., Lee, Y., Kim, J., & Park, S. (2012). Improved Wi-Fi AP position estimation using regression based approach. In *Proc. of the International Conference on Indoor Positioning and Indoor Navigation*.
- Cho, Y., Ji, M., Lee, Y., & Park, S. (2012). WiFi AP position estimation using contribution from heterogeneous mobile devices. In *Position Location and Navigation Symposium (PLANS), 2012 IEEE/ION (pp. 562–567)*. <http://doi.org/10.1109/PLANS.2012.6236928>
- Ezpeleta, S., Claver, J., Pérez-Solano, J., & Martí, J. (2015). RF-Based Location Using Interpolation Functions to Reduce Fingerprint Mapping. *Sensors*, 15(10), 27322–27340. <http://doi.org/10.3390/s151027322>
- Farid, Z., Nordin, R., & Ismail, M. (2013). Recent advances in wireless indoor localization techniques and system. *Journal of Computer Networks and Communications*, 2013.
- Gutmann, J.-S., Burgard, W., Fox, D., & Konolige, K. (1998). An experimental comparison of localization methods. In *Intelligent Robots and Systems, 1998. Proceedings., 1998 IEEE/RSJ International Conference on (Vol. 2, pp. 736–743 vol.2)*. <http://doi.org/10.1109/IROS.1998.727280>
- Han, D., Andersen, D. G., Kaminsky, M., Papagiannaki, K., & Seshan, S. (2009). Access Point Localization Using Local Signal Strength Gradient. In S. B. Moon, R. Teixeira, & S. Uhlig (Eds.), *Passive and Active Network Measurement: 10th International Conference, PAM 2009, Seoul, Korea, April 1-3, 2009. Proceedings (pp. 99–108)*. Berlin, Heidelberg: Springer Berlin Heidelberg. [http://doi.org/10.1007/978-3-642-00975-4\\_10](http://doi.org/10.1007/978-3-642-00975-4_10)
- Ji, M., Kim, J., Cho, Y., Lee, Y., & Park, S. (2013). A novel Wi-Fi AP localization method using Monte Carlo path-loss model fitting simulation. In *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC) (pp. 3487–3491)*. <http://doi.org/10.1109/PIMRC.2013.6666752>
- Knauth, S., Storz, M., Dastageeri, H., Koukofikis, A., & Mähser-Hipp, N. A. (2015). Fingerprint calibrated centroid and scalar product correlation RSSI positioning in large environments. In *Indoor Positioning and Indoor Navigation (IPIN), 2015 International Conference on (pp. 1–6)*. <http://doi.org/10.1109/IPIN.2015.7346968>

- Koo, J., & Cha, H. (2011a). Autonomous construction of a WiFi access point map using multi-dimensional scaling. In *Pervasive Computing* (pp. 115–132). Springer.
- Koo, J., & Cha, H. (2011b). Localizing WiFi access points using signal strength. *IEEE Communications Letters*, 15(2), 187–189. <http://doi.org/10.1109/LCOMM.2011.121410.101379>
- Koo, J., & Cha, H. (2012). Unsupervised locating of WiFi access points using smartphones. *Systems, Man, and Cybernetics, Part C: Applications and Reviews*, *IEEE Transactions on*, 42(6), 1341–1353.
- Kosović, I. N., & Jaguš, T. (2014). Enhanced Weighted Centroid Localization Algorithm for Indoor Environments. *World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering*, 8(7), 1184–1188.
- Ledlie, J., g. Park, J., Curtis, D., Cavalcante, A., Camara, L., Costa, A., & Vieira, R. (2011). Mole: A scalable, user-generated WiFi positioning engine. In *Indoor Positioning and Indoor Navigation (IPIN), 2011 International Conference on* (pp. 1–10). <http://doi.org/10.1109/IPIN.2011.6071942>
- Lee, M., & Han, D. (2012). Voronoi Tessellation Based Interpolation Method for Wi-Fi Radio Map Construction. *IEEE Communications Letters*, 16(3), 404–407. <http://doi.org/10.1109/LCOMM.2012.020212.111992>
- Liu, H., Darabi, H., Banerjee, P., & Liu, J. (2007). Survey of Wireless Indoor Positioning Techniques and Systems. *Systems, Man, and Cybernetics, Part C: Applications and Reviews*, *IEEE Transactions on*, 37(6), 1067–1080. <http://doi.org/10.1109/TSMCC.2007.905750>
- Liu, H.-H., & Yang, Y.-N. (2012). Study on the use of a weighted screening method for indoor positioning systems. In *The 15th International Symposium on Wireless Personal Multimedia Communications*.
- Lohan, E. S., Talvitie, J., e Silva, P., Nurminen, H., Ali-Loytty, S., & Piche, R. (2015). Received signal strength models for WLAN and BLE-based indoor positioning in multi-floor buildings. In *Localization and GNSS (ICL-GNSS), 2015 International Conference on* (pp. 1–6).
- Maple, C. (2003). Geometric design and space planning using the marching squares and marching cube algorithms. In *Geometric Modeling and Graphics, 2003. Proceedings. 2003 International Conference on* (pp. 90–95). <http://doi.org/10.1109/GMAG.2003.1219671>
- Moreira, A., & Meneses, F. (2015). Where@UM - Dependable organic radio maps. In *Indoor Positioning and Indoor Navigation (IPIN), 2015 International Conference on* (pp. 1–9). <http://doi.org/10.1109/IPIN.2015.7346751>
- Nam, S. Y. (2014). Localization of Access Points Based on Signal Strength Measured by a Mobile User Node. *Communications Letters, IEEE*, 18(8), 1407–1410.
- Savvides, A., Han, C.-C., & Strivastava, M. (2001). Dynamic fine-grained localization in Ad-Hoc networks of sensors. *Proceeding MobiCom '01 Proceedings of the 7th Annual International Conference on Mobile Computing and Networking*, 166–179. <http://doi.org/10.1145/381677.381693>
- Sibson, R., & others. (1981). A brief description of natural neighbour interpolation. *Interpreting Multivariate Data*, 21, 21–36.
- Talvitie, J., Renfors, M., & Lohan, E. S. (2015). Distance-Based Interpolation and Extrapolation Methods for RSS-Based Localization With Indoor Wireless Signals. *IEEE Transactions on Vehicular Technology*, 64(4), 1340–1353. <http://doi.org/10.1109/TVT.2015.2397598>

- Torres-Solis, J., Falk, T. H., & Chau, T. (2010). A review of indoor localization technologies: towards navigational assistance for topographical disorientation. INTECH Open Access Publisher.
- Torres-Sospedra, J., Montoliu, R., Mendoza, G., Belmonte, O., Rambla, D., & Huerta, J. (2016). Providing Databases for Different Indoor Positioning Technologies: Pros and Cons of Magnetic field and Wi-Fi based Positioning. *Mobile Information Systems*.
- Varzandian, S., Zakeri, H., & Ozgoli, S. (2013). Locating WiFi access points in indoor environments using non-monotonic signal propagation model. In *Control Conference (ASCC), 2013 9th Asian* (pp. 1–5).
- Wang, J., Urriza, P., Han, Y., & Cabric, D. (2011). Weighted Centroid Localization Algorithm: Theoretical Analysis and Distributed Implementation. *IEEE Transactions on Wireless Communications*, 10(10), 3403–3413. <http://doi.org/10.1109/TWC.2011.081611.102209>
- Werner, M. (2014). *Indoor Location-Based Services: Prerequisites and Foundations*. Springer.
- Zhang, Z., Zhou, X., Zhang, W., Zhang, Y., Wang, G., Zhao, B. Y., & Zheng, H. (2011). I am the antenna: accurate outdoor ap location using smartphones. In *Proceedings of the 17th annual international conference on Mobile computing and networking* (pp. 109–120).
- Zhao, F., Luo, H., Geng, H., & Sun, Q. (2014). An RSSI gradient-based AP localization algorithm. *China Communications*, 11(2), 100–108. <http://doi.org/10.1109/CC.2014.6821742>