# A Lightweight Indoor Localization Model based on Motley-Keenan and COST

Carlos Serodio, *Member, IAENG,* Luis Coutinho, Luis Reigoto,
Joao Matias, *Member, IAENG,* Aldina Correia, *Member, IAENG,* Pedro Mestre, *Member, IAENG*

***Abstract*—This paper presents a novel approach to WLAN propagation models for use in indoor localization. The major goal of this work is to eliminate the need for *in situ* data collection to generate the Fingerprinting map, instead, it is generated by using analytical propagation models such as: COST Multi-Wall, COST 231 average wall and Motley-Keenan. As Location Estimation Algorithms kNN (K-Nearest Neighbour) and WkNN (Weighted K-Nearest Neighbour) were used to determine the accuracy of the proposed technique. This work is based on analytical and measurement tools to determine which path loss propagation models are better for location estimation applications, based on Receive Signal Strength Indicator (RSSI).This study presents different proposals for choosing the most appropriate values for the models parameters, like obstacles attenuation and coefficients. Some adjustments to these models, particularly to Motley-Keenan, considering the thickness of walls, are proposed. The best found solution is based on the adjusted Motley-Keenan and COST models that allows to obtain the propagation loss estimation for several environments.Results obtained from two testing scenarios showed the reliability of the adjustments, providing smaller errors in the measured values values in comparison with the predicted values.**

***Index Terms*—LBS, Location Estimation Algorithms, Fingerprinting, Motley Keenan, COST.**

## I. INTRODUCTION

The trends of emerging technologies and services arise often the ubiquitous computing paradigm. So, we assist to a use of wireless technologies (IEEE802.11, Bluetooth, RFID, ...) in a widely spread way. We can say that wireless technologies became the alternative to the traditional broadband technologies. In a world even more under the ubiquitous computing paradigm it can be said that wireless communications and the characterization of location are inherent needs.

Manuscript received March 16, 2012.

C. Serodio is with CITAB - Centre for the Research and Technology of Agro-Environment and Biological Sciences, University of Trás-os-Montes and Alto Douro, Vila Real, Portugal, email: cserodio@utad.pt

L. Reigoto is with UTAD - University of Trás-os-Montes and Alto Douro, Vila Real, Portugal, email: luisreigoto@gmail.com

L. Coutinho is with UTAD - University of Trás-os-Montes and Alto Douro, Vila Real, Portugal, email: luis_ coutinho_86@hotmail.com

J. Matias is with CM-UTAD - Centre for the Mathematics, University of Trás-os-Montes and Alto Douro, Vila Real, Portugal, email: j_matias@utad.pt

A. Correia is with ESTGF-IPP, School of Technology and Management of Felgueiras Polytechnic Institute of Porto, Portugal, aldinacorreia@eu.ipp.pt

P. Mestre is with CITAB - Centre for the Research and Technology of Agro-Environment and Biological Sciences, University of Trás-os-Montes and Alto Douro, Vila Real, Portugal, email: pmestre@utad.pt

The main objective of this work is to develop a propagation model suitable to determine indoor localization, based on empirical models, which outputs are the path loss and the effect of attenuation due to obstacles, like walls and furniture, with a low computational time. Obtaining a good performance, in terms of computation time, might result in a lower accuracy, when compared with other types of models.

In wireless network coverage problems we deal essentially with two types of propagation models: (1) deterministic models which need a great amount of information about the environment, and are slightly heavier in terms of computational effort; (2) empirical models which need less information about the environment, which results in the use of simpler mathematical expressions and lower computational time and effort.

Another important feature of this work is the elimination of the off-line phase to generate the Fingerprinting Map (FM). Instead of generating the FM using the traditional procedures, which can be hard and time consuming due to the need to measure the wireless signal in all important points of the scenario, we propose the use of empirical models like COST231 Multi-Wall, COST Average Walls and Motley-Keenan [1]. Another important issue is the fact that whenever the scenario suffers changes, e.g. due to the removal or inclusion of furniture, it might imply the need for an update to the FM. This means that the FM generation procedures must be executed again [2].

These type of tasks are very simplified if empirical propagation models are used. A computing system can update the new features/labels of scenario in a easy way. Besides this, the generation of new scenarios or updating of existing ones can be made in an automate way, as presented by the authors in previous works [3],[4].

From the literature [5],[6] we can find several works related with propagation in indoor environments using empirical models and some of these models consider the attenuation due the penetration of walls. This effect depends, among other parameters, on the thickness and construction material of the walls. However, typically the models divide the wall into groups, but do not specify how to evaluate the effects of its the width. The present work takes special attention in the adjustment of propagation models in way to minimize the errors that depend on wall characteristics.

In the tests presented in this work, location estimation was made using consumer electronic devices such as smartphones and laptops and existing wireless network infrastructures.



ISBN: 978-988-19252-1-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

## II. Problem analysis

The prediction of propagation characteristics, between two transceivers, in indoor environments is important specially for the design of LBS services and applications such as ubiquitous computing environments or smart spaces.

A major limitation to the localization estimation are the irregular values for the Received Signal Strength (RSS). This is due to the short distance between the transmitter and receiver, in indoor environments, highly affected by the attenuation caused by obstacles.

Typically the propagation models are divided into two classes: (1) empirical models; (2) deterministic models. The first type is represented by simple mathematical equations which provide the path loss as output. The others are computational models which simulate the behaviour of propagation of radio signals. As one of the main objectives of this work is to obtain a lightweight model to estimate indoor localization, the use of empirical models is a better option because it minimizes the computational time and the needed resources. It might however penalize the accuracy.

On the other hand, if we simplify too much the computation stage, it could lead to a FM for which the accuracy of the location methods would be very low. At this point it is important to define what resolution do we need to obtain from the designed method.

The most simple propagation model is the classic One-slope propagation model, Eq. 1, which widely used in several environments [2][6]:

$$L_{dB} = L_{0,dB} + 10nLog_{10}d \quad (1)$$

where $L_{0,dB}$ is the path loss obtained at the distance of *1 meter* from emitter and *n* is a coefficient experimentally determined [7].

This model is a general path loss model that is good for a first approach to the problem, but does not consider the effect of obstacles like walls, windows and furniture.

If we want to consider the effects of obstacles we can change this model into the Dual-Slope Model. According to [2] the Dual-Slope model gives a better accuracy, however the generation of the FM is more complex and the computation time is higher. One reason for this is because the Dual-Slope model has different equations to LOS and NLOS and also computes the relative height of antennas [2].

## III. Test Bed

It is not an easy task to compare different indoor propagation models, because they have several and different parameters, such as the type (glass, concrete, wood,...), thickness of walls, furniture, etc. Small differences in the parameters values could be reflected in discrepancies on the output values of the different propagation models. In this work the COST Multi-wall, COST average wall and Motley-Keenan models were chosen to evaluate its suitability for the characterization of propagation in indoor environments. The comparison between these models is supported by several tests and modifications over the considered models.

Two testing scenarios were chosen, which have different configurations and type of rooms, such as classrooms, corridors and Electrical Machines Laboratory, with several electronic equipments. This means that it was considered some rooms that could be classified as hazardous regarding of view electromagnetic radiations, due the presence of several metal cabinets. The effect of metal furniture in the signal propagation can have at least two consequences: it might work as a wave guide, for example on corridors, or it is a very serious obstacle when they are near the walls.

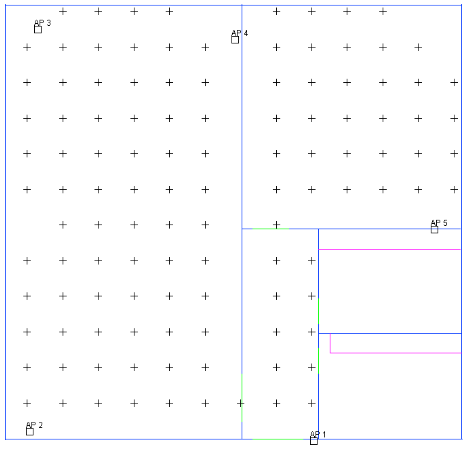

Figure 1. Scenario 1

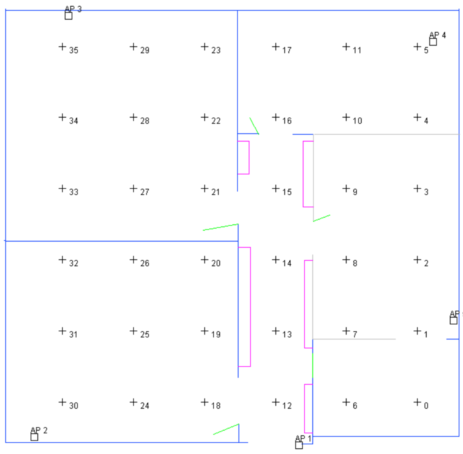

Figure 2. Scenario 2

The two testing scenarios are depicted in Fig. 1 and Fig. 2. The first one is a typical floor with classrooms and offices, while the second is constituted by corridors with metal cabinets, laboratories with electronic and electrical equipments and a higher number of walls.

At the begin it was used the values for the parameters found in the bibliography [2][6] as well as in authors' previous works [4],[8].

As starting point to Motley-Keenan model we have considered that in Free-Path Loss *n=2* and in indoor environments *"n will be bigger, closest to 3"*, as stated by [7].

The equations of the methods to be evaluated for indoor environments are:

- COST 231 Multi Wall:

$$L_{dB} = L_0 + 20Log(d) + k_f^{\frac{k_f+2}{k_f+1}-b} L_f + \sum_{i=1}^{k_w} k_{wi} L_{wi} \quad (2)$$


ISBN: 978-988-19252-1-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

- Free Space Loss:

$$L_{0,dB} = -32.44 - 20Log(f_{MHz}) - 20Log(d_{km}) \quad (3)$$

- COST Average Wall:

$$L_{dB} = L_0 + 20Log(d) + k_w L_w \quad (4)$$

- Motley-Keenan:

$$PL(d)[dB] = PL_R + 10nLog(d) + \sum_{i=1}^{k_w} k_{wi} L_{wi} 2^{Log_3 \frac{e_i}{e_{oi}}} \quad (5)$$

As stated in the equations 2 and 4 the free-space path loss is obtained by Eq. 6:

$$L_0 = \left(\frac{4\pi d_0}{\lambda}\right)^2 \quad (6)$$

where for $f = 2.4GHz, \lambda = 0.125, d_0 = 1m$ then the 6 is *40.2dB*, as in [4] and by definition.

COST and Motley-Keenan models are very similar due to the fact that they calculate the propagation loss based on path loss in free space added by the loss (attenuation) due to obstacles (walls and furniture).

Besides the difficulties to find a solution that works quite well with all models, it will be presented some approaches based on kNN (k-Nearest Neighbour) and WkNN (Weighted k-Nearest Neighbour) and the parameters needed tune them. With this work it is possible to evaluate the performance of each propagation models comparing their results in real environments. This allows also to fine tune the parameters in a more accurate way.

The first approach to COST and Motley-Keenan models are very simple because we just need identify the number of walls between the transmitter and the receiver. At this point we have the first simplification of models. We just consider two types of Walls: Thin $(-3, 4dB)$ and Thick $(-6, 4dB)$.

The second approach, that was demonstrated by the measurement test, is: *"if you have several walls between the transceivers, the effect usually it is not linear, this means that the attenuation is often small when compared with the addition of the individual effects"*. So we use empirical methods that are similar to the COST average wall method. At this experiments we only consider a bi-dimensional system (2D), therefore the third member of Eq. 2 is not considered. To simplify even more the computation time we have opted by a variation of this model that results in Eq. 4.

The generation of FM, which is the first step execute, is made using the empirical models. The greatest challenge is to generate the FM based only on the information on the blueprints [8]. While in Eq. 2 and Eq. 5 model it very straightforward, when we use Eq. 4, some interaction with the user, who decides the simplification to do, is needed.

## IV. Localization Estimation Algorithms

To estimate the indoor localization the classical k-Nearest Neighbour (kNN), [9], and Weighted k-Nearest Neighbour(WkNN) methods were used. It was performed a set of tests in which it was applied the traditional approaches as well as some variations of these methods regarding to offset compensation, and weight determination.

### A. The KNN Location Estimation Algorithm

The kNN is the simplest of all algorithms for predicting the localization. Regarding the use of this methods it must be defined the value of $k$ and the distance function. Usually the value of $k$ is 3. To define the distance the most common distance function is the Euclidean distance, Eq. 7:

$$d(x, y) = ||x - y|| = \sqrt{(x - y)(x - y)} = \left(\sum_{i=1}^{m}(x_i - y_i)^2\right)^{1/2} \quad (7)$$

where $x$ and $y$ are points in $X = R^m$, $m$ represents the dimension of the problem, this means that are performed $m$ measurement for each point to estimate. Typically the major dimension found in literature is around 20, [10], and still having a lightweight problem. In many literature this value works like a limit. However it was demonstrated that the accuracy depends on the number of samples registered in the FM, because the signal strength varies widely and these variations over the time result in the introduction of errors or discontinuities in users' trajectory [7].

The kNN LEA is based on the identification of the closest neighbour or the $k$ closest neighbours. This method consists in searching in the Fingerprinting Map for the closest $k$ neighbours following the criterion:

$$X = argmin\left(\sum_{l=1}^{N}(P_l(x, y) - P_{FMl}(x_k, y_k))^2\right) \quad (8)$$

where $P_l$ represents the RSSI value received from AP $l$, $N$ in the number of Access Points and $(x_k, y_k)$ are the coordinates on spatial domain. As this method is strongly affected by the variations of RSSI, several stages of compensation are made:

1) Compute the RSSI average for $m$ samples: $\sum_{l=1}^{m}(P_{l,dB}(x, y)/m;$
2) Use one calibration point, by default point $(0, 0)$;
3) Compensate the measurements variations, using the following procedure: the point to be computed with FM is $P_l(x, y) - \sum_{l=1}^{m}(P_l(x_{0l}, y_{0l}) - P_{rl}(x_{0rl}, y_{0rl}))/m$, where $l$ is the index of measures and $rl$ is index of the reference point in the FM;
4) Compute the previous point with all points of FM, and identify the $k$ closest points;
5) Finally the point to be considered will be the point resulting from the average of the $k$ closest points;
6) Repeat the procedure for all points.

To evaluate the performance of the proposed adjustments to Motley-Keenan and COST models (accuracy and precision) the Means Relative Error (MRE) and the Standard Deviation $(\sigma)$ between measured values and calculated values were computed. The adjustments resulting from the first test, which works as a validation to the adjustments, are presented in Tab. I. These results were obtained in Scenario 1 using a HTC smartphone. The methodology used to estimate the

ISBN: 978-988-19252-1-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Table I
KNN IN SCNENARIO1 WITH HTC

| Model | Mean err. | $\sigma$ | Max. err. | % err.< res. |
|---|---|---|---|---|
| COSTMW n=2; $Lw$=3.4 | 3,188 | 1,889 | 9,014 | 46.67 |
| COSTAW n=2; $Lw$=3.4 | 2,611 | 1,512 | 6,719 | 60.00 |
| Motley n=2; $Lw$=10 | 2,792 | 1,378 | 5,069 | 43.33 |
| Motley n=2.5; $Lw$=10 | 2,774 | 1,473 | 5,590 | 53.33 |
| Motley n=2.9; $Lw$=10 | 2,852 | 1,389 | 5,590 | 50.00 |

Table II
KNN IN SCNENARIO1 WITH PC

| Model | Mean Err. | $\sigma$ | Maxi. err. | %err. < res. |
|---|---|---|---|---|
| COSTMW n=2; $Lw$=3.4 | 2.859 | 1.793 | 7.682 | 50.00 |
| COSTAW n=2; $Lw$=3.4 | 3.216 | 1,943 | 9.317 | 56.67 |
| Motley n=2; $Lw$=10 | 2,979 | 1,711 | 7,861 | 53.33 |
| Motley n=2.5; $Lw$=10 | 3,067 | 2,079 | 9.310 | 56.67 |
| Motley n=2.9; $Lw$=10 | 2,864 | 1,556 | 7,453 | 50.00 |

Table III
CALIBRATION OF KNN IN SCNENARIO1 WITH HTC

| Model | Mean err. | $\sigma$ | Max.err. | % err. < res. |
|---|---|---|---|---|
| COSTMW n=2; $Lw$=3.4 | 3,141 | 1,801 | 8,448 | 40.00 |
| COSTAW n=2; $Lw$=3.4 | 2,695 | 1,630 | 6,719 | 60.00 |
| Motley n=2; $Lw$=10 | 2,610 | 1,760 | 6,298 | 53.33 |
| Motley n=2.5; $Lw$=10 | 3,017 | 1,670 | 6,832 | 53.33 |
| Motley n=2.9; $Lw$=10 | 2,920 | 1,603 | 6,832 | 50.00 |

Table IV
KNN AND WKNN TEST IN SCNENARIO1 WITH SONY

| LEA | kNN | | WkNN | |
|---|---|---|---|---|
| Model | Mean Err. | $\sigma$ | Mean Err. | $\sigma$ |
| COSTMW n=2; $Lw$=3.4 | 3,657 | 2,650 | 3,550 | 2.423 |
| COSTAWn=2; $Lw$=3.4 | 2,683 | 2,043 | 2,950 | 1.839 |
| Motley n=2; $Lw$=10 | 3,280 | 2,031 | 3,000 | 1.896 |
| Motley n=2.5; $Lw$=10 | 3,365 | 2,110 | 2,832 | 1.810 |
| Motley n=2.9; $Lw$=10 | 3,334 | 1,992 | 3,133 | 2.025 |
| Motley n=2.9; $Lw$=16 | 2,921 | 1,942 | 3,093 | 1.924 |

point was based on the average of measurements (signal domain) and the average of the $k$ closest points.

The same tests were made using an IBM laptop and with a Sony Android smartphone. In the case of the laptop the best results were obtained with kNN but with the $k$ closest points determined using the mean of $m$ measurements at each point, Tab. II. For the smartphone, using kNN, the best results were obtained with Motley-Keenan ($n = 2.5$, $L_w = 10$, mean error=3.09, $\sigma = 1.633$). When performed the calibration using kNN with the mean deviation the best results were also obtained with Motley-Keenan.

Regarding to the HTC calibration and from Tab. III the better results are obtained with Eq. 4 and Eq. 5 ($n = 2; L_w = 10$). It was defined that resolution of FM works like the threshold (the lower than the resolution is the result, the better it is). Since the Motley-Keenan model is the main object of study of this paper it was made an effort to enhance its performance, i,e., obtain more points with error value below the resolution. Besides this from the analysis of other output data it was noted that the measured reference point $(0, 0)$ did not match exactly with the reference point. The first approach was to consider that deviation as an *offset* and apart this as measurement values present variations we use the *mean-deviation* either to make the compensation task. The procedure is as follows:

1) Step (1), (2), (3) and (4) are similar to the previous methodology;
2) After that, it is computed the mean of the $k$ closest points to determined the reference point, by default $(x_0, y_0)$;
3) Determine the value of the mean deviation between the $k$ closest points to the reference point: $Mean.Deaviation = Mean(ABS(X_{0l}, Y_{0l}) - Mean(X_{0i}, Y_{0i}))$, $P_{0,md}(x_0, y_0) = \left( \sum_{l=1}^{k} |(x_{0l}, y_{0l}) - \frac{\sum_{i=1}^{k} P(x_{0i}, y_{0i})}{k}| \right) /k$
4) For all the other points calculate the kNN. The point that is effectively considered is obtain with the expression: $P_{l,md}(x, y) = \sum_{l=1}^{k} P_l(x_{0l}, y_{0l})/k - P_{0,md}(x_0, y_0) + P_{rl,md}(x_{0rl}, y_{0rl})$
5) Finally the point of the spatial domain is obtained by multiplying the coordinates of $P_{l,md}(x, y)$ by the resolution of FM (in present case 2.5 meters).
6) Repeat the procedure for all points.

To evaluate the suitability of the calibration procedure, these tests were made with a Samsung Android at Scenario 2, and the results are shown in Tab. IV. Since the global results were a little worse, it was obtained a new balance between $n$ and $L_w$. Using kNN the best results were obtained using Motley-Keenan ($n = 2.5, L_w = 6$) and with the calibration the best case was Eq. 4. These new parameters values were tested with HTC and PC and we had obtained lower values for *mean error* and $\sigma$. However due the fact that scenario 2 is very hazardous, from the propagation point of view, the maximum error is higher, more than 3 times the spatial resolution of FM. In this case it is harder to find a correct balance between free space loss and losses due walls and metal cabinets.

### B. The WKNN Location Estimation Algorithm

The Weight k-Nearest Neighbour (WkNN) is a variation of kNN in which each of the $k$ closest neighbours has different influence in the determination of the final point. This means that after the measurements it must be taken in account the significance of each point in way to determine the correct weight of each neighbour.

Several approaches to define the weight of each neighbour exist: static attribution (e.g. $w_1 = 0.7, w_1 = 0.2, w_1 = 0.1$ based on empirical tests [4]) or dynamic attribution where each point has also a different weight, but the weight is computed on the fly during the online phase. However, in

ISBN: 978-988-19252-1-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

the last approach we can have two situations: the weight depends on the order of the $k$ closest neighbour (the first closest has the highest weight, the second lower, and so on); or it can be defined an algorithm that considers the $k$ closest neighbours as a set and then computes them to find which one could be the most representative, and then attribute to it the higher weight. This approach results from the observation of variations of the measured points, because frequently the real nearest point does not appear as the closest (appears as second, third, or $k-th$ significant point).

As the measurement is the RSSI value, then it must be transformed into a position using LEA techniques presented at the previous sections. These approaches are based on Dudani's work in which it is proposed a distance-weighted kNN method which assigns to the $i-th$ nearest neighbour a weight $w_i$ as a function of its distance, as defined by Eq. 9:

$$w_i = \frac{d_k - d_i}{d_k - d_1} \tag{9}$$

for $d_k \neq d_1$ and equal to 1 when $d_k = d_1$. Dudani demonstrated that this approach allows to yield lower error rates than obtained in kNN.

Other weighting function that can be employed by the distance-weight k-Nearest Neighbour is given by Eq. 10:

$$w_i = \frac{1}{d_i}, d_i \neq 0 \tag{10}$$

and the weights are bigger for distances $d_i$ closer to zero which could be redundant with the kNN and in some case has the constraint of the division by zero. In this case, it would be useful adding a small constant to the denominator. Latter, Baily and Jain [11] showed that the distance-weight is not always better than the traditional kNN. Macleod made a modification to Eq. 9 in [12]. Based on the advantages and disadvantages of the methods above described it was applied Eq. 11 method [13] which uses an exponential weighting following the rule:

$$w_i = exp(-\alpha d_i^{\beta}) \tag{11}$$

in this work it was used $\beta = 1$.

However, the distance measure can not be computed directly because it is not monotonic, which implies the need to be interpreted. The real nearest neighbour may lie far away, and not appear as the k-closest or often could not be included in the set of $k$ nearest points. To minimize these problems, it must be a linear approximation based on the following issues: $i)$ instead using the measured k-closest point directly we firstly determine the addition of the difference between them; $ii)$ use the parameter $\alpha$ from Eq. 11 as compensation. Since $\alpha$ can switch between 1 or $-1$, depending on previous computation and it used to compensate the $x$ axis.

To validate all approaches several tests, using different modes of distance-weight determination, were made.

*1) Experiment 1:* To perform the first set of tests it was considered that the distance-weights are determined as following: (1) compute the average of the x-coordinates and y-coordinates of the k-closest points. The distance to be evaluated is obtained by de difference between the $x_i$ or $y_i$) and the correspondent average value. $x_i = x_l - \sum_{l=1}^{k} x_l/k$ (2) based on the signal of new points $(x_j; y_j)$ it is defined the value of $\alpha$. After that, the Euclidean distance for each point $x = \sqrt{x_i^2 + y_i^2}$ is calculated. Then the individual distance-weight is determined using Eq. 11 with $\beta = 1$ and $\alpha$ as compensation parameter. Finally the point to be considered as the position is obtained using Eq. 12:

$$P(x,y) = (\sum_{i=1}^{k} x_i w_i; \sum_{i=1}^{k} y_i w_i) \tag{12}$$

*2) Experiment 2:* In the second set of tests it was considered some variations in the determination of the distance-weights, which are determined as following: (1) for each measured point compute the difference of x-coordinates and y-coordinates between itself and the other $(k-1)$-closest points. The point that presents the lowest values of $(x, y)$ is assumed to be the more significant point, this means that it is the closest point and must have the higher weight. The distance to be evaluated is obtained by de difference between the $x_i$ or $y_i$ and the correspondent average value.

After that, the Euclidean distance for each point $x = \sqrt{x_i^2 + y_i^2}$ is calculated. Then the individual distance-weight is determined using Eq. 11 with $\alpha = \beta = 1$. Finally the point to be considered as the position is obtained by:

$$\forall i \in \{1, 2, 3, ...k\}, x_i = \sum_{j=1}^{k} x_i - x_j \tag{13}$$

After this point repeat the procedures of previous experiment, i.e., calculate the Euclidean distance for each point and then calculate the individual distance-weight using Eq. 11 with $\beta = 1$ and $\alpha$ to compensate the non monotonic behaviour of the measures. Finally the point to be considered as the position is obtained using Eq. 12. The results obtained in experiment 1 and 2 are showed in Tab. V.

Analysing the results, we can conclude that they are very encouraging. We have got several position estimations with the following characteristics: *Average Error = around the Resolution*; *standard deviation = closest to the half of resolution*. Regarding to the *Maximum Error* the objective is to obtain a value not greater than *twice the resolution*, but in some cases it still is a little bit far away from the objective.

Finally, to validate the approaches described in Experiment 1 and 2 it was chosen other smartphone, a Samsung Android, with the previous propagation models and LEA (Location Estimation Algorithms) to be applied to the scenario 2. The objectives are: test the independence of the models to the characteristics of the devices; promote the application on hazardous scenarios (from the propagation point of view).

From Table VI we can identify two undesired factors: the average error is slightly higher than the desired value; the maximum error is too big. This is related with the ratio between $n$ and *attenuation due the wall* balance. So, to minimize or even eliminate this problem it was made another test bed with Motley-Keenan model with $n = 25; L_w = 6$ and the results was better: *Experiment 1: δ=2.887; σ=1.762; Error_max=8.939*; *Experiment 2: δ=2.681; σ=1.478; Error_max=9.378*

ISBN: 978-988-19252-1-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Table V
RESULTS OF WKNN OF HTC AND PC ON SCNENARIO2 UNDER EXP1 AND EXP 2 CONDITIONS

| | COSTMW | | | COSTAW | | | Motley n=2 | | | Motley n=2,5 | | | Motley n=2,9 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\delta(m)$ | $\sigma(m)$ | $E_m(m)$ | $\delta(m)$ | $\sigma(m)$ | $E_m(m)$ | $\delta(m)$ | $\sigma(m)$ | $E_m(m)$ | $\delta(m)$ | $\sigma(m)$ | $E_m(m)$ | $\delta(m)$ | $\sigma(m)$ | $E_m(m)$ |
| HTC Exp1 | 3.128 | 1.785 | 9.014 | 2.423 | 1.501 | 6.610 | 2.651 | 1.340 | 5.409 | 2.733 | 1.442 | 5.591 | 2.846 | 1.387 | 5.591 |
| PC Exp1 | 3.128 | 1.816 | — | 2.532 | 1.515 | — | 2.604 | 1.350 | — | 2.746 | 1.472 | — | 2.871 | 1.419 | — |
| HTC Exp2 | 3.353 | 1.667 | 8.545 | 2.781 | 1.619 | 7.478 | 2.649 | 1.296 | 5.00 | 2.680 | 1.676 | 6.310 | 2.898 | 1.419 | 6.072 |
| PC Exp2 | 3.034 | 1.733 | 7.903 | 3.102 | 1.733 | 7.900 | 2.801 | 1.405 | 7.903 | 2.795 | 1.454 | 7.903 | 2.491 | 1.683 | 7.975 |

Table VI
RESULTS OF SAMSUNG ANDROID ON SCENARIO2 UNDER EXP1 AND EXP 2 CONDITIONS

| | COSTMW | | | COSTAW | | | Motley n=2 | | | Motley n=2,5 | | | Motley n=2,9 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\delta(m)$ | $\sigma(m)$ | $E_m(m)$ | $\delta(m)$ | $\sigma(m)$ | $E_m(m)$ | $\delta(m)$ | $\sigma(m)$ | $E_m(m)$ | $\delta(m)$ | $\sigma(m)$ | $E_m(m)$ | $\delta(m)$ | $\sigma(m)$ | $E_m(m)$ |
| Samsung Exp1 | 3.851 | 2.580 | 11.562 | 2.553 | 2.006 | 11.069 | 3.178 | 1.922 | 10.056 | 3.437 | 2.035 | 10.274 | 3.469 | 1.998 | 10.274 |
| Samsung Exp2 | 3.140 | 1.591 | 10.306 | 2.552 | 1.280 | 7.923 | 3.130 | 1.590 | 7.923 | 3.356 | 1.735 | 9.879 | 3.255 | 1.571 | 9.879 |

After this test and because the maximum error was higher than the desired value, the map was analysed and three points that were misclassified, which caused a sharp deviation in relation to the actual point, were identified. This situation must be corrected by the user intervention. So, it was proved that the approaches described can be implemented under any environment and with any device. The major challenge is obtain a reliable FM from the blueprints.

## V. CONCLUSION AND FUTURE WORK

In this paper it was presented a comparison of different location methods based on empirical propagation models and on location estimation algorithms for indoor environments.

For the models that consider the attenuation due to walls, the best results for the accuracy were obtained using average or simplified models in comparison with original COST231.

It was demonstrated that the localization estimation using empirical models for indoor environments with multiple walls can be obtained without the need of knowing in advance the penetration loss of the wall.

The COSTAW and Motley-Keenan models were a good choice for use with automatic generation of the FM. Dynamic weight based algorithms produce an enhancement to the accuracy over the other presented methods, and leads to a better localization estimation accuracy. This approach seems to be suitable to compensate the variations due either the fluctuation measurement values and the fact that the measures have not a monotonic behaviour, even for sequential points. In this case the Zavrel distance-weight approach demonstrated to be suitable to do the compensation task based on $\alpha$ to define the signal of tendency.

With the new approaches and assuming that FM resolution working like a *threshold* it was obtained a satisfactory accuracy: *Average Error = close to the resolution value*; *standard deviation = closes to the half of resolution*, and the *Maximum Error= close to twice the resolution value*.

Finally, from the overview of the results we can conclude that the propagation models represent a good solution to generate Fingerprinting Maps which eliminate the need for data collection in all significant points of the scenario, and eliminate the hard work need to generate FM in off-line. This solution gives a good approach to the FM by computing the empirical models based on the blueprints. Even when mapping RSSI values to position labels, the computation time and needed resources are low.

## ACKNOWLEDGEMENT

This work is supported by European Union Funds (FEDER/COMPETE - Operational Competitiveness Programme) and by national funds (FCT - Portuguese Foundation for Science and Technology) under the project FCOMP-01-0124-FEDER-022696.

## REFERENCES

[1] A. Motley and J. Keenan, “Personal communication radio coverage in buildings at 900 MHz and 1700 MHz,” *Electronics Letters*, vol. 24, no. 12, pp. 763 –764, jun 1988.

[2] C. Andrade and R. Hoefel, “Ieee 802.11 wlans: A comparison on indoor coverage models,” in *Electrical and Computer Engineering (CCECE), 2010 23rd Canadian Conference on*, may 2010, pp. 1–6.

[3] C. Serodio, L. Coutinho, H. Pinto, and P. Mestre, “A comparison of multiple algorithms for fingerprinting using IEEE802.11,” in *Lecture Notes in Engineering and Computer Science: Proceedings of The World Congress on Engineering 2011, WCE 2011, 6-8 July, 2011, London, U.K.*, 2011, pp. 1710–1715.

[4] P. Mestre, H. Pinto, C. Serodio, J. Monteito, and C. Couto, “A multi-technology framework for lbs using fingerprinting,” in *Industrial Electronics, 2009. IECON ’09. 35th Annual Conference of IEEE*, 2009, pp. 2693 –2698.

[5] M. Lott and I. Forkel, “A multi-wall-and-floor model for indoor radio propagation,” in *Vehicular Technology Conference, 2001. VTC 2001 Spring. IEEE VTS 53rd*, vol. 1, 2001, pp. 464 –468 vol.1.

[6] A. Zaballos, G. Corral, A. Carné, and J. L. Pijoan, “Modeling new indoor and outdoor propagation models for WLAN,” in *OPNETWORK’2004*, 2004, pp. 1–6.

[7] F. Evennou and F. Marx, “Improving positioning capabilities for indoor environments with WiFi,” in *IST Summit*, 2005.

[8] P. Mestre, L. Coutinho, L. Reigoto, J. Matias, A. Correia, P. Couto, and C. Serodio, “Indoor location using fingerprinting and fuzzy logic,” in *Advances in Intelligent and Soft Computing*, no. 107. Springer-Verlag Berlin and Heidelberg GmbH & Co. K, 2011, pp. 363–374.

[9] C. Elkan, “Nearest neighbor classification,” 2007, online, available at http://cseweb.ucsd.edu/ elkan/151/nearestn.pdf.

[10] A. M. Kibriya and E. Frank, “An empirical comparison of exact nearest neighbour algorithms,” in *PKDD’07*, 2007, pp. 140–151.

[11] T. Baily and A. K. Jain, “A note on distance-weighted k-nearest neighbor rules,” *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 8, no. 4, pp. 311 –313, april 1978.

[12] J. Macleod, A. Luk, and D. Titterington, “A re-examination of the distance-weighted k-nearest neighbor classification rule,” *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 17, no. 4, pp. 689 –696, july 1987.

[13] J. Zavrel, “An empirical re-examination of weighted voting for k-nn,” in *Proceedings of the 7th Belgian-Dutch Conference on Machine Learning*, 1997, pp. 139–148.