Using neural networks and Active RFID for indoor location services

Alejandro Santos Martínez Sala, Raúl Guzman Quirós, Esteban Egea López, Polytechnic University of Cartagena, Spain

Abstract

Indoor RTLS (Real Time Location Systems) are the foundations of promising context-aware and ambient intelligent services. In this work the feasibility of applying Active RFID and neural networks to develop a RTLS service is discussed. In most of the Active RFID systems available on the market, the Readers can measure the Received Signal Strengh (RSS) from a beacon transmitted from a Active tag and send the data gathered to a location server. The RSS measurements can be processed to infer the position of a tag by means of a positioning algorithm. In this research work we discuss and show how to use RSS measurements from the Readers to calculate the tag position by means of a neural network, based on a Multilayer Perceptron which is trained and tested with a radiomap and learns to compute the tags position. By means of simulation, we study the proper MLP architecture and the mean error positioning estimation and precision that is achieved depending on the number of Readers. With 8 Readers deployed in an indoor area of 576 m^2 we get an error less than 1.75 meters in the 75% of the target area.

1 Introduction

Location systems are key for developing context aware services [1]. Promising services in indoor environments are foreseen such as navigation guide systems in museums, airports, for visual impaired, tracking of patients or high-value equipment in hospitals, etc. However, all of these require a positioning technology in order to monitor the position with a reduced and bounded error. GPS technology does not work out indoors because the signals from the satellites are too weak or even null. Therefore other wireless positioning technologies suitable for indoor environments are needed.

We propose using Active RFID at 2,4 GHz and a processing engine based on neural networks to infer the tag position. The only assumption is that the Readers can measure the Received Signal Strength from a tag beacon. Our proposal can be extrapolated to any Active RFID system available on the market that complies with this assumption.

A positioning system calculates the position by means of measuring a signal generated between a tag (attached to a person or a mobile object) and the network of Readers; second, a location algorithm processes the gathered signals to estimate the position (x,y). There are several kind of measurements from a RF signal which are useful for a location algorithm: *Angle of Arrival* (AoA), *Time of Arrival* (ToA) y *Received Signal Strength* (RSS) [1].

The indoor wireless channel at 2,4GHz is rather complex to characterize and model due to its stochastic nature. There are challenging factors to take into account such as the common Non Line of Sight communication link between the tags and Reader because of the building geometry, obstacles like walls, furniture or even mobile objects or people; also the multipath and fading effects on the RF signal, etc. Therefore it is well known that the RSS measure does not follow an ideal deterministic model and its real behavior is environment dependant.

The positioning algorithms could be classified into distance based or pattern recognition problem. A distance based algorithm determines the distances between the tag and three or more reference points and it locates tags through triangulation. Due to the random variability of an indoor wireless channel, a distance based algorithm suffers from high errors and it is not advisable for indoor positioning.

A positioning algorithm based on pattern recognition employs the gathered signals from known (x,y) points from a building (called the radiomap), like the Received Signal Strength, which conforms a particular pattern. This technique is called the fingerprint method based on RSS and it requires an intensive measurement campaign, called calibration phase, in order to build the radiomap of the target environment.

Once we have the radiomap database, during the online phase the matching algorithm compares the characteristics of the observed signal with the existing fingerprints in the radio map and it chooses the reference point that matches best with the observed data and estimates the tags current position. The main advantage of the fingerprint method is that the radiomap characterizes and takes into account the random features and complexity of the indoor wireless channel. The main drawback of the fingerprint method is that it needs a big amount of data collection and manual labor for creating the radiomap. The positioning accuracy depends on the number of samples of the dataset.

There are several proposals for positioning pattern matching algorithms based on RSS fingerprint: RADAR [2] was one of the first proposals of indoor location systems with WiFi networks and it uses the K-Nearest Neighbors (KNN) algorithm: it is the most basic algorithm which compares an observed RSS vector with all available fingerprints in the reference radiomap and finds a reference point from the radiomap with the smallest Euclidean distance. LANDMARC [3] is an indoor Active RFID system that uses a network of Readers with a grid of reference tags (landmarks) that use the KNN algorithm. EKAHAU [4] is a commercial Active RFID system based on WiFi standard that uses probabilistic methods based on bayesian networks. Aeroscout [5] is another commercial Active RFID system, based on WiFi with RTLS services, which employs the RSS fingerprint method.

Battiti *et. al.* [6] was the first proposal to employ a neural network for location estimation using a WLAN infrastructure and mobile devices WiFi-enable in an indoor environment. In [7] and [8] a neural network is used for location estimation in a WLAN but they employ a fix amount of WiFi Access Point and they don't study the impact of the number of Access Point on the positioning error.

In this paper a Multilayer Perceptron (MLP) neural network is evaluated as a positioning method. The impact of learning parameters and different MLP architectures is considered in the positioning estimation performance within a typical indoor environment of *576m²*. By simulation we analyze the influence of the number of deployed Readers on the mean error and precision.

The rest of the paper is organized as follows: In section II the implemented indoor channel and physical model are presented, as well as the main features of the simulation tool. Then the performance metric to evaluate the Active RFID positioning system is defined. In section IV the MP architecture and its main configuration parameters are explained. In section V the results are discussed and, finally Sect. VI shows the conclusions.

2 Channel and environment model

2.1 Scenary description

In figure 1 the 2D layout of the simulated environment is depicted. It is an area of $576m^2$ of teaching labs from the Polytechnic University of Cartagena composed of a main corridor and several rooms at both sides of the corridor.

2.2 Channel Model at 2.4GHz

In [9] an intensive measurement campaign at 2,4 GHz in our target environment is explained and it is justified that a path-loss shadowing model that takes into account signal attenuation due to walls and obstacles [10], characterizes the wireless channel behavior with high precision. In equation (1) the expresion of this model is represented, where L(d) is the signal attenuation (expressed in decibels), at a distance d between the emitter and the receiver.

$$L(d) = L_0 + \sum L_{obs} + 10\alpha \ \log_{10}(d) + X \qquad (1)$$

The formula has four terms: L_0 represents signal losses at a reference distance of 1 meter, L_{obs}(dB) is the contribution of walls and obstacles to the signal attenuation, in [10] the values depending on the kind of obstacles and materials are defined. The logarithm term is the path-loss at a distance d with path-loss coeficient α. Finally X is the term related to the shadow fading and it is modelled with a gaussian random variable of 0 mean and variance σ_X .

$$f_x(x) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{x^2}{2\sigma_x^2}} \qquad (2)$$

2.2 Simulator Tool

The main assumptions concerning the Active RFID system are: a) an Active RFID tag is allowed to move through the area and can transmit periodically RF beacons at a fix power. b) Tag and Reader antennas are isotropic. c) Readers are at a fix position and a height of 2.2 meters. Also, a tag moves at a mean height of 1.5 meters. d) Readers are capable of measuring the RSS from a tag

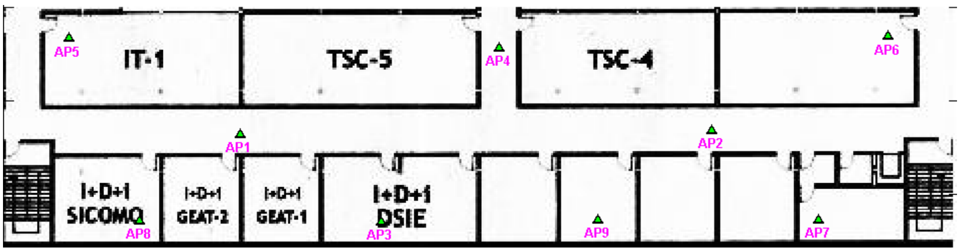

Fig.1 Floor Plan of the Building used in our simulations.

beacon sending the data gathered to a location server connect to the Ethernet network.
The simulation tool has been implemented with Matlab and it uses the Neural Network toolbox [11] for the development of the positioning engine. The simulator allows to build a 2D layout of the indoor environment made up of plans, corridors, room, obstacles and walls. A 3D cartesian reference system is used where the Readers are placed at fix positions.
The physical layer follows the channel model from [8] and takes into account the gain and radiation pattern from the antennas of both tag and Readers, the RF power transmission of a tag, and the radio's Readers threshold sensitivity. Once the digital map (2D layout) and the physical parameters have been defined, the radiomap is built: from a collection of points (i.e. (x,y) locations chosen from a grid of points), the transmission of a beacon is simulated at each point, then depending on the Euclidean distance between the point (x,y) and the Readers, the RSS measured at the radio's Readers is calculated. Therefore, for each point of the radiomap there is a RSS vector, where the i_{th} element is related to the i_{th} Reader. The radiomap is the source data for training and learning the positioning engine.

3 Metrics for positioning performance characterization

For performance evaluation of the location service we define three figures of merit: Mean Absolute Error, Root Mean Square Error, and Precision.
Mean Absolute Error (MAE) is defined as the mean Eucliden Distance between the true (x,y) position and the estimated (x′,y′) position taking into account a collection of points test from the radiomap. The MAE is calculated once the MLP has been trained.
Root Mean Square Error (RMSE) shows what the variance of the error is like. For example, if a high RMSE is achieved but with a low MAE, it is because there are areas where the positioning engine has a poor precision but there are other areas with good precision.
Precision (or Accuracy) is related to the cumulative distribution function of the random variable error. If the positioning engine gets a precision of E meters with P probability, it means that any estimated (x′,y′) position is within a disc of center (x′,y′) and a radius equal E with a probability P.
It is important to remark that the selection of the collection of test points is an important issue in order to evaluate the performance of the positioning engine properly.

4 Positioning estimation engine

Neural networks are self-learning techniques which, starting from an environment radiomap, are effective for the localization of problems, since they act as universal interpolators. One of their main characteristics is that no prior knowledge about environment geometry (position of rooms, walls and obstacles) is needed. Knowledge about propagation channel and Reader positions is not necessary either. In our system, different Readers use the RSS measurements from a given tag to determine its location within the work area.

4.1 Multilayer Perceptron

The MLP is a kind of ANN very useful to solve function approximation problems. This network can achieve a good interpolation and is able to extrapolate values of the function for function inputs never shown before. This property is called generalization ability and it is what makes these mathematical tools very useful.
Therefore, it starts from the previous knowledge of certain points of the function one wants to interpolate. In this case, from the radiomap of the environment, which provides the relationship between the RSS measured by different Readers and the location (x,y) from where the beacons were sent.
The architecture of the MLP is made of input layers, hidden layers and output layer. Each layer is made of units or processing neurons, where the outputs from the previous layer are multiplied by its respective weights *wij* and added and then fed to a transfer function. In short, it is a network of internal weights *w* which, given the fact that they are appropriately chosen, is able to approximate the function that relates the inputs of the input layer with the outputs of the output layer. The weights are computed during the training stage (offline phase). In that stage, a set of samples whose outputs are known are fed to the network so that the network can learn the relaitonships. By the iterative algorithm, the neuronal weights of each layer are evolving progessively and minimizing an error function, which is usually the MSE (Mean Square Error).

$$MSE = \sum_{n=1}^{N} (t_n - o_n(\omega))^2 \qquad (3)$$

As can be seen, this function depends on both the difference between the outputs estimated by the network at a given instant, which is a function of

the weights *w* at that moment, and the desired outputs (also called targets), associated to the input.
After the training, the network will be able to estimate the position (x,y), with a given error, for any input vector (RSS1, RSS2, RSS3,...,RSSn), never previously seen in the training stage. The resulting error is called test error or generalization error.
In our case, for the training and testing of the trained network, the Matlab Neural Network Toolbox [11] has been used.

5 Assumptions and Working Methology

In this work, we have evaluated how to integrate and process the information of an Active RFID system to estimate the location of tags in a 2D area, by simulation. In the following sections the methodology and work steps are described.

5.1 Environment Characterization and 2D Layout

First, the environment is modeled with the appropriate accuracy: description of floors, rooms, walls, windows, doors and so on. The accuracy depends on the propagation model used. In our case, it is enough to define the walls that separate each room, since our model does not take into account materials, but the number of walls traversed by the signal.

5.2 Hardware parameters and Readers placement

Second, the transmission power of the tag and its antenna gain are configured. In addition, the sensitivity and gain of the Reader is configured. In our simulations we have fixed a transmission power of 10 dBm, and a sensitivity of -110dBm. We assume that the antennas are isotropic. The gains of Reader and tag are fixed to 5dBi and 0dBi respectively.
Finally, the Readers are positioned on the map. Nine Readers were positioned, avoiding shadow areas and selecting positions that provide good results as shown in [12] and in Figure 2. It must be remarked that it is assumed that good coverage is available in all the area, otherwise it would affect negatively to the system, which is out of the scope of the present work.

5.3 Selecting Channel Propagation Model

Several propagation models are available to be used: path-loss, free-space, path-loss doble slope, partition loss, etc. We have selected a path loss model with the parameters estimated in [9] for our environment.

5.4 Space Discretization

The space discretization of the area can be made in two main ways: either by selecting points directly on the map or by setting a grid whose resolution on X and Y can be tuned. In our case, we put the origin of the coordinate system in the left upper corner and then set up a grid of 0.8m x 0.8m. The set of samples is obtained from that grid. All the points have a prefixed height that simulates the average height of a tag carried by an iser. It has been fixed to 1.5m.

5.5 Simulation and getting sample datasets

Once the physical layer has been modeled, the radiomap is generated in order to obtain the pairs of postions (x,y) and RSS vectors measured by the deployed Readers. The network is trained and validated with the dataset generated in the next stage. With the 0.8m-resolution grid we get a quasi-uniform set of 231 training samples and 88 validating samples. Since generating training samples from simulation is direct, we have generated another set of 230 samples selected directly from the modeled environment for the system testing stage.

5.6 Selecting and training the MLP

The next step is to select an architecture for the neural network that will be trained to approximate the function that relates RSS vectors and position (x,y), from where the signal is emitted. After several trials and errors we conclude that a 16-4-2 architecture is powerful enough and provides good results. The proposed architecture is made of two hidden layers with hyperbolic tangent transfer functions and 16 and 4 neurons respectively. The output layer is made of 2 neurons (estimated X and Y) , whose transfer function is the identity so as not to limit the output range. In the input layer we have as many units as Readers we are using to estimate the position. The MLP training uses two datasets, one for training and the other for validation, with 231 and 88 samples respectively, which are generated from the simulations and stored in the

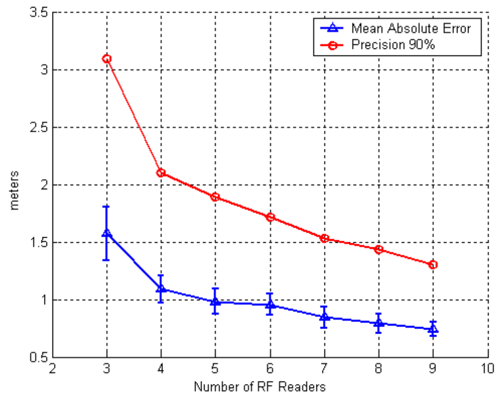

Fig. 2 MAE and Precision (probability 90%) of the Active RFID Location System vs Number of Readers (231 test samples)

radiomap. The training algorithm
selects the weights that minimize the risk function, which is defined as the MSE of the output (X',Y') with respect to the real position (X,Y).
To get a minimal generalization error we have used the early stopping technique. This algorithm uses the validation samples (usually a 20% of the total number of samples) and stops the training stage if the risk function for the complete set of samples is not reduced during a given number of cycles. On the contrary, if it keeps reducing itself, the training finishes when a given number of cycles has been reached, 400 in our case. This technique is used to avoid the overfitting effect, which reduces the generalization ability of the network. The training algorithm used has been the Levenberg Marquardt which provides very good results [13]. It is based on the backpropagation technique. Finally, the samples have been normalized and denormalized in a range [-1,1] with a pre and post processing since it has been shown that it improves the generalization ability of the network after the training.

5.7 Testing and getting results

The test set is generated to evaluate a space point where the system should obtain a good result (apart from the points selected for training) like certain points in the aisle, corners and central parts of the room. This selection can be made given some spacial and quality of service constrains (accuracy), and the operation characteristics that are to be obtained with the system in certain zones.
In our study, we simulate a total of 221 test samples located at points where the system should achieve good estimates.

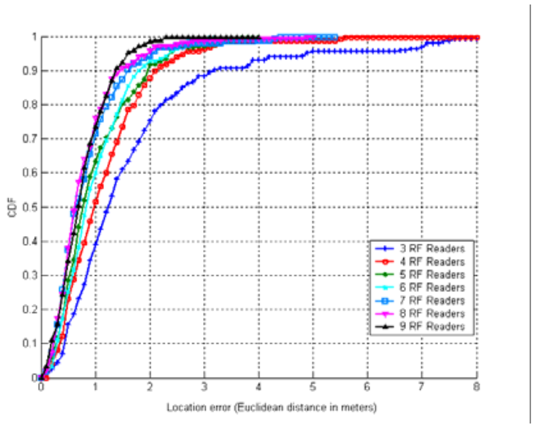

Fig. 3 Cumulative Distribution Function of the VA Error vs Number of Readers

5.8 Analysis and error trends of results

The results and trends in the error statistics and their correlation with the network architecture as well as the main configuration parameters are analyzed.
With this procedure, new hypothesis are drawn about the effects on the generalization error provided by a given architecture (number of hidden layers and number of neurons in each layer) and the combination of parameters. Afterwards, the previous steps are repeated until an optimal configuration of the networks is achieved.

5.9 Validation of results

Once it has been decided that a given architecture and configuration parameters provide good results, the process is repeated again, now changing the testing samples in order to validate the results.

6 Results and discussion

In this section we evaluate the influence of the number of Readers deployed in the environment shown in Section 2.2 over the estimation error yielded by the location system based on MLP. For this study, the simulator tool developed will be used, following the working methology proposed. This way, the power of the simulator tool to make this kind of parametric studies and to evaluate indoor location systems is demonstrated.

6.1 Number of Readers

We first show the trend of the location system error as the number of Readers deployed in the enviroment is increased. Fig.1 shows nine Readers distributed through the building. The tendency of the location system error is presented in Fig.2 and Fig.3. In Fig.2, MAE and Precision (90%) are presented. These metrics have been calculated from 248 test samples (samples different from training set). In this graph, it is shown that the greater the number of Readers is, the more accurate is the system (less error is committed). In fact, simulation results predict that, with seven Readers (RF1 to RF7 in Fig.1), we get 1.5 m precision for 90% of the test set (248 samples). For more information, MAE and CI95% (Confidence Interval to 95%) are represented. This measure indicates the average error which is generated by the system at test locations. The CI95% indicates that the real system average error is between those two boundaries with 95% probability. Obviously, MAE and CI95% follow the same decrease tendency with respect to the number of Readers, as precision describes. In addition, the confidence interval is tighter as the number of Readers increases. This effect is because the greater the number of Readers that measure the signal, the greater the information that is extracted about the position of the user's transmitter. However, this information might become redundant if the number of Readers is very large and the error reduction between both situations negligible. We can observe those effects through the precision or MAE tendency between 5 and 9 Readers. The improvement between one situation and another is only a few centimeters. Moreover, the decreasing CI95% means that the error committed by the system is more homogeneous in the scenary area. This is because CI is directly proportional to RMSE. We can verify this result visually in Fig.5, where we represent an error map of the area, presenting the error, for 7 Readers case, as the Euclidean distance between the real location and estimated location in the discretized points of a mesh distributed homogenously in the whole area. This grid has a resolution of 0.25 x 0.25 m.

Fig. 3 verifies the latest conclusions since the CDF (Cummulative Distribute Function) approach. CDF represents the system behaviour, in precision, with respect to the error. As the number of Readers increases, the CDF of the Error aproaches 1 more quickly. If we trace a vertical line from a 2m error which intersects with the different CDFs (each case of number of Readers), we can observe that with three Readers, the system gets a maximum error of 2m in 75% of the test samples. For 4 Readers, the CDF increases quickly, getting the 90% of test location equal or below 2m. In the last case of 9 Readers deployed, the system is able to estimate the 98% of the test locations with an error not exceeding 2m.

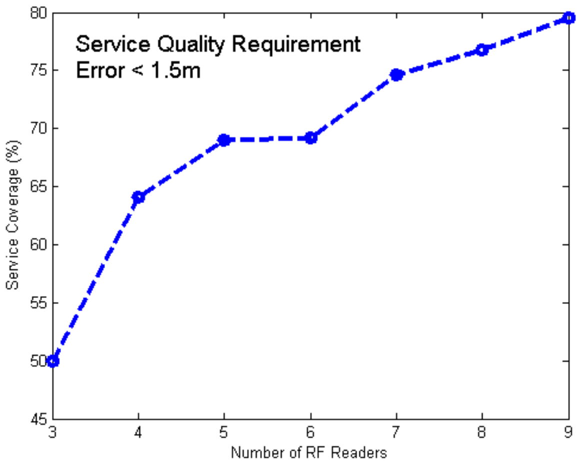

Fig. 4 Percentage of Area Coverage with Error ≤ 1.5m

In Fig. 5 we observe certain zones (corners, corridors or some room locations), where the error is bigger than others. These errors are due to a bad error generalization at these locations, since the vectors stored in the Radio Map from the signal space at discrete locations of the building, and from which to train the MLP network (training set), are not sufficiently representative of the relation between signal space and position at some areas. For this reason, the MLP will not be able to approximate quite well the function between position and RSS vector at these regions.

Finally, the study is approached from the point of view of a global service coverage provided by the system. The service coverage of the location system is defined as the area where the system accomplishes a required service quality, which in our case is the maximum location error allowed. For our simulation, the quality requirement imposed is a maximum error of 1.5m, since such error is permissible for this kind of services. In Fig.5 the service coverage is represented as the number of Readers that are considered. An increasing tendency of service area is observed as the number of Readers is augmented. Again, these results verify the previous conclusions.

7 Conclusion

Indoor location systems have attracted great commercial interest for context aware services. In this paper, an indoor location system based on active RFID using neural networks is used to estimate users' location. The goal is to study, in a fast, powerful and comfortable way, the effects generated by different parameters (such as hardware parameters, number and emplacement of Readers, or antenna pattern diagrams) of these systems on the system error. To this aim a full simulator tool has been developed with which we are able to model the indoor environment, channel propagation effects and location engine.

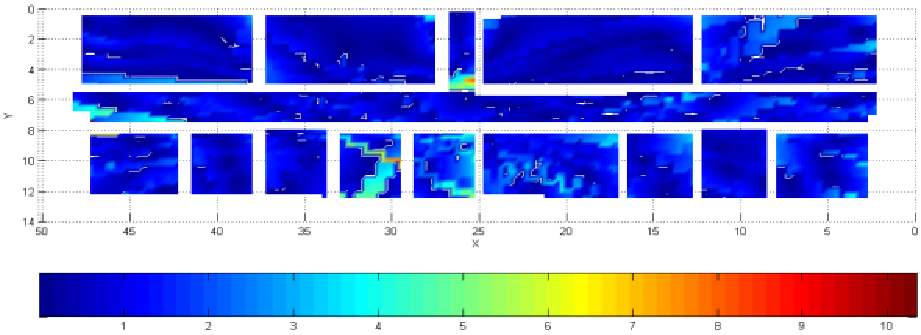

Fig. 5 Map error (blue equal low error, red high error) for a system with 7 Readers

In order to demonstrate the capabilities of our tool, a study about the number of Readers deployed in a complex indoor environment has been carried out. The error system trend was obtained by the simulator analysis, which led us to the conclusion that as the number of Readers increases, the system is able to get a maximum error of 1.5m in 75% of the total area with 8 Readers located through the building. Therefore, global error diminished as the precision increased in the whole service area.

In the future, our work might be extended in other directions. One of the most interesting studies would be the prediction of Readers location to improve the system accuracy, extending the service coverage under restrictions such as indoor environment obstacles, maximum transmitter power or hardware constraints.

Acknowledgements

This work has been funded by the Spanish National Project TEC2007-67966-01/TCM (CON-PARTE-1) and by the Fundación Séneca throught the “Becas de Innovación" and “Programa de Ayudas a Grupos de Excelencia de la Región de Murcia”.

References

[1] K. W. Kolodziej, J. Hjelm, *Local Positioning Systems: LBS Applications and Services, ed.* CRC, 2006

[2] P. Bahl, V.N. Padmanabhan, *RADAR: An in-building RF-based user location tracking system*, IEEE INFOCOM, pages 775-784, 2000

[3] L. M. Ni,Y. Liu,Y. C. Lau, and A. P. Patil, *LANDMARC: Indoor location sensing using active RFID*, Wireless Networks, vol. 10, no. 6, pp. 701–710, Nov. 2004.

[4] EKAHAU, http://www.ekahau.com

[5] Aeroscout, http://www.aeroscout.com

[6] R. Battiti and A. Villani and T. Le Nhat, *Neural network models for intelligent networks: deriving the location from signal patterns*, in Proceedings of Annual Symposium on Autonomous Intelligent Networks and Systems, 2002

[7] M. Borenović, A. Neskovic, D. Budimir, and L. Zezelj, Lara, *Utilizing artificial neural networks for WLAN positioning.* In 19th International IEEE Symposium on Personal, Indoor and Mobile Radio Communications, 2008, 15-18 September 2008, Cannes, France.

[8] C.Y. Tsai, S.Y. Chou, S.W. Lin and W.H. Wang, *Location determination of mobile devices for an indoor WLAN application using a neural network,* Knowledge and Information Systems, pp. 81-93, vol. 20, nº 1, 2009.

[9] Jose-Maria Molina-Garcia-Pardo, José Víctor Rodríguez and Leandro Juan Llácer, *Polarized Indoor MIMO Channel Measurements at 2.45 GHz*, IEEE Transactions on Antennas and Propagation, vol. 56, pp. 3818-3919, 2008

[10] T. Rappaport, *Wireless Communications: Principles and Practice,*ed. Prentice Hall, 2001

[11] MATLAB Neural Tool Box; http://www.mathworks.com/

[12] Baala, O. Et al: “The Impact of AP placement in WLAN-based Indoor Positioning System”, IEEE , 8th International Cibference on Networks, 2009

[13] M.Hagan, M.Menhaj, *Training Feedforward networks with the Marquardt algorithm*, IEEE T. Neural Networks 5 pag 989-993, 1994

Vitae

Alejandro S. Martinez-Sala received the Electrical Engineering degree in 2000 and the Ph.D. in Telecommunications from the Polytechnic University of Cartagena (UPCT), Spain, in 2006. In October 2001 he joined the UPCT where he is assistant professor of the Department of Information Technologies and Communications. He has been involved in several National research projects related to RFID, location systems, and wireless sensor networks.

Raúl Guzman Quirós received the Telecommunications Engineering degree in 2009, from the Polytechnic University of Cartagena (UPCT). He his a PhD student and since 2010 has a research grant from Fundacion Séneca. His research interest is focused on wireless location systems and intelligent reconfigurable antennas.

Esteban Egea Lopez received the Telecommunications Engineering degree in 2000, from the Polytechnic University of Valencia (UPV), Spain, the Master Degree in Electronics in 2001, from the University of Gavle, Sweden, and Ph.D. in Telecommunications in 2006 from the Polytechnic University of Cartagena. Since 2001, he is an assistant professor of the Department of Information Technologies and Communications at the Polytechnic University of Cartagena. His research interest is focused on ad-hoc and wireless sensor networks