NLOS Classification Based on RSS and Ranging Statistics Obtained from Low-Cost UWB Devices

Valentín Barral, Carlos J. Escudero, José A. García-Naya
University of A Coruña, Department of Computer Engineering, A Coruña, Spain.
E-mail: {valentin.barral, escudero, jagarcia}@udc.es

***Abstract*—Ultra-wideband (UWB) devices have been largely considered for indoor location systems due to their high accuracy. However, as in other wireless systems, such accuracy is significantly degraded under non-line-of-sight (NLOS) propagation conditions. Therefore, the identification of NLOS conditions is essential to mitigate inaccuracies due to NLOS propagation. Nonetheless, most of the techniques considered to identify NLOS situations are based on the study of the channel impulse response (CIR), which is not practical and even becomes unfeasible when employing low-cost UWB hardware. This is precisely the main motivation of this work, to introduce a classification system based on the statistics of both the received signal strength (RSS) and range available from low-cost UWB devices.**

We analyze the effect of considering different statistic sets of both the RSS and range as features to feed a support vector machine (SVM) classifier, which is experimentally evaluated by means of measurements carried out in a real scenario where both line-of-sight (LOS) and NLOS conditions are present.

***Index Terms*—Ultra-wideband, NLOS Classification, RSS, ranging, SVM.**

I. Introduction

Ultra-wideband (UWB) location systems have become increasingly popular due to their capabilities to provide accurate positioning based on the signal time of arrival (TOA) or time difference of arrival (TDOA). When there is a clear line-of-sight (LOS) between emitter and receiver devices, the performance of UWB location systems is excellent. Unfortunately, multipath effects are frequently found in indoor environments, particularly in non-line-of-sight (NLOS) situations. Although UWB systems are more robust against multipath than those relying on received signal strength (RSS) (e.g., those based on WiFi, Bluetooth or radio-frequency identification (RFID)), in the absence of a clear LOS, the receiver might select a secondary delayed path instead of the primary one, leading to an erroneous estimation.

Therefore, NLOS detection has revealed as one of the most important tasks in order to mitigate the error caused by multipath propagation and hence to improve the accuracy of UWB location systems. NLOS detection can be addressed considering different techniques, being statistical analysis applied to range estimations one of them [1], [2]. Alternatively, several detectors based on the estimated channel impulse response (CIR) have been proposed in the literature [3]–[6]. All of them rely on the idea that, under LOS propagation conditions, the energy of the first path is significantly larger than that of the secondary paths, but this behaviour is not found in NLOS situations. NLOS detection can be performed at a higher level, once the ranging estimations have been integrated into a location algorithm. In this case, additional information such as the geometry of the scenario or historical data (in the case of tracking algorithms) is also used [7].

In this work, we consider machine learning techniques to detect NLOS conditions using measurements carried out with low-cost UWB hardware in an indoor scenario. More specifically, we propose a support vector machine (SVM) classifier based on statistics of both the RSS and range parameters provided by UWB devices. Considering a real indoor scenario, the classifier is experimentally evaluated when identifying LOS and NLOS situations. Additionally, the measurement data are publicly available [8].

The paper is structured as follows: Details of the hardware used in the experiments and a description of the data supplied by them are provided in Section II. Section III describes the measurement scenario and the methodology used to record the measurements. Details of the SVM classifier are found in Section IV, including the combination of the parameter statistics and the methodology employed to process the measurements. Section V shows the results obtained and, finally, Section VI is devoted to the conclusions.

II. Hardware

This paper focuses on Pozyx hardware [9], a low-cost hardware based on modules integrating a UWB transceiver and several inertial sensors. They can be used as an Arduino board or connected to a computer via USB. The low cost of these devices has converted them into a very popular solution to perform indoor localization. They rely on a two-way (round trip) TOA-based algorithm to estimate the distance (range) between two nodes: a tag and an anchor. First, a poll signal is sent from the tag towards the anchor. Next, the anchor returns immediately a response signal to the tag. Finally, the so-called final message is sent from the tag to the anchor to compute the time of flight, from which the distance between the nodes is estimated. In a location system, multiple anchors (a minimum of three for a 2D location, or four for 3D) are considered, being possible to estimate the position of a given tag from the ranges between the tag and each of the anchors.

The Decawave UWB transceiver [10] inside the Pozyx hardware provides the RSS and range estimations from the incoming signal corresponding to the first path whose energy exceeds a predefined threshold value. Given that both the RSS and range parameters are affected by noise and bias [11], the final estimate of the tag position will not be exact.

An additional hindrance arising in NLOS situations is multipath, which might originate a situation in which the energy of the signal corresponding to the first path does not exceed the aforementioned threshold, whereas the energy of a signal from a secondary delayed path does surpass such a threshold, yielding an overestimation of the distance between the nodes. Although Pozyx devices are able to output a set of samples of the CIR when a new signal arrives, range estimations based on the provided CIR present the following constraints:

- To obtain the CIR information it is necessary to gather a significant amount of data through the serial port (4064 bytes [12]) , hence originating a long latency of about one second just to transfer the data. This leads to a long delay for estimating the range, making this approach unfeasible when the estimated position needs to be updated at a high frequency.
- Pozyx devices can be used in two different modes: local and remote. In the local mode, a tag starts a two-way ranging protocol with an anchor, obtaining the range, the RSS and possibly the CIR. However, in the remote mode, a monitoring node commands a tag the beginning of a ranging process with an anchor. The tag will return the obtained ranging and RSS values to the monitoring node, but the available CIR data corresponds to the wireless link between the tag and the monitoring node instead of the CIR of interest, which is the one from the tag and the anchor.

Therefore, under the above-mentioned constraints, relying on CIR estimations is not a feasible approach in real-time applications when dealing with low-cost UWB solutions, such as the one from Pozyx. On the other hand, both the RSS and range are readily available even when the nodes operate in the remote mode. Additionally, because of the small number of bytes required by each range, up to 20 measurements per second can be acquired, thus making feasible to obtain statistics from both the RSS and range when considering real-time applications.

III. Experimental Setup

We have carried out a measurement campaign indoors, in the corridors of the so-called Scientific Area Building, placed at the Elviña Campus at the University of A Coruña, Spain. Five Pozyx devices were distributed as shown in Fig. 1. A monitoring device (M) was attached to a computer to record the range measurements between the tag (T) and the other three anchors (namely A, B and C) working in the remote mode. This way, the tag can be moved down the aisle without requiring a computer attached to it to store the measurements. The anchors A, B, and C remain static in their positions during the whole measurement campaign, whereas the measurements between the tag and each of the anchors were recorded at different tag positions spaced 0.5 m apart along the aisle, storing measurements for each position during 90 seconds at 20 measurements per second, which is a period of 50 ms.

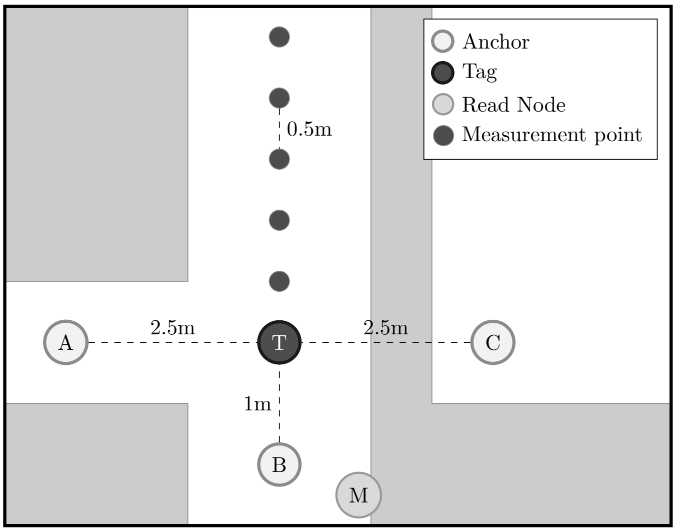

Fig. 1. Experimental setup.

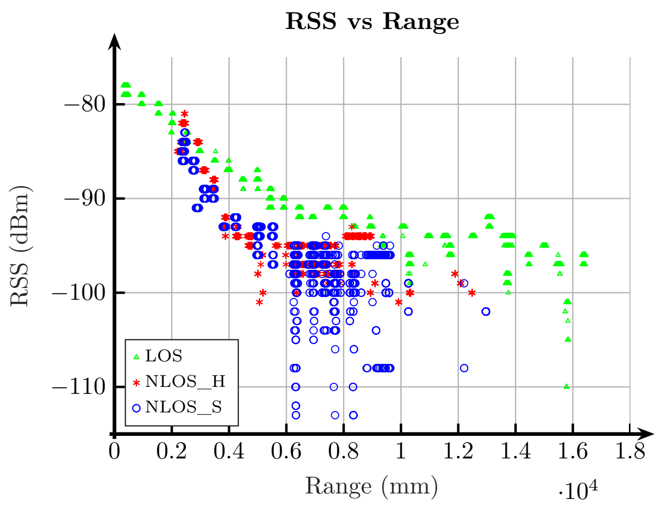

Fig. 2. Parameter measurements.

At the host side, an Arduino Uno device was attached to the monitoring device (M) to read the measurements through the serial port and to store them in a computer via USB. In such a computer, a Robot Operating System (ROS) [13] instance was in charge of reading, parsing and storing both the range and RSS values together with the timestamp and the actual distance value between the tag and the anchors. Once the measurements were recorded, they were exported to be processed by algorithms developed in Matlab. Such data files are publicly available [8].

A. NLOS Model

Fig. 2 shows the measurements of the RSS and range parameters obtained from the scenario shown in Fig. 1. From Fig. 2, it is not obvious how to distinguish between LOS and NLOS conditions, especially if we consider two different

NLOS categories previously proposed in the literature [2], [14]:

- Soft NLOS (NLOS-S) is a condition found when both the first and secondary paths are attenuated by the same obstacle (e.g., a wall, as in our scenario). In this case, the receiver can detect correctly the first path if the energy level is sufficiently high for the signal to be decoded.
- Hard NLOS (NLOS-H) appears when the emitter and the receiver are placed at positions without clear LOS between them, thus all signals impinging the receiver correspond to reflections. Therefore, the secondary paths can reach the receiver with more o similar level of energy than the first path (which is blocked in most of the times).

This classification of NLOS is very important because the mitigation of its effects will be done differently. However, the problem observed in Fig. 2 shows that this identification is not easy. On the other hand, the LOS condition, where both the emitter and receiver have an open space between them such that the first path can be received without obstacles, is much easier to identify.

In the scenario of our indoor measurement campaign, the three considered anchors have been strategically placed to obtain a representation of each of the three possible propagation conditions with respect to the tag: Node A experiences NLOS-H, Node B corresponds to LOS and Node C models NLOS-S.

IV. Machine Learning

Machine learning techniques are well known in classification problems [15]. In particular, we consider the classic SVM algorithm [16], a supervised algorithm which tries to find the hyperplane that maximizes the distance between the values of its input features. The algorithm performance is based on a set of parameters (kernel type, regularization, and sigma parameters mainly) that can be adjusted after a process of cross-validation, selecting the one that causes the lowest misclassification rate.

The SVM algorithm typically follows the steps below:

1) A set from the original measurements is selected as the training data, whereas the remaining data is used as the test set. Notice that different measurements at the same position can be included in both sets. Although this condition might be seem unrealistic, it does not impact on the parameters comparative.
2) The training set is used to train the SVM model using a cross-validation schema. This means that, in an iterative process, the training set is split into several subsets in order to compare different realizations of the SVM model depending on some configurations of the algorithm. The most successful configuration is the one selected to be used in the test stage.
3) Once the training stage has finished, the SVM model is applied over the test set to get an estimation of its real performance.

In our case, all the data obtained from the measurement campaign will be shuffled in order to obtain the training and the test sets with a mixture of NLOS-H, NLOS-S and LOS conditions.

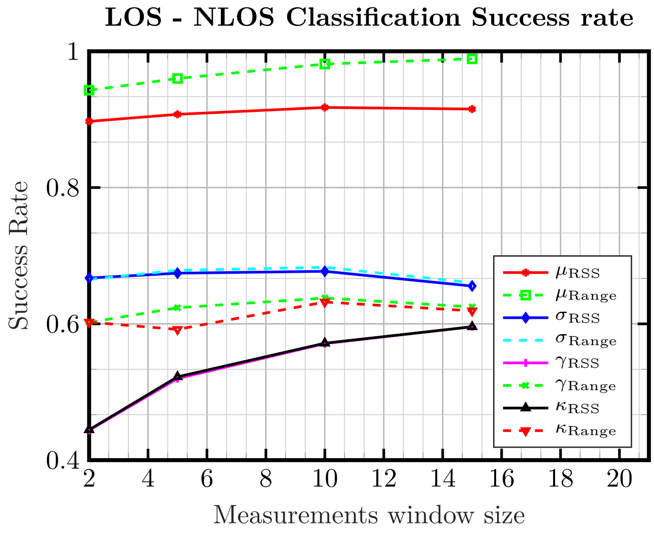

Fig. 3. Success rate of a LOS-NLOS classifier with individual statistics (features).

In our work, we consider the following statistics, which will be the input features for the SVM algorithm:

- μ_{rss}, μ_{ran}: Mean RSS and mean range, respectively.
- $\sigma_{rss}, \sigma_{ran}$: Standard deviation of the RSS and range, respectively.
- $\gamma_{rss}, \gamma_{ran}$: Skewness of the RSS and range, respectively.
- κ_{rss}, κ_{ran}: Kurtosis of the RSS and range, respectively.

Note that, as mentioned in Section II, Pozyx devices produce noisy values. Therefore, to generate the above-mentioned statistics, it is necessary to consider a window size to group the raw measurement values, similarly to what has been already done in other works such as [17] for the RSS in a WiFi system. Consequently, the window size will be an additional system configuration parameter impacting on the latency of the classification process.

V. Results

In order to assess the classifier performance when separating LOS and NLOS conditions, Fig. 3 shows the success rate for the different statistics (features) mentioned in Section IV when they are considered individually. Fig. 3 reveals that the separation between LOS and NLOS is performed very well since, according to Fig. 2, these two conditions are easier to identify. In particular, the statistics corresponding to the RSS and range means and variances, μ_{rss}, μ_{ran}, σ_{rss} and σ_{ran}, respectively, are the ones providing the best results.

However, it also follows from Fig. 2 that separating NLOS-H from NLOS-S is much more challenging, as revealed in Fig. 4. As expected, this classification is more difficult to perform because both NLOS conditions are similar. However, our classifier behaves well again for the same statistics that worked also well in the results shown in Fig. 3.

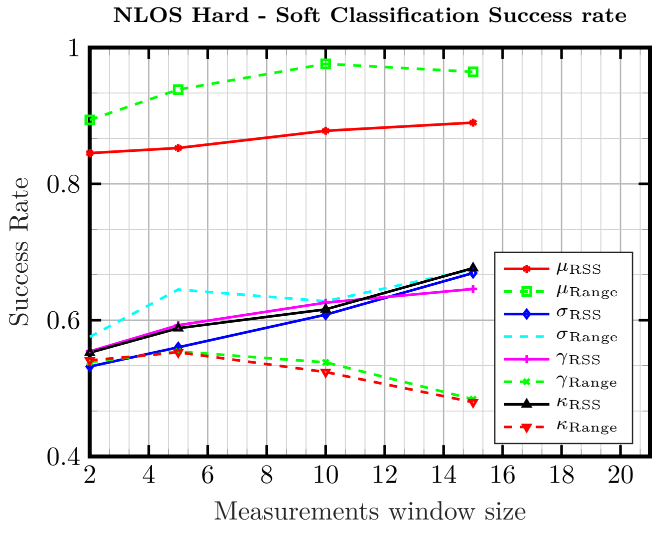

Fig. 4. Success rate of a NLOS-H and NLOS-S classifier with individual statistics.

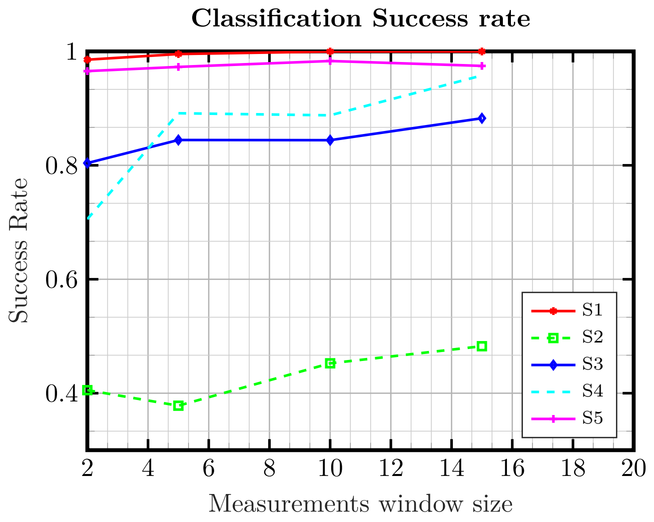

Fig. 5. Success rate a joint classifier of LOS, NLOS-H and NLOS-S when the sets of statistics specified in Table I are considered.

TABLE I
FEATURES SETS

Feature	Set 1	Set 2	Set 3	Set 4	Set 5
μ_{rss}	✓		✓		✓
μ_{ran}	✓			✓	✓
σ_{rss}		✓	✓		✓
σ_{ran}		✓		✓	✓

Finally, considering that μ_{rss}, μ_{ran}, σ_{rss} and σ_{ran} perform the best according to the results shown in Figs. 3 and 4, we now consider combinations of them (see Table I) as the input features of the classifier. We discard the skewness and the kurtosis statistics for the RSS and ranging because they do not provide an improvement in the performance of the classifier. In this experiment, the performance of the classifier is evaluated when considering the joint identification of LOS, NLOS-H and NLOS-S. Fig. 5 shows that, even in this more complex situation where the three propagation conditions have to be jointly identified, our classifier does a good job. In this scenario, the sets $S1$ and $S5$ are the most suitable, providing the highest robustness levels for different propagation situations. In fact, in this experiment, we have used a restricted group of combinations, but we could complete Table I considering other combinations which, in other scenarios, could behave better.

According to Figs. 3 to 5, the performance of the classifier improves with the increase of the window size in most of the cases. However, the larger the window size, the longer the latency introduced in the estimation process. Although it is necessary to consider the observed trade-off between the classifier performance and the introduced latency due to the window size, when considering the sets $S1$ and $S5$ (see Fig. 5), the classifier performance reaches values close to its maximum for windows with a size equal to 5, which can be considered a low window size.

VI. CONCLUSIONS

In this work, we analyzed the viability of using machine learning techniques to identify NLOS conditions in an actual indoor scenario when low-cost UWB devices are employed and CIR estimations are not feasible. In particular, special interest is shown in classifying two different NLOS categories as this will be of utmost importance for mitigating their effects. To this end, a measurement campaign has been carried out in an indoor scenario and three possible propagation conditions were considered: LOS, hard NLOS and soft NLOS. In order to classify these conditions, only the RSS and the range have been considered, since they are the only parameters available from low-cost UWB hardware. Different statistics were calculated from the measured parameters and they were used as the inputs (features) of a classifier based on the well known SVM algorithm.

The results prove that the joint classification among LOS, hard NLOS and soft NLOS is performed at a very high success rate (close to 1) when the mean of both the RSS and range are considered.

ACKNOWLEDGMENT

This work has been funded by the Xunta de Galicia (ED431C 2016-045, ED431G/01), the Agencia Estatal de Investigación of Spain (TEC2016-75067-C4-1-R) and ERDF funds of the EU (AEI/FEDER, UE).

REFERENCES

[1] J. Borras, P. Hatrack, and N. B. Mandayam, “Decision theoretic framework for NLOS identification,” in *Proc. of the 48th IEEE Vehicular Technology Conference (VTC)*, vol. 2, 1998, pp. 1583–1587.

[2] S. Venkatesh and R. M. Buehrer, "Non-line-of-sight identification in ultra-wideband systems based on received signal statistics," *IET Microwaves, Antennas and Propagation*, vol. 1, no. 6, pp. 1120–1130, Dec. 2007.

[3] A. Maali, H. Mimoun, G. Baudoin, and A. Ouldali, "A new low complexity NLOS identification approach based on UWB energy detection," in *Proc. of the IEEE Radio and Wireless Symposium*, Jan. 2009, pp. 675–678.

[4] S. Al-Jazzar and J. Caffery Jr, "New algorithms for NLOS identification," in *Proc. of the 14th IST Mobile and Wireless Communications Summit*, 2005.

[5] İ. Güvenç, C.-C. Chong, F. Watanabe, and H. Inamura, "NLOS identification and weighted least-squares localization for UWB systems using multipath channel statistics," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, no. 1, p. 271984, Aug. 2007. [Online]. Available: https://doi.org/10.1155/2008/271984

[6] S. Marano, W. M. Gifford, H. Wymeersch, and M. Z. Win, "NLOS identification and mitigation for localization based on UWB experimental data," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 7, pp. 1026–1035, Sep. 2010.

[7] C.-D. Wann and C.-S. Hsueh, "NLOS mitigation with biased kalman filters for range estimation in UWB systems," in *Proc. of IEEE Region 10 Conference (TENCON)*, 12 2007, pp. 1 – 4.

[8] V. Barral, C. J. Escudero, and J. A. García-Naya, "NLOS classification based on RSS and ranging statistics obtained from low-cost UWB devices," 2019. [Online]. Available: http://dx.doi.org/10.21227/swz9-y281

[9] "Pozyx modules." [Online]. Available: http://www.pozyx.com/.

[10] "Decawave. scensor dw1000." [Online]. Available: http://www.decawave.com/.

[11] V. Barral, P. Suárez-Casal, C. Escudero, and J. A. García-Naya, "Assessment of UWB ranging bias in multipath environments," in *Proc. of the International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Alcalá de Henares, Spain, 2016.

[12] "Pozyx register overview." [Online]. Available: https://www.pozyx.io/product-info/developer-tag/datasheet-register-overview

[13] "Robot operating system." [Online]. Available: http://www.ros.org/.

[14] W. Yang, C. Peipei, Z. Xinwei, Z. Qinyu, and Z. Naitong, "Characterization of indoor ultra-wide band NLOS channel," in *Proc. of the IEEE Annual Wireless and Microwave Technology Conference*. IEEE, 2006, pp. 1–5.

[15] S. B. Kotsiantis, I. Zaharakis, and P. Pintelas, "Supervised machine learning: A review of classification techniques," *Emerging artificial intelligence applications in computer engineering*, vol. 160, pp. 3–24, 2007.

[16] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proc. of the fifth annual workshop on Computational learning theory*. ACM, 1992, pp. 144–152.

[17] Z. Xiao, H. Wen, A. Markham, N. Trigoni, P. Blunsom, and J. Frolik, "Non-line-of-sight identification and mitigation using received signal strength," *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1689–1702, 2015.