2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN), 24-27 September 2018, Nantes, France

WLAN Indoor Ranging Dataset for Evaluation of Time of Arrival Estimation Algorithms

1st Ahmed Makki
School of Computer Science
University College Dublin
Belfield, Dublin 4, Ireland
ahamza44@gmail.com

2nd Chris J. Bleakley
School of Computer Science
University College Dublin
Belfield, Dublin 4, Ireland
chris.bleakley@ucd.ie

***Abstract*—This paper describes a dataset of WLAN (Wireless Local Area Network, WiFi, IEEE 802.11g/n) ranging signals recorded in an indoor environment for the purposes of time of arrival (TOA) estimation. The dataset contains signals captured in four experimental configurations, with two different clocking techniques, at ranges from 0.5 m to 25 m, and using two different symbols - the 802.11 standard long training sequence and an impulsive symbol. It is expected that the dataset will facilitate development and benchmarking of improved algorithms for TOA estimation in WLAN systems.**

***Index Terms*—indoor localization, IEEE 802.11, positioning, WLAN, time of arrival, time difference of arrival.**

I. Introduction

In recent years, the number of mobile computing devices with high rate wireless communication capabilities use has increased dramatically. Companies and research groups are actively exploring the possibility of using WLAN technology to provide high resolution indoor position estimation on mobile devices. However, accurate estimation of the indoor position of a mobile computing device using WLAN remains a significant technical challenge.

WLAN is particularly attractive for indoor positioning due to the large number of low cost WLAN-enabled devices already deployed. Solutions exploiting existing network infrastructure to determine the position of a mobile user obviate the need for an additional dedicated network for positioning applications. Most previous work on WLAN positioning has focused on use of the Received Signal Strength Indicator (RSSI) for location estimation since RSSI readings are available at the application layer on all standards compliant devices. However, the metric correlates poorly with position within buildings due to multipath, interference, and noise [1]. As a result, WLAN fingerprinting position estimate accuracy is limited to 1.6-5 m [2]. Fingerprinting methods are also onerous in terms of mapping effort due to the large number of RSSI reference measurements must be taken before positioning can be performed.

Time-based WLAN positioning has been proposed as an alternative solution [3]. Time-based positioning requires that the time of arrival of the WLAN radio signal be measured at the receiver. Typically, the TOAs of signals from multiple synchronized access points (APs) with known positions are measured at a mobile WLAN device. The position of the mobile device can then be estimated based on the time differences of arrival (TDOAs) of the signals. While time-based approaches do not require mapping, they do require very high accuracy TOA estimation, i.e. sub-3 ns for sub-meter ranging. This is challenging for two main reasons. Firstly, the time precision required is less than the baseband sampling period. In other words, sub-sample TOA resolution is required in the baseband. Secondly, the receiver signal is subject to significant multipath. This makes it difficult to separate the direct path signal from indirect signals caused by reflections from the walls. Calculating range based on an indirect path TOA significantly over-estimates the ground-truth range giving rise to large positioning errors.

Despite these challenges, a number of methods have been proposed for estimation of the TOA of WLAN RF signals. One of the difficulties developing improved WLAN TOA algorithms has been in obtaining suitable WLAN baseband signals for analysis. The baseband signal is not normally provided by WLAN transceivers. Access to the signals requires firmware changes only available to the manufacturer. Even if the signals are accessible, performing an experimental campaign to record typical indoor signals in a systematic way is time consuming and error prone. In previous work [4], [5], the authors of this paper developed a TOA estimation algorithm for WLAN transceivers. The dataset was recorded using the Wireless Open-access Research Platform (WARP) developed at Rice University. Conveniently, the platform provides full access to the baseband signals. As part of our work, we recorded a dataset of WLAN baseband signals with ground truth ranging information. We are now making that dataset available to the community for the purposes of research. As well as facilitating the development and testing of new TOA estimation algorithms, we hope that the dataset will allow benchmarking of novel and existing TOA estimation methods. The dataset is now available at [6]. To the best of the authors' knowledge, this the first publicly available dataset containing real-world WiFi symbols recorded in an indoor environment together with ground-truth range measurements.

II. Related Work

A. WLAN Ranging

The reader is referred to [7] for a complete survey of time-based 802.11 indoor positioning methods. Herein we focus on

978-1-5386-5635-8/18/$31.00 ©2018 IEEE

physical layer time-based methods since these offer greatest accuracy as they avoid the timing variation inherent in methods applied at higher layers.

Conventional physical layer methods employ cross-correlation of the received signal with a reference transmitted signal locally stored in the receiver. Typically, the short training sequence (STS) or long training sequence (LTS) are used for TOA estimation. LTS was used by Reddy *et al* [8] to estimate the channel impulse response (CIR). In Matlab simulations, the algorithm achieved an absolute error less than 50 ns (approximately 15 m) in 90% of cases. However, no experimental study was conducted.

Cross-correlation, if used alone, limits the TOA estimate to a time resolution equal to the sampling period duration. To improve upon this, researchers have sought to combine it with super-resolution signal processing methods, such as interpolation [9].

In an attempt to accurately estimate TOA, frequency domain super-resolution methods [10] have been applied to the conventional baseband signal. These methods split the received signal into a signal subspace and a noise subspace using eigendecomposition of the correlation matrix. MUSIC [11] and ESPRIT [12] super-resolution methods have been reported to achieve accuracy in the range 1-5 m [13].

Xiong *et al* [14] proposed ToneTrack, a channel stitching to minimize the effect of the limited bandwidth of 802.11 transceivers. The equivalent of a wideband radio was constructed by transmitting symbols sequentially on multiple WiFi channels and stitching the received symbols together before applying a super-resolution MUSIC algorithm. Exploiting one, two, and three channels, the median accuracies reported were 1.9 m, 1.3 m, and 0.9 m, respectively. Vasisht *et al* also employed channel stitching in [15]. Frequency hopping over 35 channels (in the 2.4 GHz band and the 5 GHz band) was exploited. The method achieved median positioning errors of 65 cm and 98 cm in LOS and non-line of sight (NLOS) conditions, respectively.

The method presented in [5] uses CLEAN deconvolution to separate the multipath components of the signal. The method was found to achieve an accuracy of 0.14-3.15 m

B. WLAN Datasets

To date, all previously published WLAN positioning datasets have focused on RSSI measurements. RSSI readings can be obtained from commercially available software packages running on a laptop or smart phone. Positioning is performed by matching the "fingerprint" of observed RSSI readings of nearby WLAN APs with a previously captured reference database of fingerprints and ground-truth position coordinates. Notable WLAN RSSI datasets include: the UJI-IndoorLoc dataset [16]; the IPIN 2016 datasets [17], [18]; and the Tampere-New Imaging Technologies datasets [19].

C. Radio Channel Models

The alternative to recording a dataset of real-world received symbols is to generate synthetic received signals. This can be done by synthesizing the transmitted symbol and convolving it with artificial impulse responses generated according to a radio frequency channel model. A radio frequency channel model uses statistical parameters to describe the radio frequency environment [20], [21]. Based on this model, typical impulse responses can be generated using random parameter values.

While artificial radio channel models are useful for quickly testing large numbers of possible radio frequency environments, it can be difficult to model all signal impairments present in the transmitter, channel, and receiver - especially non-linear effects. Therefore, it is desirable to confirm the results of simulation studies using a real-world dataset such as the one described herein.

III. Experimental Method

A. Experimental Set-up

The wireless experiments were carried out on the ground floor of the 4-story O'Brien Science Center East at University College Dublin. The normal WLAN infrastructure in the building was operational during the experiments.

The Wireless Open Access Research Platform (WARP) [22] was used to obtain experimental measurements. The platform consists of a 802.11g/n transceiver with FGPA baseband implementation. It allows easy access to the physical layer signals. The ADC has a 40 MHz sampling frequency. WARPlab (ver 7.4) mode was used in the experiments. This allows a block of transmit samples to be sent from one node and a block of receive samples to be recorded simultaneously at a remote node. Due to hardware memory restrictions, the length of the block is fixed.

WARP v3 inter-node RF and sampling clocks synchronization can be achieved using a simple clock module called CM-MMCX. It can be optionally used, via a wired connection, to enable sourcing and sinking the RF reference and/or sampling clocks for inter-node synchronization [23]. In these experiments, the transmitter and receiver shared a common sampling clock via a wired connection. Sharing the sampling clock means that the time reference is constant, facilitating experimentation. Note that, sharing the sampling clock between the transmitter and receiver need not be done when using TDOA positioning, as described in [4]. The phase of the received signal with respect to the receiver sampling clock varies as the transmitter-receiver distance changes. The transmitter and receiver could optionally share RF carrier frequency via another wired connection. Not sharing the RF carrier frequency can lead to phase offset in the received signal.

In our experiments, the transmitter WARP board was used to trigger recording of a block at the receiver board via a wired connection. As in [24], the trigger signal was used to provide a common timing reference at the transmitter and receiver.

Coaxial RF cables, directional flat panel antennas (P2415T [25]) with beamwidth of 34°, and RN-SMA-4 omnidirectional antennas [26] were used for range estimation.

B. Symbol Design

Two ranging symbols were used for TOA estimation measurements: the 802.11g standard preamble long training sequence (LTS) and an impulsive symbol.

The LTS symbol is included in all 802.11a/g frames and has good auto-correlation properties as it was designed for communications time synchronization. The LTS symbol was designed for 802.11a/g and so has a bandwidth of 16.25 MHz. In time-based 802.11 positioning research, the LTS is commonly used for ranging [7]. In these experiments, since the ADC has a 40 MHz sampling frequency, LTS symbols were over-sampled by 2 in the time domain to facilitate signal processing (i.e 128 sample duration in time domain).

The impulsive symbol has a bandwidth of 33.75 MHz and is sampled at 40 MHz, as in 802.11n receivers. It was generated by setting the complex amplitude of all active sub-carriers (108 sub-carriers) to 1. This resulted in an impulsive symbol with 128 samples in time domain. The timing of the impulsive peak within the symbol can be adjusted by means of a circular time delay. For example, a circular time delay of 64 samples centers the impulse within the symbol. Our experiments indicate that the position of the peak within the symbol has little effect on TOA estimation accuracy. The impulsive symbol is not supported by the 802.11 standard. If added just after the original LTS, it would incur less than a 0.2% overhead in a 1500-byte payload frame [14].

C. Transmission and Reception

The WARPLab 7.4 reference design hardware was used in these measurements. This can support storage of 32 k samples (819.2 μs duration). Therefore, a transmit block of 30k samples (750 μs duration) was synthesized in Matlab. The block consisted of 150 LTS/impulsive symbols with gaps of 72 samples between consecutive symbols. Measurements were obtained for 300 transmitted blocks (45,000 symbols in total). The delay between symbols within a block is fixed (72 samples) whereas the delay between blocks (i.e between the last symbol in a block and the first symbol in the succeeding block) is random due to variations in the delay of the operating system of the PC used to transmit the data.

The received data was captured at the output of the receiver ADC and stored for off-line processing.

D. Experiments

Four experiments were conducted for the purposes of recording WLAN signals.

In all four experiments, two clocking setups were used - sharing both the RF and sampling clocks between the transmitter and the receiver; and sharing only the sampling clock between the transmitter and receiver.

In the first experiment, cables of various lengths were used as channels. These channels have very low multipath and so allow assessment of the accuracy of a method for the direct path component only. Recordings were made for eighteen cable lengths, from 1.83 m to 19.15 m. A cable of length of 0.57 m was used for calibration. Actual cable lengths are presented in Table I.

TABLE I: Actual Cable lengths.

Test No	Cable Length [m]	File name
1	0.57	Impulsive_SA_001
2	1.83	Impulsive_SA_002
3	3.66	Impulsive_SA_003
4	5.49	Impulsive_SA_004
5	7.32	Impulsive_SA_005
6	9.15	Impulsive_SA_006
7	19.15	Impulsive_SA_007
8	17.32	Impulsive_SA_008
9	15.49	Impulsive_SA_009
10	13.66	Impulsive_SA_010
11	11.83	Impulsive_SA_011
12	10	Impulsive_SA_012
13	15	Impulsive_SA_013
14	5	Impulsive_SA_014
15	6.83	Impulsive_SA_015
16	8.66	Impulsive_SA_016
17	10.49	Impulsive_SA_017
18	12.32	Impulsive_SA_018
19	14.15	Impulsive_SA_019

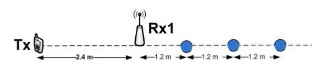

Fig. 1: 1D Experiment Setup

In the second experiment, directional antennas at both nodes were used to create a low multipath environment. Using directional antennas, the transmitter position was fixed and the receiver was moved over nineteen points in a straight line in steps of 1.2 m. The first point was at a transmitter-receiver separation of 2.4 m as illustrated in Figure 1.

In the third experiment, omni-directional antennas at both nodes with LOS were used. The transmitter position was fixed and the receiver was moved over twenty points in a straight line in steps of 1.2 m. The first point was at a transmitter-receiver separation of 2.4 m.

In the fourth experiment, omni-directional antennas at both nodes with NLOS were used. The transmitter-receiver configuration was the same as in the LOS experiments. NLOS conditions were created by putting a metallic plate (185 cm x 60 cm) between the transmitter and the receiver at a distance of 1.8 m from the transmitter.

IV. The Dataset

A. Dataset Organization

The released dataset is saved on a folder called DATASET which contains a four separate folders corresponding to each scenario; in addition, another folder contains the transmitted symbols. In each folder SYNC_ALL contains the measurements obtained when sharing both the RF and sampling clocks while SYNC_SA contains the measurements obtained when sharing sampling clocks only. In each folder there are two sub-folders. One for the LTS symbols and one for the impulsive symbols. This hierarchy is depicted in Figure 2 . Table I provides examples of the file name format and the corresponding cable length. This naming method is followed for all cables scenarios. For the wireless scenarios, the first file

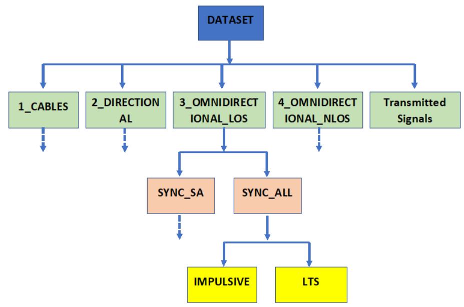

Fig. 2: Dataset Organization

(indexed 001) contains the measured data at the first point (i.e. the receiver is 2.4 m away from the transmitter) and then the index increases by one as the distance increases by 1.2 m. In the NLOS scenario, the measurements at the first point were obtained in a LOS environment for calibration.

Each file is saved as MATLAB formatted data (.mat) and contains one variable called Received which contains a 300x30000 matrix, where every row is a received block.

B. Data Processing

A wired trigger signal was used to trigger the transmitter to transmit a block of samples and to trigger recording of a block at the receiver. The transmitter delays transmission for a fixed amount of time to insure the buffers of the receiver are open and ready to record the data. This operation creates an offset at the receiver measurements. This offset should be measured at the calibration point and should be subtracted at all other points.

Users are recommended to calculate the TOA on the test cable/point and subtract the mean TOA on the calibration cable/point and compare to the range using the speed of light (in case of the cable the speed should multiplied by a factor of (0.7) approximately [27]).

V. Conclusion

This paper describes a public domain dataset for evaluation of time of arrival estimation algorithms for indoor ranging using IEEE 802.11g/n. Experiments were conducted to record the received baseband signals in four environments - using a cable, using directional antennas, using omni-directional antennas with LOS, and using omni-directional antennas with NLOS. Two ranging symbols were used - the standard LTS and a new impulsive symbol.

Acknowledgments

This work was supported by Science Foundation Ireland under Grant Number SFI/11/US/I2220.

References

[1] W. T. Lopes, D. A. Formiga, R. D. Lucena, and W. J. Queiroz, "Localization in IEEE 802.11 networks by using the nelder-mead method," in *ACM Work. Perf. Monitoring and Meas. of Heterogeneous Wireless and Wired Netw.*, pp. 211–216, 2013.

[2] D. Lymberopoulos, J. Liu, X. Yang, R. R. Choudhury, V. Handziski, and S. Sen, "A realistic evaluation and comparison of indoor location technologies: experiences and lessons learned," in *Int. Conf. on Info. Proc. in Sensor Networks*, pp. 178–189, 2015.

[3] A. Günther and C. Hoene, "Measuring round trip times to determine the distance between WLAN nodes," in *Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications Systems*, pp. 768–779, 2005.

[4] A. Makki, A. Siddig, M. Saad, J. R. Cavallaro, and C. J. Bleakley, "Indoor localization using 802.11 time differences of arrival," *IEEE Trans. Instr. and Meas.*, vol. 65, pp. 614–623, March 2016.

[5] A. Makki, A. Siddig, and C. J. Bleakley, "Robust high resolution time of arrival estimation for indoor wlan ranging," *IEEE Trans. Inst. and Meas.*, vol. 66, no. 10, pp. 2703–2710, 2017.

[6] A. Makki and C. Bleakley, "WLAN Indoor Ranging Dataset." https://www.kaggle.com/ahmedmakki/toa-dataset/data. Accessed 2018-05-09.

[7] A. Makki, A. Siddig, M. Saad, and C. Bleakley, "Survey of WiFi positioning using time-based techniques," *Computer Networks*, vol. 88, pp. 218–233, 2015.

[8] H. Reddy, M. G. Chandra, S. Harihara, P. Balamuralidhar, J. Sen, and D. Arora, "WLAN-based local positioning using distorted template," in *Int. Symp. on Commus. and Info. Tech.*, pp. 1043–1048, 2007.

[9] A. Ferrero, S. Salicone, and S. Toscani, "A fast, simplified frequency-domain interpolation method for the evaluation of the frequency and amplitude of spectral components," *IEEE Trans. Instr. and Meas.*, vol. 60, no. 5, pp. 1579–1587, 2011.

[10] X. Li and K. Pahlavan, "Super-resolution TOA estimation with diversity for indoor geolocation," *IEEE Trans. Wireless Comms.*, vol. 3, no. 1, pp. 224–234, 2004.

[11] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.

[12] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoustics, Speech, and Sig. Proc.*, vol. 37, no. 7, pp. 984–995, 1989.

[13] F. Zhao, W. Yao, C. C. Logothetis, and Y. Song, "Comparison of super-resolution algorithms for TOA estimation in indoor IEEE 802.11 wireless LANs," in *Int. Conf. Wireless Communications, Networking and Mobile Computing*, pp. 1–5, 2006.

[14] J. Xiong, K. Sundaresan, and K. Jamieson, "Tonetrack: Leveraging frequency-agile radios for time-based indoor wireless localization," in *Int. Conf. Mobile Computing and Networking*, pp. 537–549, 2015.

[15] D. Vasisht, S. Kumar, and D. Katabi, "Decimeter-level localization with a single wifi access point," in *13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16)*, pp. 165–178, 2016.

[16] J. Torres-Sospedra, R. Montoliu, A. Martínez-Usó, J. P. Avariento, T. J. Arnau, M. Benedito-Bordonau, and J. Huerta, "UJIIndoorLoc: A new multi-building and multi-floor database for wlan fingerprint-based indoor localization problems," in *Indoor Positioning and Indoor Navigation (IPIN), 2014 Int. Conf.*, pp. 261–270, IEEE, 2014.

[17] R. Montoliu, E. Sansano, J. Torres-Sospedra, and O. Belmonte, "Indoorloc platform: A public repository for comparing and evaluating indoor positioning systems," in *Indoor Positioning and Indoor Navigation (IPIN), 2017 Int. Conf.*, pp. 1–8, IEEE, 2017.

[18] J. Torres-Sospedra, A. R. Jiménez, S. Knauth, A. Moreira, Y. Beer, T. Fetzer, V.-C. Ta, R. Montoliu, F. Seco, G. M. Mendoza-Silva, *et al.*, "The smartphone-based offline indoor location competition at IPIN 2016: Analysis and future work," *Sensors*, vol. 17, no. 3, p. 557, 2017.

[19] G. M. Mendoza-Silva, P. Richter, J. Torres-Sospedra, E. S. Lohan, and J. Huerta, "Long-term wifi fingerprinting dataset for research on robust indoor positioning," *Data*, vol. 3, no. 1, p. 3, 2018.

[20] H. Hashemi, "The indoor radio propagation channel," *Proc. of the IEEE*, vol. 81, no. 7, pp. 943–968, 1993.

[21] R. Akl, D. Tummala, and X. Li, "Indoor propagation modeling at 2.4 ghz for ieee 802.11 networks.," in *Wireless and Optical Comms.*, 2006.

[22] "WARP Project." http://warpproject.org. Accessed 2016-11-20.

[23] "WARP v3." http://mangocomm.com/products/kits/warp-v3-kit. Accessed 2017-10-20.

[24] A. Makki, A. Siddig, M. M. Saad, J. R. Cavallaro, and C. J. Bleakley, "High-resolution time of arrival estimation for OFDM-based transceivers," *Electronics Letters*, vol. 51, pp. 294–296, February 2015.

[25] "Directional WiFi antenna panel." http://www.simplewifi.com/indoorpanel.html. Accessed: 2015-01-13.

[26] "Microchip RN-SMA-4 antenna." http://ie.farnell.com/microchip/rn-sma-4/antenna-2-4ghz-4in-sma-male-conn/dp/2143321?Ntt=2143321. Accessed: 2016-10-13.

[27] "Data sheet of RG316 Coax Cable." https://www.pasternack.com/images/productPDF/RG316-DS.pdf. Accessed 2018-05-05.