Indoor TOA Error Measurement, Modeling, and Analysis
Ian Sharp and Kegen Yu, Senior Member, IEEE

Abstract—This paper presents a comprehensive investigation on the error characteristics of time-of-arrival (TOA) measurements obtained from positioning systems in indoor environments where rich multipath and nonline-of-sight propagation exist. By careful analysis of the measured data from three different settings of measurements, a model for the range errors is developed. The model has two components, one associated with the TOA measurement in the receiver and another associated with the delay excesses accumulated along the propagation path. By modeling multipath signals mathematically, and application of the central limit theorem, it is shown that statistical shape of the leading edge of a received pulse is essentially independent of the details of the scattering environment. Application of this statistical shape of the leading edge allows estimates of the statistical distribution of the TOA measurements to be calculated, either analytically or numerically from simulations. The second component of the delay excess is a model based on the number of walls along the path. Statistical performance based on this combined model is shown to be in good agreement with measured data collected from three different systems that have different RF frequencies, signal bandwidths, and TOA detection algorithms. The insights afforded by the theory assist in the design of more accurate positioning systems.

Index Terms—Experimental verification, leading edge algorithm, multipath and nonline-of-sight (NLOS) propagation, RF propagation through walls, TOA (time-of-arrival) error modeling and analysis.

I. INTRODUCTION

THE development of mobile radio systems in recent years has been dramatic, both for wide-area outdoor use (data and position determination) as well as indoor data communications based on the IEEE 802.11 standards. While applications, such as tracking people and assets [1]–[5] have been suggested, indoor positioning is still mainly in its infancy, partly due to the difficult operating environment for radio transmissions. In such indoor environments the global positioning system does not work well since the satellite signals are blocked by the buildings, and typically nonline-of-sight (NLOS) multipath conditions predominate. The design and performance estimation of indoor positioning systems is challenging as the rich multipath indoor radio propagation environment makes accurate range measurements difficult. In particular, the scattering of the radio signals results in ranging measurement errors consisting of a biased component as well as the random error typical of line-of-sight (LOS) outdoor positioning systems. The measurement and modeling of the ranging performance in both LOS and NLOS conditions has been much reported in [6]–[18], but the results are difficult to interpret in a generic sense due to the variations in measurement techniques, radio frequencies, signal bandwidths, range, and the characteristics of the building. Because of the random nature of the received signal, the measurements are typically modeled as range error statistical distributions, which are the empirically fitted to a statistical model. However, such data are difficult to apply in other operating environments, as the model parameters are only appropriate to the measurement technique and the particular indoor environment. In practice measured errors will have contributions associated with both the scattering environment (a function of the building architecture and the radio frequency) and the measurement method for determining the time-of-arrival (TOA) of the signal at the receiver. However, by applying the statistical model proposed by this paper, it is shown that the measured statistical results can be explained without resorting to direct parameter fitting of the statistical models to the measured data. Note that the model is intended to apply to indoor positioning systems with ranges similar to that of NLOS Wi-Fi data transmissions using IEEE 802.11 (say of the order of 50 m), and not short-range (say up to 10 m) typical of IEEE 802.15.4 (Zigbee).

The proposed method of estimating range errors consists of splitting the range errors into two components, namely those associated with NLOS propagation from the transmitter to the receiver, and second, the errors associated with the detection of the signal within the receiver. While these two components are coupled in the measured data, this paper describes methods of decoupling the relative contributions, which allows comparisons of the model with the measured data. Furthermore, it will be shown that in a NLOS environment the statistical characteristics of the leading edge of pulse detected by the receiver are essentially independent of the building characteristics, so that parameters, such as the zero-range bias error and its associated statistical variation, can be predicted for any particular leading edge TOA detection algorithm. Additionally, it is also shown range errors associated with the propagation path can be explained by propagation delay excesses associated with the number of walls along the path. As the number of walls along a path is both a function of the architecture of the building and...
the range, the path component to the measured range errors can be estimated (at least in theory) by use of a map of the building and the delay excess per wall. As a consequence, performance of any particular positioning system, from relatively narrow bandwidth Wi-Fi based systems to ultrawide band (UWB), in any particular building can be estimated by application of the general model proposed using the appropriate parameters.

Thus, in summary, the main contributions of this paper are as follows.
1) Presentation of measurements of range errors in typical indoor operating conditions for two different actual positioning systems, rather than using simulated data.
2) Analysis of the measured data to obtain insights into the mechanisms causing range errors.
3) Description of a leading edge threshold TOA algorithm with superior performance characteristics, and providing analytical expressions to determine its probability density function (PDF) of range errors.
4) Development of an analytical model for the shape of the leading edge in NLOS conditions.
5) Development of a statistical model of delays associated with propagation through walls.
6) Comparison of measured and model statistical data for three different sets of measured data showing good correlation.

The details of the TOA error analysis and statistical models are presented in the remainder of this paper. Section II briefly reviews the theory and concepts of indoor multipath radio propagation, and presents three sets of measured TOA data. Section III develops an analytical statistical model of the shape of the leading edge of the received multipath pulse; this shape model can be applied to estimate the performance of any leading-edge TOA algorithm. Section IV applies the theory of Section III to analyze the performance of the leading edge algorithms actually used for the measurements described in this paper. Section V develops a statistical model of the ranging excess errors associated with internal walls of buildings. Section VI compares the measured TOA statistical results to the theory developed in the previous sections. Finally, Section VII provides a discussion and conclusion on the consequences of the range error modeling and analysis described in this paper.

II. INDOOR RADIO PROPAGATION MEASUREMENTS

The concept of this paper is to develop a generic statistical model of NLOS TOA range errors that would be applicable to indoor mesh positioning systems with ranges similar to Wi-Fi data transmissions. To guide the development of the model, radio propagation and TOA measurements were performed at a number of RF frequencies and bandwidths from that typical of Wi-Fi channels to 4 GHz UWB. By analyzing the measured data documented in this section, the generic mathematical model is developed in Sections III–V.

A. Indoor Propagation Characteristics

Indoor radio propagation is complex, with multiple scattering, diffraction, and signal attenuation associated with walls and other obstacles along the path from the transmitter to the receiver. Because of this complexity, analysis requires considerable simplifications. One possible approach is to adopt the uniform diffraction theory [19], so that the radio signals can be represented as rays, incorporating both reflected and diffracted paths. With this interpretation of the scattered signal, the received signal can be expected to be composed of multiple individual components with varying amplitudes, phases, and delays depending on the path from the transmitter to the receiver.

Now consider a typical example of a UWB measurement in an indoor environment, as shown in Fig. 1. The data in this figure was measured using a Network Analyzer method [27, Appendix B], whereby the windowed broadband spectrum is converted to the impulse response using an inverse Fourier transform. The impulse response shows that signal scattering results in a large number of received signals, with a general trend of decreasing amplitude as a function of the delay excess. With these high-bandwidth measurements, individual scattered signals down to the measurement resolution (in this case 1 ns) can be observed. Clearly to accurately measure the TOA, the system should use the first detectable signal, namely using the leading edge of the impulse. Similar measurements reported in [11], [13], and [14] used the first detectable peak instead of the leading edge, but the resulting standard deviations (STDs) in the range-errors for the same bandwidth are 2–4 times worse compared with the examples in Section II-B. Indeed, using measured data in a NLOS environment it is shown in [20] that leading edge algorithms significantly outperform other common methods using peak data, including super-resolution methods, such as MUSIC.

B. Measured Range Error Data

To provide a variety of measured TOA data for later analysis, three separate data sets will be used, covering a wide variety of RF frequencies, signal bandwidth, and TOA detection algorithms. Because of the extensive nature of the data sets, only a limited sample is included in the following subsections. Unlike most reported TOA measurements, two of the data sets relate to actual positioning systems, while...
The third UWB case is extracted from extensive measurements described in [21]. An overall summary of the systems and their parameters is presented in Table I. For more details on the UWB, wireless ad hoc system for positioning (WASP), and the precision location system (PLS), refer to Section VI. The results shown in Fig. 2(a)–(c) are representative of the larger set of measurements, all of which have similar characteristics. The data in the figures show the scatter of range errors plotted as a function of range. Also shown is the linear trendline (solid line) and the two lines ± one pulse rise-time relative to the trendline (dashed lines). The data are analyzed to determine the zero-range intercept bias error (δ0), the range error STD relative to the trendline (σr), and the slope (λ) of the trendline. Note also that all the measurements have NLOS conditions, i.e., there is at least one wall between the transmitter and the receiver. Observe that when the data are presented in this manner, despite the different parameters described above, the generic characteristics are quite similar, suggesting common underlying principles.

1) UWB Measurements: The UWB measurement data in Fig. 2(a) are extracted from an extensive series of measurements reported in [21]. This reference contains data, which includes the range errors with an associated range, and for the NIST building a map of the measurement locations, allowing later comparisons with theory. The measurement method [27, Appendix B] is based on Gaussian windowing the spectrum from a Network Analyzer, with the pulse determined by an inverse Fourier transform. The TOA is determined using a threshold set at a multiple (usually 12) of the noise root-mean-squared (RMS) level. Other similar but less comprehensive UWB data are given in [22].

The zero-range intercept bias for the NIST building is 0.41 ± 0.21 m. This value seems unusually large for UWB, and probably implies some opaque walls rather than the cinder block construction stated. Data for two other buildings given in [21] have intercept biases of 0.21 ± 0.17 and 0.08 ± 0.17 m. Note that the large statistical variation in these estimates shows that the zero-range intercept parameter is difficult to determine accurately from measured data. Similar bandwidth data in [22] also reports the NLOS bias increases linearly range, with a zero-range bias of 0.019 m and a slope of 0.022.

2) WASP Measurements: The measurements in Fig. 2(b) are based on the WASP [23]–[25]. This experimental system has been designed to achieve an indoor positioning accuracy of 1 m (or better) based on a mesh network of nodes. The concept is based on using Wi-Fi analog chip radios, but generating an effective wider 125 MHz channel by quickly scanning over time eight 20 MHz Wi-Fi channels [25], and then generating a pulse with a rise-time of 13.5 ns by appropriate signal processing of the spread-spectrum signal. The measured range data are from intermodal round-trip-time measurements [25], [26], [28] using TOA detection based on a threshold set at a level of 18 dB below the local peak near the leading edge, as described in [24] and Section IV-A.

3) PLS Measurements: The measurements in Fig. 2(c) are based on the CSIRO developed PLS, which operates in the 2.4-GHz ISM band. This system was designed for outdoor sporting applications, and in particular horse racing. The designed positional accuracy outdoors is 0.5 m. The system is not intended for operation indoors, but testing was performed to assess its performance indoors.1 The pulse with a rise-time of 25 ns is generated by despreading a 40 Mchips per second direct sequence spread-spectrum signal. In the testing, range was determined using a round-trip-time technique [26] between two base stations, similar to that used by the WASP. The TOA measurements are based on the Ratio Algorithm described in Section IV-C.

C. Interpretation of Measured Data

While the measurement systems described in Section II-B are quite different, it is evident that the overall characteristics of the range error data are quite similar. In particular, note the following characteristics.

1) Nearly all the data lie within the two boundary lines defined relative to the least-squares (LSs) fitted linear trendline and the pulse rise-time. This characteristic is to be expected from a good leading-edge algorithm which is not affected by signal delays greater than the rise-time of the pulse.

2) The trendline delay excesses are a linear function of range. This characteristic is interpreted as the direct path signal being below the detection threshold of the leading edge TOA algorithm due to signal attenuation (such as through walls) and scattering along the straight-line NLOS path from the transmitter to the receiver.

1The unsatisfactory PLS performance indoors prompted the development of the WASP for indoor applications.
This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT

Fig. 2. (a) UWB range errors (NIST building). Frequency 5 GHz, Gaussian window channel-width 500 MHz (250 MHz 3-dB bandwidth), and pulse rise-time 3 ns (0.9 m). Range error STD (detrended) 0.42 m, slope 0.0355, and zero-range bias 0.41 m. Number of samples is 38 (2 bad). Cinder block walls. (b) WASP range errors (Marsfield Campus buildings, CSIRO). Frequency 5.8 GHz, channel (reconstructed) bandwidth 125 MHz, and pulse rise-time 13.5 ns (4.0 m). Range error STD (detrended) 1.14 m, slope 0.0209, and zero-range bias 0.87 m. Number of samples 191 (1 bad). (c) PLS range errors (Pawsey building, CSIRO). Frequency 2.4 GHz, channel bandwidth 60 MHz, and pulse rise-time 25 ns (7.5 m). Range error STD (detrended) 3.0 m, slope 0.156, and zero-range bias 0.84 m. Number of samples 128 (3 bad).

3) The scatter about the linear trendline is broadly similar, and largely independent of range, although some extreme range errors (termed bad data in Fig. 2) tend to be at longer ranges. This characteristic is interpreted as a range-independent variation associated with the measurement of TOA in the receiver, and the path scattering of errors being independent of range.

4) The zero-range intercept is nonzero. This characteristic is interpreted as being associated with the TOA detection algorithm as it is independent of the delay excess, which increases linearly along the path.

Based on the above discussion and Fig. 2(a)–(c), indoor ranging errors have two main components, namely the errors that accumulate along the path associated with scattering and passing through walls, and errors associated with the determination of the TOA in the receiver. Thus, NLOS ranging errors can be expressed as the sum of these two types of errors, namely

\[ \varepsilon_{\text{NLOS}} = \varepsilon_{\text{TOA}}(\tau_{\text{pulse}}) + \varepsilon_{\text{path}}(R, N_{\text{walls}}; f_c, \text{BW}). \]  

The errors from TOA determination will be a function of the TOA detection algorithm and the rise-time of the leading edge of the receiver baseband pulse (\(\tau_{\text{pulse}}\)). The path errors due to NLOS propagation will be a function of the path length (\(R\)) and it is proposed by the number of walls through which the path passes (\(N_{\text{walls}}\)), while the RF signal carrier frequency (\(f_c\)) and the signal bandwidth (\(\text{BW}\)) will be constant parameters for any given positioning system. Note that often in the literature (such as in [11], [12], [14], and [19]), the bias errors are averaged over all data ranges, so this range-related effect cannot be inferred from the published results.

To assist in the understanding of the detailed development of the range error model (1) in the following sections, Fig. 3 summarizes the flow of the mathematical development of the model in terms of the TOA detection algorithm, and the path delay excesses through walls in Sections III–V. In general, the mathematical development determines the mean (bias) ranging errors and the random variations in the delays expressed in terms of the PDFs, and hence the associated means and STDs. The complete mathematical model is then compared with the measured performance derived from the three sets of raw range error data summarized in Section II-B. Also note that the \(\sim\) symbol signifies a random variable or a function associated with a random process throughout the following text and equations.

III. NLOS LEADING EDGE ANALYSIS

This section provides a theoretical analysis of the shape of the leading edge of the received pulse in a NLOS environment. The aim of the analysis is to obtain a more general insight into the characteristics of the leading edge of the NLOS multipath pulse shape. In the following analysis, it is assumed that the nominal (no multipath) received pulse is triangular in shape due to bandlimiting. For direct-sequence spread-spectrum systems (such as the WASP and the PLS) this is a good approximation to the despread signal, while for the inverse fast Fourier transform (FFT) method of detection (UWB case in Section II) the shape depends on the weighting function, which is Gaussian in this case. The triangular shape assumption greatly simplifies the mathematics without greatly affecting the results due to the consequences of the central
limit theorem (CLT) associated with the summing of the many multipath signals.

A. Equi-Amplitude With Random Phase Case

The initial (over simplified) case considered is where all the multipath signals are of equal amplitude but with random phase; this simplifying assumption is used in the following analysis, but variation in the amplitude is considered later in Section III-B.

In a severe multipath environment there will be a large number of interference signals, the amplitude of which tends to decrease with increasing delay excess (Fig. 1). For theoretical analysis the indoor NLOS multipath signal amplitude distribution can be approximated as an exponential function of delay excess, with the exponential parameter defining the rate of decay. If the pulse rise-time is much less than this decaying period, then to a first-order approximation they can be considered of equal amplitude. For example, the WASP system described in Section II-B2 has a pulse rise-time of 13.5 ns, which is small compared with the delay spread shown in Fig. 1.

Now consider the signal phase characteristics. As the path lengths are much greater than the wavelength, even small changes in the path length will result in a rapid change in phase. Thus, it is reasonable to assume that the interference signals will have random phase, typically assumed to have a uniform distribution over \([0, 2\pi]\).

As only interference signals with delays up to about the pulse rise-time \(\tau_{\text{pulse}}\) can affect the performance of a leading edge algorithm, without loss of generality only interference signals with delays up to \(\tau_{\text{pulse}}\) are considered. Furthermore, these signals are assumed to be uniformly distributed in time throughout this period with a separation of \(\delta \times \tau_{\text{pulse}}\). Thus in the following analysis, it is convenient to normalize the time by \(\tau_{\text{pulse}}\), so the time separation of individual signal components is \(\delta\) in normalized time. In addition, it is convenient to normalize the pulse amplitude to unity. While the delay excesses are assumed to be equally spaced, simulations showed that other random distributions (such as statistically uniform distribution of delays) have little effect on the characteristics of the leading edge of the pulse.

First consider that all the UWB delta function components of the signal (similar to that shown in Fig. 1) are in phase and of unit amplitude. The actual transmitted signal will be of narrower bandwidth, so the received signal will be the UWB signal (delta functions) convolved by the impulse response of the bandlimited baseband signal, which as stated in the introduction to this section is assumed to be triangular in shape. For this special case the convolution is easy to calculate, namely the summation of the delayed unit-amplitude signals, so that starting at the leading edge the cumulative pulse amplitude at the normalized times \(0, \delta, 2\delta, 3\delta, \ldots\) will be

\[
0, \delta, 2\delta + \delta, 3\delta + 2\delta + \delta, \ldots \delta \sum_{n=1}^{N} n \]

(2a)

where there are \(N\) delta functions in the UWB signal within the bandlimited pulse rise-time. Note that the amplitude of first (associated index \(n = 1\)) convolved delta function component has an amplitude of \(N\) at time index position \(N\), and conversely the delta function component at \(n = N\) has unit amplitude. This is a consequence of the convolution process, which has a index \(N-n\) associated with the triangular
convolution process. These concepts are also shown in Fig. 4, showing the components of a (simulated) NLOS pulse. If the
nth UWB signal has a random phase $\phi_n$, resolving the signal into real and quadrature components the magnitude of the
multipath pulse at pulse time $\tau = N\delta$ is given by

$$
\tilde{M}_N = \delta \left[ \sum_{n=1}^{N} n \cos \phi_{N-n} \right]^2 + \left[ \sum_{n=1}^{N} N \sin \phi_{N-n} \right]^2.
$$

(2b)

The calculation of the statistics of the magnitude of pulse at
index position $N$ using (2b) is difficult due to the nonlinearity
of the square-root function, so an alternative approximate
method is used. Because the components in (2b) are the
summation of a number of random variables, each component
will have an approximate Gaussian distribution due to the CLT.
Furthermore, the expected value of the summation will be
zero, as the expected values of the cosines and sines are zero.
Under such circumstances the magnitude of the multipath pulse at the $N$th index position will exhibit approximately
a Rayleigh statistical distribution, provided $N$ is sufficiently
large; this simplification is later confirmed through simulations.
At this stage, it is also worth noting that because of
the randomizing effect of the random phase. Using this Rayleigh
statistical distribution, due to the CLT this result also applies to any
random phases, the expected value of the magnitude-squared
is given by

$$
E[\tilde{M}_N^2] = \delta^2 \sum_{n=1}^{N} n^2 = \delta^2 S(N) \approx \delta^2 N^3/(N \gg 1)
$$

(4)


While the expected value of the magnitude cannot be easily
calculated, by assuming a Rayleigh statistical distribution the
variance to the mean-squared ratio has a known value of
$4/\pi - 1$. Thus

$$
\frac{\sigma^2}{\mu^2} = \frac{4}{\pi} - 1 = \frac{E[\tilde{M}_N^2]}{E[\tilde{M}_N]^2} - 1.
$$

(5)

Combining (3) and (5) gives the mean of the magnitude as

$$
\mu(N) \approx \sqrt{\frac{\pi}{4} E[\tilde{M}_N]} = \delta \sqrt{\frac{\pi}{2} S(N)} \approx \frac{\delta}{2N} \sqrt{\frac{\pi}{3} N^3} (N \gg 1).
$$

(6)

Note that (6) also applies at any position on the leading edge
(namely at the $n$th position) as well as the $N$th position.
In the context of Fig. 1 where the high-resolution impulse has $N$ distinguishable impulses (delta functions) in the time-
normalized leading edge, then $\delta = 1/N$ and $\tau_n = n\delta = n/N$.
As the pulse amplitude does not affect the TOA measurement,
the $N$th position can be used to normalize the amplitude of the $n$th position of the multipath pulse, so the expected value of
the magnitude of the normalized pulse is given by

$$
\hat{\mu}(\tau) = \frac{\mu(n)}{\mu(N)} = \frac{N^{3/2}}{n^{3/2} N} \equiv \tau^{3/2}
$$

(7)

where the last expression in (7) is the equivalent continuous
analytical function, so that it can be used in theoretical calculations later. Finally, assuming a Rayleigh statistical distribution,
the ratio of the STD to the mean can be used to estimate the STD of the leading edge normalized shape, namely from (5) and (7)

\[ \hat{\sigma}(\tau) \approx \frac{4}{\pi} \sqrt{1 - 1 \tau^{3/2}}. \]  

(8)

An important consequence of the above analysis on the design of a leading edge TOA algorithm is the rapid increase in the STD of the multipath pulse noise as a function of delay from the epoch of the pulse. Clearly, the best quality measurements are near the epoch of the pulse, and the worst data are near the peak. This observation explains the large TOA variation for algorithms based on the peak of the first detectable signal.

### B. Random Amplitude and Phase Case

While the equi-amplitude analysis provides useful estimates of the shape of the leading edge, actual propagation conditions will have varying amplitudes (Fig. 1), so that the statistical characteristics with varying signal amplitudes and phases are analyzed in this section. From the analysis in Section III-A, the expected value of the magnitude-squared is

\[ E[\tilde{M}_N^2] = \delta^2 E \left[ \sum_{n=1}^{N} n^2 \tilde{a}_{N-n}^2 \right] = \delta^2 \sum_{n=1}^{N} n^2 E[\tilde{a}^2]. \]  

(9)

The expected value of the signal amplitude-squared in (9) can be evaluated if the statistics of the signal amplitude are known. For example, with a Rayleigh distribution of the amplitude the corresponding magnitude-squared is

\[ E[\tilde{M}_N^2] = \frac{1}{3} N(2N + 1)(N + 1)(\delta \sigma)^2 \approx \left( \frac{2}{3} (\delta \sigma)^2 \right) N^3. \]  

(10)

If (10) is compared with (4) it can be observed that the only difference is the constant term, which is also true for any other amplitude statistical distributions, albeit with a different constant. Thus, if the multipath pulse is normalized in the same manner as described in Section III-A, the resulting mean shape will be the same, namely as defined by (7). Thus provided there is a sufficiently large number of scattering sources such that statistics of large numbers is reasonably valid (about five sources, see [27 Sec. 2.1.4]), the mean shape of the leading edge is independent of the signal amplitude statistics. The STD in the shape of the leading edge can also be calculated in a manner similar to that described in Section III-A.

### C. Simulation of Multipath Pulses

To show the above concepts, Fig. 4 shows the results of a simulation using the concepts described in Section III-B. The figure shows the nominal triangular pulse shape, and the multipath pulse corrupted by noise. The multipath signals have a Rayleigh amplitude distribution, while the phases have a uniform distribution over \([0, 2\pi]\). The time axis is normalized by the nominal pulse rise-time of the triangular pulse, and the multipath pulse amplitude is normalized to unity. Observe the concave curvature of the multipath pulse, which is also delayed relative to the nominal pulse. Each random pulse will have a different shape, but the above analysis shows the on average the shape is defined by (7). The TOA detection algorithm task is to estimate the arrival time with minimal errors compared with the ideal pulse shape, despite the multipath pulse distortion and the effects of receiver noise.

### IV. Performance of TOA Algorithms

Section III established the shape of the leading edge pulse, plus its statistical variation; this model is now applied to two TOA algorithms as used in the actual measurement systems described in Section II. Note, however, that the merits of these algorithms are not being assessed, but rather the analysis allows comparisons with measured data.

Before considering particular TOA algorithms, it is useful to define desirable characteristics of leading edge algorithms. Ideally, the measurement of TOA for a positioning system should:

1. scan from the noise region towards the signal region to locate the first detectable signal for further processing;
2. be independent of the pulse amplitude;
3. be independent of the noise level before and in the signal pulse;
4. should try to minimize both the effects of noise and multipath interference; this typically requires a compromise in performance between these two effects;
5. at high signal-to-noise ratio (SNR) and low multipath, the TOA should approach a constant value, called the epoch of the pulse;
6. all nodes in a network should have similar characteristics.

There are many possible TOA detection algorithms, but the two described in the following subsections are based on the three particular measurement examples used in this paper. The merits of each algorithm are not particularly relevant for this paper. Nevertheless, all the algorithms are based on a small section of the impulse response near the leading edge. In all three sets of data described in Section II, the TOA pulse is derived using an inverse Fourier transform from various wideband/spread-spectrum signals. As an FFT is very computationally efficient, the pulse generation is also very efficient. In addition, interpolation of the bandlimited signal can be efficiently performed using the FFT method. The actual determination of the TOA of the pulse requires a trivial additional amount of time compared with the basic pulse generation.

#### A. Threshold Algorithm

The concept of a TOA Threshold Algorithm [9], [12], [20] is simple—the epoch of the pulse is defined by the position where the leading edge of the (normalized) pulse crosses (exceeds) a specified threshold level. In the nominal pulse without multipath or noise corruption, the threshold amplitude \((a_{th})\) can be defined by the relationship to the peak \((A)\) of the pulse, namely

\[ a_{th} = \alpha A \]  

(11)

where \(\alpha\) is a fixed parameter of the algorithm so that the threshold time position (pulse epoch) is independent of the
pulse amplitude. The corresponding reference epoch is $a \tau_{\text{pulse}}$ from which ranging errors can be determined for a multipath pulse. Hence, in the ideal case the errors will be zero. However, in the presence of receiver noise and multipath signals defining the threshold level is more difficult. In particular, in a multipath environment, the peak signal can be remote from the leading edge (Fig. 1), so that applying (11) will result in a threshold different from that in the error-free case, resulting in range measurement errors. Second, receiver noise before the leading edge can be falsely interpreted as the leading edge of the pulse.

Clearly, from the above analysis of the variation in the shape of the leading edge the threshold should be as low as possible, but conversely should be well above the noise level. Furthermore, the multipath noise will result in noise-related errors in determining the peak amplitude ($A$), which in turn will result in errors in defining the threshold level and consequentially the pulse epoch. Based on this discussion and Fig. 4, the proposed Threshold Algorithm is summarized as follows.

1) Using the position of the pulse peak amplitude as a time reference guide, the noise section of the pulse (well before the peak) is located, and the noise ($\sigma_n$) determined. An initial noise threshold is defined as $k \sigma_n$, where $k$ is typically 3 or 4. The value of $k$ is used by the system designer to select between better multipath performance (smaller $k$), or better rejection (larger $k$) of false detection because of noise before the signal pulse.

2) The time position where the pulse leading edge exceeds the noise threshold is determined. To minimize false detection, a total of three adjacent samples is checked, and the search continues if all three do not exceed the threshold.

3) Due to the curved nature of the multipath leading edge [(7) and Fig. 4], this position will be somewhat greater in time than the true start of the leading edge, and will also be a function of the pulse SNR.

4) Using this noise threshold time, define a time position advanced by the known rise-time of the nominal pulse. The peak of the pulse within this time interval is located, and defined as the nominal peak $A$ for use in (11). Note that $a_{th} > k \sigma_n$ is assumed to hold; otherwise define the threshold amplitude as $a_{th} = k \sigma_n$.

5) The scan as in step (2) above is repeated with the updated threshold level to locate the pulse epoch.

This somewhat complex Threshold Algorithm significantly outperforms an algorithm with a fixed threshold in a multipath environment [20].

### B. Random Variation in the TOA Performance

While the TOA performance analysis can be based on the mean pulse shape described in Section III, this approach fails to account for the statistical variations in the leading edge. Thus, a more rigorous theoretical approach must apply the variability in the leading edge to the TOA algorithm, which results in the statistical variation in the epoch of the pulse; from this statistical variation (as defined by its PDF) the true mean performance can be calculated.

The Threshold Algorithm as defined above is complicated, so a rigorous analytical mathematical analysis is difficult; thus some simplifications are necessary. The determination of the epoch of the signal using the Threshold Algorithm is based on scanning the leading edge of the pulse from the noise section until the amplitude exceeds a threshold. As the leading edge in a multipath environment is random, the statistical performance should be based on the analysis of many such random pulses. However, an alternative equivalent approach that simplifies the mathematics is to reverse the order of the analysis, so that the statistical variation at each point on the leading edge is analyzed first, and then the probability that the statistical pulse exceeds the threshold is calculated.

The analysis in Section III showed that the pulse amplitude at each point on the leading edge has approximately Rayleigh statistics, so using (8) the complementary cumulative distribution function of the probability that a threshold amplitude $a$ will be exceeded at normalized time $\tau$ is given by

$$
\tilde{C}_e(\tau) = \exp \left[ -\frac{\pi a^2}{2 (4 - \pi) \tau^3} \right] \quad (\tau \geq 0) \quad (12a)
$$

and differentiating (12a) the corresponding PDF is

$$
\tilde{p}_e(\tau, a) = \frac{3 \pi a^2}{2 (4 - \pi) \tau^4} \exp \left[ -\frac{\pi a^2}{2 (4 - \pi) \tau^3} \right]. \quad (12b)
$$

The threshold value $a$, as given by (11), depends on the peak amplitude ($A$), which in this case is a statistical random variable at $\tau = 1$. As before, $A$ (and hence $a$) will have approximately a Rayleigh statistical distribution, so the PDF of the threshold amplitude is

$$
\tilde{p}_a(a, a) = \frac{\pi a}{(4 - \pi) \sigma} \exp \left[ -\frac{\pi a^2}{2 (4 - \pi) \sigma^2} \right]. \quad (13)
$$

Using this definition of the threshold amplitude, the estimate of the PDF of the pulse epoch is given by

$$
\tilde{P}_e(\tau, a) = \int_0^\infty \tilde{p}_a(a, a) \tilde{p}_e(\tau, a) da = \frac{3 a^2 \tau^2}{(a^2 + \tau^2)^2}. \quad (14)
$$

The corresponding mean and STD of the TOA using (14) are given by

$$
\mu_e(a) = \tau_{\text{pulse}} \left( \frac{2 \pi}{3 \sqrt{3}} a^{2/3} - a \right) \quad (15a)
$$

$$
\sigma_e(a) = \tau_{\text{pulse}} \left( \frac{2}{3} \sqrt{\pi (\sqrt{3} - \pi/3) a^{2/3}} \right). \quad (15b)
$$

These analytical results are compared with measurements in Section VI.

### C. Performance of Ratio Algorithm

This section considers the performance of another type of leading edge TOA algorithm (used by the PLS in Section II-B3), which is based on the concept that the shape of the leading edge is largely independent of the pulse amplitude and the receiver noise. In particular, the Ratio Algorithm described in detail in [27, Ch. 4] is based on determining the ratio ($\rho$) of the pulse amplitude at two points on the leading edge,
and compares it to a value $\rho_0$, which represents a zero-error set-point. The set-point value (typically about 0.4) is chosen to optimize performance (a compromise between noise and multipath tracking errors). A feedback loop adjusts the phase of a local clock to try to maintain the measured ratio at the set-point, namely $\rho = \rho_0$. The local clock then gives the TOA. The statistical performance analysis of the Ratio Algorithm is given in Appendix A.

To show an alternative to the analytical methods used for the Threshold Algorithm, the performance of the Ratio Algorithm will be evaluated by a Monte-Carlo simulation using the statistical characteristics of the leading edge established in Section III; simulation methods are appropriate when analytical calculations are difficult. The resulting cumulative distribution function (CDF) using parameters appropriate for the PLS is shown in Fig. 9.

V. PATH PROPAGATION

The analysis in Sections III and IV shows that in a heavily multipath NLOS environment the shape of the leading edge is distorted, resulting in TOA measurement errors. These effects are essentially independent of the length of the propagation path, but additional errors occur due to the characteristics of the multipath scattering along the path, resulting in a delay excess above the straight line path delay. This section analyses this delay excess based on the number of walls along the path, with each wall contributing an increment in the delay excess.

The path propagation error can be considered as having a mean (bias) error and a random component. The mean bias component will be a function of the range and the mean number of walls ($\bar{N}_{\text{walls}}$) along the path of range $R$. The random component for a given range would be associated with the statistical variation in the number of walls for that particular range, and the variation in the excess delay associated with the signal scattering at each wall. The path range errors associated with walls can thus be expressed in the form

$$\tilde{e}_{\text{path}} = (\gamma_0 + \Delta \tilde{\gamma})(N_0(R) + \Delta \tilde{N}_{\text{walls}})$$

(16)

where $\gamma_0$ is the mean delay excess per wall along a path in the particular indoor environment, $N_0(R)$ is the expected number of walls along a path of length $R$, $\Delta \tilde{\gamma}$ is a random variable (with zero mean) which represents the variation in mean delay excess per wall along various paths of length $R$, and $\Delta \tilde{N}_{\text{walls}}$ is the random variation in the number of walls along various paths of length $R$. These statistics can be analyzed both theoretically and from measured data. Also, the statistical variation in the number of walls along a path in a particular environment can be determined from architectural maps of the building.

The overall statistics for the range error along indoor paths can determined by expanding (16), and applying the statistical distribution of the two random components, namely $\Delta \tilde{\gamma}$ and $\Delta \tilde{N}_{\text{walls}}$. However, the analysis is considerably simplified by observing that in both cases the random component is much smaller than the mean of the component, so that the product of the random components can be ignored to a good order of accuracy, namely

$$\tilde{e}_{\text{path}} \approx \gamma_0 N_0(R) + [N_0(R) \Delta \tilde{\gamma} + \gamma_0 \Delta \tilde{N}_{\text{walls}}]$$

$$= N_0(R) \Delta \tilde{\gamma} + \gamma_0 \tilde{N}(R).$$

(17)

Thus, the range error associated with the path can be expressed as weighted sum of two random variables, one associated with the variation in the number of walls along the path and another associated with the variation in delay excess per wall. The PDF of the sum of these random components will be the convolution of their individual PDFs. The statistical analysis of each component is given in the following subsections.

A. Range Error Bias Models

The bias error model considered below (see measured results in Section II and elsewhere [29]–[34]) is that the mean bias errors can be modeled as a linear function of range. This model assumes that the scattering effects increase the bias linearly with range, but does not explicitly consider any specific mechanisms, such as the effects of walls and other large objects. This model can be expressed as

$$\tilde{e}_{\text{bias}} = \tilde{\delta} r_0 + \lambda R$$

(18)

where $\tilde{\delta} r_0$ is bias at zero range and $\lambda$ is the model parameter determined by a LS fit to the measured data, and will vary from building to building due to variations in the architecture. An alternative building architecture-based model uses the number of walls along a path and the apparent delay excess per wall as a proxy for the more complex scattering along the propagation path and through walls. The data in Section V-B suggests a mean bias model as

$$\tilde{e}_{\text{bias}} = \tilde{e}_0 + \gamma \tilde{N}_{\text{walls}}(R)$$

(19)

where the mean number of walls are calculated in range bins as explained in Section V-B. It will be shown in the next two sections that models (18) and (19) are in fact different manifestations of the same underlying statistical processes; model (19) is preferable theoretically as the parameters of the model are directly related (at least in principle) to the building architecture, while model (18) is practically easy to apply using measured data, as shown in [33].

B. Walls-Based Statistics Along a Path

Consider determining the parameters in (19) for a particular operating environment and positioning system. The mean bias errors can be determined from measurements (for example, as shown in Section II), and a map of the building can be used to determine the number of walls along each measurement path, so the $\gamma$ parameter in (19) can be estimated by a LS fit to the data. To reduce the statistical variation, it is useful to group the data into range bins, and analyze the mean bias error in a bin versus the mean number of walls associated with a range bin, as shown in the example in Fig. 5. The total measurement noise is reduced considerably by averaging the data in each range bin (about 30 samples per bin in this case), so the variation is reduced by about a factor of $1/\sqrt{30}$. As can
be observed in Fig. 5, there is apparently a linear variation as defined by (19).

The data, such as shown in Fig. 5, can provide important information about the nature of the bias errors. Observe that the delay excess for the first wall (about 0.6 m) is much more than the delays for subsequent walls (about 0.15 m). The interpretation is that the intercept bias \( \bar{\varepsilon}_0 \) at walls = 0 is not related to walls, but is due to some other process, namely the characteristics of the leading edge algorithm previously analyzed in Sections III and IV. Furthermore, it is concluded that the noise in the raw range error data is associated with both the TOA measurement variation and the variation in the delay excess associated with walls.

The data in Fig. 5 can also give further insight into the nature of the causes of delay excesses as a function of propagation along the path from the transmitter to the receiver. From Fig. 5, the estimated delay excess per wall is about 0.16 m, which is large compared with the thickness of the walls (about 5 cm), but small compared with the sizes of rooms and the distance separation between walls. These observations inferred from multiple measurements over a wide area are confirmed by specific individual measurements through one or two walls described in [28]. For example, the median delay excess (measured using the WASP hardware) through a glass wall was 0.22 m, and through a double-brick wall 0.15 m. For propagation through two internal walls the measured delay excess was 0.4 m, or about twice that of a single wall, and in agreement with the data in Fig. 5. Clearly, these measured delay excesses show that the main component of the excess is not associated with the delay expected from a single ray and the electrical properties of the material of the walls, but is a more complex relationship associated with multiple paths. The exact mechanisms associated with walls are beyond the scope of this paper, except the observation that the path delays are closely related to the number of walls along the path.

The number of walls for a path between two points can be determined from the bit map (black/white) generated from the architectural drawings of the building. A wall is a black bit (or bits) along the path, so that the number of walls can be easily counted. By averaging over a large number of random paths within the building (typical of paths from a mobile device to a base station), the building-wide average walls per unit length can be computed. Using this average parameter, the statistical variation in the number of walls for a given path length can be estimated. This is the classic description of a Poisson statistical problem, namely the actual number of walls along a path will be a random variable, which exhibits a Poisson statistical distribution.

An example of the statistical distribution of the number of walls in a range bin is shown in Fig. 6, together with the Poisson distribution with the same mean number of walls in the range bin data set. The similarity of the two distributions supports the assumption of Poisson statistics.

C. Statistics of Path Range Errors

This section provides a statistical analysis of the path range errors based on the model (19). This model has two random variables, namely the number of walls along a path of length \( R \) and the delay excess per wall. The number of walls per unit length (\( \eta \)) is considered to be a constant throughout the coverage area of the system, and can be determined from a map of the building. As a consequence the expected number of walls along the path of length \( R \) is \( N_0 = \eta R \), but the actual number of walls \( N(R) \) is expected to have a Poisson statistical distribution, namely

\[
\tilde{P}(n, N_0) = e^{-N_0} N_0^n / n! \quad (n = 0, 1, 2, \ldots).
\]
The delay excess at each wall will be random variable of unknown statistical variation, but as the delay accumulates at each wall along the path the overall average delay per wall for the path is expected to be approximately a Gaussian random variable due to the consequences of the CLT. Thus, the $\Delta \gamma$ random variable in (17) is assumed to have the normal distribution $N(0, \sigma_{\gamma})$, where the STD $\sigma_{\gamma}$ is determined from the measured data. Note, however, that as the delay excess is always positive the STD in the normal distribution should be limited to about $\sigma_{\gamma} \leq \gamma_0/3$, where $\gamma_0$ is the mean delay per wall. Now consider the overall statistics of the range errors associated with a path. From (20) the expected (mean) error is

$$E[\tilde{\varepsilon}_{\text{path}}] = \gamma_0 N_0(R) \approx \eta \gamma_0 R = \lambda R.$$  \hspace{1cm} (21)

As from (1) there is also a contribution to the range errors from the TOA algorithm, the expected value of the range errors cannot be modeled by (21) alone. However, as the TOA algorithm statistics are independent of the path range, with an offset bias of $\varepsilon_0$ and a zero-mean random component, the total range error can be modeled as

$$\tilde{\varepsilon}_{\text{NLOS}}(R) = \varepsilon_0 + \eta \gamma_0 R + \text{Noise}(R)$$  \hspace{1cm} (22)

where the Noise($R$) has zero expectation at all ranges. Observe that (22) is the same as model (18), which shows that the range-based model (18) can be predicted from the wall-based model. Thus, plotting the measured range error against range is expected to show a noisy linear trendline, which indeed is what is seen in the measured data in Section II and in [31], [33], and [34]. Further, performing a LS fit to the data using the linear model (22) allows estimates of parameters $\varepsilon_0$ and $\lambda = \eta \gamma_0$. Additionally, as the $\eta$ parameter can be estimated from the map of the building, an estimate of the mean delay excess per wall ($\gamma_0$) can also be determined from the measured range error data.

Now consider determining the PDF of the random component of the path errors $\tilde{\varepsilon}_{\text{path}}$ based on (17). From the above discussion, and using $\varepsilon$ as a dummy variable in the following analysis, the statistical distribution of the first component in (17) is a Normal (Gaussian) distribution

$$\tilde{g}(\varepsilon; \sigma_{\gamma}) = N(0, \gamma_0 R \sigma_{\gamma}).$$  \hspace{1cm} (23)

The PDF of the second component is a Poisson distribution, so the mean delay excess through $n$ walls is $\tilde{\varepsilon}_{\text{walls}}(n) = \gamma_0 n$ with a probability of $\tilde{P}(n, N_0)$. Finally, as from (17) the random component of total delay excess error is the sum of the Gaussian and Poisson distributed components, the PDF of the total path error is the convolution of the two PDFs given by (20) and (23), namely

$$\tilde{C}(\varepsilon, n) = \tilde{P}(n, N_0) \tilde{g}(\varepsilon - n \gamma_0; \sigma_{\gamma}).$$  \hspace{1cm} (25)

The total PDF is obtained by summing (25) over all $n$, namely

$$\tilde{p}_{\text{path}}(\varepsilon) = \sum_{n=0}^{\infty} \tilde{P}(n, N_0) \tilde{g}(\varepsilon - n \gamma_0; \sigma_{\gamma})$$  \hspace{1cm} (26)

which is effectively the weighted sum of shifted Gaussian distributions. Note that this distribution includes the LOS case ($n = 0$); if only NLOS cases are included, then the summation in (26) will commence at $n = 1$. The expectation (mean) of the ranging error is

$$E[\varepsilon] = \int_{0}^{\infty} \varepsilon \tilde{p}_{\text{path}}(\varepsilon) d\varepsilon = \sum_{n=0}^{\infty} \tilde{P}(n, N_0) \int_{-n \gamma_0}^{\infty} (x+n \gamma_0) \tilde{g}(x; \sigma_{\gamma}) dx.$$  \hspace{1cm} (27a)

Now as the usual case is that $n \gamma_0 \gg N_0 \sigma_{\gamma}$ (see above comment on limiting $\sigma_{\gamma}$) the lower limit can effectively be
extended to $-\infty$ with little error, so the last integral in (27a) can be evaluated to be $n\gamma_0$, and the expectation then reduces to

$$E[\varepsilon] \approx \gamma_0 \sum_{n=0}^{\infty} n \tilde{P}(n, N_0) = \gamma_0 N_0$$  \hspace{1cm} (27b)$$
as the summation in (27b) can be recognized as the mean of the Poisson distribution. The result in (27b) is in agreement with (21). A further consequence is that the combined distribution will have a peak near $N_0\gamma_0 = \eta\gamma_0 R$, as shown in Fig. 7. A similar analysis shows that the variance of the range errors is approximately

$$\text{var}[\varepsilon] \approx N_0\gamma_0^2 + N_0^2 \sigma^2.$$  \hspace{1cm} (28)$$
Examples of the distribution with typical parameters is shown in Fig. 7. As for a Poisson distribution the mean number of walls is linearly related to the range, the mean and STD are given by

$$\mu_{\varepsilon} = \eta\gamma_0 R$$
$$\sigma_{\varepsilon} = \eta R \sqrt{\frac{\gamma_0^2}{\eta R} + \sigma^2} \approx \gamma_0 \sqrt{\eta R}$$  \hspace{1cm} (29)$$
where the last approximation applies as (typically) $\gamma_0 \gg \sigma\sqrt{N_0}$. In this case the STD of the range error along the path only increases slowly with range, but this variation can be masked by a larger contribution to the STD from the TOA algorithm, which is independent of range.

Fig. 8 shows the STD typically for the three measurement data sets (UWB, WASP, and PLS) described in Section II. The results depend on the relative effects of the TOA algorithm and the path statistics. Observe in particular that the WASP has a STD which is practically independent of range. This effect is evident in Fig. 2(b), and also has been noted previously [33], [34] for measured data from the WASP system, contrary to a common assumption that the STD is proportional to range. This agreement between theory and measured data lends support to the analysis described above.

VI. STATISTICAL COMPARISON BETWEEN MEASUREMENTS AND THEORY

This section describes the comparison between theory and measurements of ranging errors statistical distributions, and are based on the raw range error measurements described in Section II and the statistical models described in Sections III–V. As the model has two components, one associated with the TOA algorithm and one associated with the propagation path, the combined model (as shown in Fig. 3) will include a particular TOA PDF [such as (14) for the Threshold Algorithm] and the Path model (24) to obtain the combined statistical model (PDF) of the ranges errors minus the bias errors. The two PDFs are combined by a convolution of the PDFs of these two statistically independent processes.

A. Leading Edge TOA Measurements Comparison

The measured raw range errors presented in Section II have contributions from both the leading edge detection method and errors associated with path propagation. These two random effects cannot be decoupled, but the random errors will be mainly associated with leading edge detection provided the path length is not too long (say ranges up to 10 m). A summary of the short-range leading edge data is given in Table II, comparing the measured and theoretical values based on the model in Section III. The mean and STD of the range errors contributable to the TOA are given (and the associated 90% confidence interval), as well as the model predictions. For the UWB case the match is somewhat poor as path errors dominate, and thus in this case a comparison with the leading edge theoretical results is not possible. However, the WASP theoretical and measured mean and STD data in Table II agree well, as the path errors are comparatively small in this case. For the PLS case Fig. 9 shows the CDF of the measured range errors and the corresponding Monte-Carlo simulation results for the Ratio Algorithm based on the leading edge theory in Section III. The results are quite
相似，表明该比值比算法模拟基于在第III节中描述的理论统计分布的测量数据。理论和理论平均值和标准差对于PLS在第表II中给出的值也相当接近。A Chi-squared goodness-of-fit analysis was also performed to check if the measured range errors are from the TOA model distribution. The Chi-squared test [34] is based on \( k \) range-bins counts in the histogram bins of the measured range errors, and the associated counts using the model PDF distribution. Then, it can be shown [34] that the summation \( V = \sum_{i=1}^{k} \left( \text{Measured}_i - \text{Model}_i \right)^2 / \text{Model}_i \) has approximately a Chi-squared statistical distribution \( \chi^2_{k-1} \). If \( V \) is not greater than the Chi-squared distribution at the \( 1 - \alpha \) level of significance, then the measured data are deemed consistent with the model statistical distribution. For the two distributions in Fig. 9, at the \( \alpha = 0.1 \) level of significance the values are \( [V, \chi^2_{k-1, 1-\alpha}] = [8.00, 10.65] \), so the Chi-squared test confirms that the measured range error data are consistent with the model statistical distribution.

### B. Comparison of Statistical Variations in Range Errors

The theory in Sections III–V suggests that range errors have contributions from two sources, namely the leading edge TOA detection process, and secondly from delay excesses associated with NLOS path propagation through walls. The statistical model for the first process has no parameters, which need to be determined from measured data, while the second process has two parameters \((\gamma_0, \sigma)\), which require matching to the measured data; a third parameter \((\eta)\) for the path propagation model can be determined from an architectural map of the building.

The following statistical results are obtained after the removal of linear trendline bias effects. In this case the residual errors are limited to those defined by the random component of the leading edge theory of Section III for a particular TOA algorithm in Section IV, and the path statistical distribution, namely as given by (24). It is shown in [33] that for mesh positioning systems the bias errors can be estimated and removed, so that the residual statistical ranging errors will effectively define the positional accuracy. The parameters \((\gamma_0, \sigma)\) are determined from a LS fit to the measured data, but there is little variation in the fit due to \( \sigma \) (changing this parameter mainly affects the tails of the distribution). In the model the mean number of walls \((N_0)\) is set at the overall mean for the building.

As with the comparisons of the statistical distributions in Fig. 9, the measured and model distributions shown
The CDF between the measured and model is 0.024, and 0.041, and the Chi-squared distribution must be reduced by two.

The Path model parameters are \( \gamma \) and \( \sigma_\gamma \). The RMS error in the CDF between the measured and model is 0.020, and 0.05 m, and \( N \_8 \). 092 m, and \( N \_16 \). (c) CDF of the bias-corrected range errors. The measured and statistical model CDFs for the random components of the WASP errors are shown in Fig. 10(a). The large number of samples (190) means that a close match should be possible, and indeed the model and measured distributions match very well, and the Chi-squared test passes; these good results provide further confidence in the validity of the model.

4) Summary of Statistical Results: The residual range error statistical distributions for the three different systems generally show a good match between the measurements and the theory. The residual errors have contributions from both the TOA algorithm and the path propagation, but the relative contributions depend on the specific parameters of the particular system. Nevertheless, the residual statistics are quite similar despite the bandwidths varying from 40 to 250 MHz. One noticeable characteristic common to all is that the statistical distributions are not Gaussian, with a noticeable upper tail to the distribution. However, the leading edge algorithms constrain the tail to about the pulse rise-time in each case, although the relative tail size somewhat decreases with the signal bandwidth.

VII. Conclusion

This paper investigated the effects of NLOS multipath scattering in an indoor environment on range measurements. Based on observations from measured data, a generic mathematical statistical model was developed, and was tested at three different RF frequencies, signal bandwidths, and TOA detection algorithms; good matches between the model and the measured data were observed. Based on the observed characteristics of the scattered signal with NLOS propagation it was shown that the leading edge of the pulse can be predicted (both the mean shape and its statistical variation) without any need for information about the propagation environment. Furthermore, a statistical model of scattering and propagation through walls correctly predicts a mean linear increase in the delay excess along the path, and also predicts the observed statistical variation in the measured data.

The consequences of the analysis in this paper have impacts on the performance of indoor positioning systems. First, although range bias errors have long been observed indoors, it was not appreciated that the expected zero-range bias is
similar for all paths. The consequence of this observation is
that pseudorange [range plus an (unknown) constant range
offset] positioning will be largely unaffected, as this bias can
be incorporated into the pseudorange constant. Second, it is
shown that the linear bias with range previously observed
in [31], [33], and [34] is associated with internal walls and as
such is expected to be common to most indoor situations. It is
shown in [33] that if this bias effect is correctly compensated
for, the positional errors associated with range-based position
fixing are approximately halved. Finally, these results mean
that designers of positioning systems can make a priori
performance predictions in particular operating environments
without having to resort to measurements or complicated
numerical simulations. However, the current model has one
parameter (mean delay per wall) which must be matched to the
measured data, although the data in this paper could be used
to estimate this parameter. In theory, this wall delay parameter
could be estimated by theoretical radio-wave analysis, and is
suggested as a topic for further investigation.

APPENDIX A

STATISTICAL PERFORMANCE OF THE RATIO ALGORITHM

This Appendix considers the performance of the Ratio
Algorithm introduced in Section IV-C, and described in detail
in [27, Ch. 4]. This TOA algorithm is based on determining
the ratio of two points on the leading edge, with the signal
epoch defined when this ratio is equal to the required set-
point value \( \rho_0 \). For a bandlimited signal used in real systems
the leading edge needs only to be sampled at a period of about
half the pulse rise-time, so that nominally only two points are
located above the noise on the leading edge, although the pulse
can be interpolated without error due to the Sampling theorem.
Thus, the Ratio Algorithm can efficiently use all the available
information about the leading edge.

The analysis of the performance of the Ratio Algorithm
is based on the statistics of the amplitude of samples (suf-
ciently separated in time) as established in Section III, namely
Rayleigh statistics. Thus, the statistics of the ratio is deter-
mined by deriving its PDF from two statistically independent
Rayleigh samples

\[
\rho = \frac{A_1}{A_2} \Rightarrow \frac{R(a, \sigma_1)}{R(a, \sigma_2)}.
\]  (A-1)

The determination of the required PDF is based on the
method described in [35, Ch. 9]. Consider two transformations
\( Y_1 = \frac{A_1}{A_2} \) and \( Y_2 = \frac{A_2}{A_2} \) such that random variables
are \( a_1 = Y_1 Y_2 \) and \( a_2 = Y_2 \). Then, the joint probability of the
product is given by

\[
fy_1 y_2 (y_1, y_2) = |J| f_{a_1 a_2} (y_1 y_2, y_2) \tag{A-2}
\]

where \( |J| \) is the Jacobian of transform equations.
As the two variables have Rayleigh statistics their joint
probability is

\[
f_{a_1 a_2} (a_1, a_2) = \frac{a_1 a_2}{\sigma_1^2 \sigma_2^2} \exp \left[ - \left( \frac{a_1}{2 \sigma_1^2} \right) - \left( \frac{a_2}{2 \sigma_2^2} \right) \right] \tag{A-3a}
\]

and thus from (A-2)

\[
f_{y_1 y_2} (y_1, y_2) = |y_2| \frac{y_1 y_2^2}{\sigma_1^2 \sigma_2^2} \exp \left[ - \frac{1}{2} \left( \frac{(y_1 y_2)^2}{\sigma_1^2} + \frac{y_2^2}{\sigma_2^2} \right) \right]. \tag{A-3b}
\]

Finally, the required PDF can be determined by integrating
out the dummy \( y_2 \) variable in (A-3b), namely

\[
f_{y_1} (y_1) = \frac{y_1}{\sigma_1^2 \sigma_2^2} \int_{-\infty}^{\infty} |y_2| y_2^2 \exp \left[ - \frac{1}{2} \left( \frac{(y_1 y_2)^2}{\sigma_1^2} + \frac{y_2^2}{\sigma_2^2} \right) \right] dy_2
\]
or

\[
f(y) = \frac{y_1}{\sigma_1^2 \sigma_2^2} \int_{0}^{\infty} x e^{-ax^2} dx = \frac{2\sigma_2^2 \rho}{(\rho^2 + \omega^2)^2} (\rho \geq 0) \tag{A-4}
\]

where \( a = 1/2[(\rho/\sigma_1)^2 + (1/\sigma_2)^2], \omega = \sigma_1/\sigma_2, \) and
\( \rho = y_1 \) from the original transformation. The corresponding
CDF obtained by integrating (A-4) is

\[
\text{CDF}_\rho (\rho) = \frac{\rho^2}{\rho^2 + \omega^2} (\rho \geq 0). \tag{A-5}
\]

The PDF defined by (A-4) is zero at ratios of both zero
and infinity, but has a large slowly converging upper tail,
characteristic of measured range error data. Using (A-4) it
can be shown that the mean ratio is \( \mu_\rho = (\pi/2)\omega \), but the
STD does not exist (infinite) due to the slow convergence in
the upper tail. The 90% CDF occurs when \( \rho = 3\omega \) and the
99% when \( \rho = 10\omega \).

While the above analysis is based on the NLOS perform-
earance, the design of the epoch tracking function must be
based on the ideal LOS case where by definition the tracking
error is zero. Thus, the ratio set-point \( \rho_0 = (t_1/t_2) \) \( = ((t_2 - \Delta)/t_2) \) defines the two set-point locations \( (t_1 \) and \( t_2) \)
on the leading edge, where \( \Delta \) is the separation between the two
points in time. The set-point should be chosen such that the
location of the first point is not too close to the noise threshold,
and the second point is not too close to the peak where the
effects of multipath interference are greatest. Typical values of
\( \rho_0 \) [27 Sec. 4.6] are in the range 0.3–0.5 with a separation
of \( \Delta = 0.4 \).

For the NLOS it would be advantageous if the mean
ratio \( \bar{\rho} \) was the set-point ratio \( \rho_0 \), which would result in
an unbiased estimate of \( \rho \). Applying (7) to obtain the
values of the Rayleigh parameters, the requirement for this
condition is

\[
\rho_0 = \bar{\rho} = \frac{\pi}{2}\omega = \frac{\pi}{2}\frac{\sigma_1}{\sigma_2} = \frac{\pi}{2} \left( \frac{t_1}{t_2} \right)^{1.5} = \frac{\pi}{2} \rho_0^{1.5} \tag{A-6}
\]

which has the solution \( \rho_0 = 4/\pi^2 = 0.405 \), and is in the
desirable range described above.

REFERENCES

wideband precision asset location system,” in Proc. IEEE Conf. UWB


Ian Sharp is a Senior Consultant on wireless positioning systems. He has more than 30 years of experience in engineering research in radio systems. His initial involvement in positioning technology was in aviation, and with the Interscan microwave landing system in 1980. From 1980 to 1990, he was the Research and Development Manager for the Quiktrak covert vehicle tracking system. This system is now commercially operating worldwide. From 1990 to 2007, he was with Commonwealth Scientific and Industrial Research Organisation (CSIRO), Melbourne, VIC, Australia, mainly on developing experimental radio systems. He was the Inventor and Architect Designer with CSIROs precision location system (PLS) for sports applications. The PLS has been successfully trailed in Australia and the USA. He holds a number of patents relating to positioning technology. He is a co-author of the book Ground-Based Wireless Positioning (Wiley and IEEE Press, 2009).

Kegen Yu (SM’12) received the Ph.D. degree in electrical engineering from the University of Sydney, Sydney, NSW, Australia, in 2003. He is currently a Professor with the School of Geodesy and Geomatics, Wuhan University, Wuhan, China. He has worked for Jiangxi Geological and Mineral Bureau, Nanchang, China, Department of Electrical Engineering at Nanchang University, CWC at the University of Oulu, CSIRO ICT Centre, Department of Electronic Engineering at Macquarie University, and School of Surveying and Geospatial Engineering (now integrated within the School of Civil and Environmental Engineering) and the Australian Centre for Space Engineering Research within the School of Electrical Engineering and Telecommunications, University of New South Wales. He has been an Adjunct Professor with Macquarie University since 2011.

Dr. Yu is currently on the editorial boards of the EURASIP Journal on Advances in Signal Processing, the IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS, and the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. He is the Lead Guest Editor for a Special Issue of Physical Communication on Navigation and Tracking and for a Special Issue of the EURASIP Journal on Advances in Signal Processing on GNSS Remote Sensing.