Peak and leading edge detection for time-of-arrival estimation in band-limited positioning systems

I. Sharp  K. Yu  Y.J. Guo
Wireless Technologies Laboratory, CSIRO ICT Centre, Australia
E-mail: kegen.yu@mq.edu.au

Abstract: The performance of the peak and leading edge detection methods for time-of-arrival (TOA) estimation in band-limited systems is examined. Analytical expressions for the detection performance in the presence of both random noise and multipath interference are derived. A dimensionless performance factor is presented that allows simple comparisons of the TOA estimation algorithms. These equations allow the performance tradeoff analysis to be undertaken without the need for simulations. It is shown that the leading edge detection method has significantly better multipath mitigation characteristics than the peak detection one, but at the expense of inferior noise performance.

1 Introduction

The global positioning system (GPS) has been widely used in many military applications and civilian services. In a variety of circumstances, however, GPS is not an option for wireless positioning and tracking because of adverse environmental conditions, cost, security or tracking accuracy, so that a local area wireless positioning system is required. One of the earliest non-GPS-based wireless positioning systems was reported for locating objects in a construction area [1]. It is expected that many future applications including sports, security and emergency services will demand such systems more. The design of non-GPS local area wireless positioning systems is rather different from wide-area systems such as GPS, as the applications are typically indoors where multipath propagation conditions predominate. Since the positional accuracy of such systems is directly related to the bandwidth, wideband techniques such as ultra-wideband (UWB) seem appropriate, but the limitations in allowable transmitter power restrict the UWB operating range to typically less than 10 m. An alternative approach is to use 2.4 and 5.8 GHz ISM bands, which allow moderate bandwidth and transmitter power, thus resulting in longer ranges, but with reduced positional accuracy. Since these unlicensed ISM bands are restricted to spread-spectrum signals to minimise mutual interference from simultaneous users, the design of practical positioning systems requires tools for assessing spread-spectrum time-of-arrival (TOA) estimation performance. This paper provides the theory for evaluating the performance of peak and leading edge detection algorithms in band-limited direct-sequence spread-spectrum systems.

The design and performance analysis of TOA-based positioning systems can be found in the literature [2–17]. One of the key challenges is to develop a practical and accurate TOA estimation scheme. Although there are many TOA estimation techniques, peak detection and leading edge detection are the most common low-complexity practical methods [18–27]. Although it is possible to evaluate the performance of various algorithms through simulation, this does not allow comparative performance to be determined easily. The objective of this paper is to develop a set of concise analytical equations for the performance analysis of the peak and leading edge detection algorithms. To the best of our knowledge, these closed-form analytical expressions do not exist in the literature. It is shown that, in a multipath hostile environment, particularly under non-line-of-sight propagation conditions, the leading edge detection algorithm outperforms the peak detection significantly. On the other hand, in line-of-sight conditions...
2 Overview of radio positioning

The determination of position using radio technology is particularly challenging because of complex radio propagation characteristics. Two main problems include excess losses and excess delays. The excess propagation loss constrains the propagation range, particularly for small low-powered devices such as the ones used in wireless sensor networks. The position determination process in two dimensions requires a minimum of three assisting nodes, and preferably five or more for redundancy. The signal-to-noise ratio (SNR) at the receiver of a distant node can be very low. As a consequence, the analysis of the receiver TOA estimation performance in the presence of noise is important, since the long-range measurements will be SNR limited. Further, the excess delays because of scattering result in positional inaccuracy. To minimise these effects, it is clear that the receiver should attempt to detect the first signal above the noise floor on the leading edge of the signal impulse response. If a direct-sequence spread-spectrum technique is used, the TOA algorithm should be based on detection of the leading edge of the correlogram.

Because of the correlation properties of direct-sequence spread-spectrum signals, the ability for the detector to reject multipath signals is directly related to the chip period of the modulating signal. The chip rate should be designed as high as possible, within the bandwidth limitations imposed by the spectral regulations and the radio hardware. Therefore UWB propagation appears particularly attractive for indoor applications in terms of the time resolution, but the required restrictions on the radiated power and the high signal processing requirements mean that the UWB technology has a rather small coverage area per node. The alternative is to use the 2.4 or 5.8 GHz ISM bands, which have much narrower bandwidths of 80 and 125 MHz in the USA respectively [28]. These narrower bandwidths result in lower chip rates and lower consequential accuracy, but the higher permissible transmitter power leads to longer ranges and more nodes contributing to the position determination. The increased number of nodes results in more redundancy, which in turn improves the position fix accuracy by allowing 'bad' TOA data to be rejected. Using more nodes also reduce the geometric dilution of precision, again improving the accuracy. Clearly, this alternative using the ISM bands provides a compromise between coverage and accuracy. However, with a maximum bandwidth of 125 MHz, the typical chip rate will be limited to about 60 Mchips/s, which has a nominal multipath delay rejection period of about 17 ns (or about 5 m); these band-limited signals cannot resolve the fine structure of indoor multipath delays. Further, for simple chip radios [29], the bandwidths are typically limited to 20–40 MHz to reduce complexity and cost, so the signals are often even more band limited. Thus, the TOA performance of band-limited spread-spectrum signals is of fundamental importance.

Indoor propagation of radio signals produces scattering of signals from obstacles along the path, so that the received signal will have a spread of delays relative to the direct path. For accurate position determination, it is essential to detect a signal as close to the direct path as possible. Therefore for accurate position determination in indoor environments, a leading edge detection algorithm is highly desirable. The basic binary direct-sequence signal has an infinite bandwidth, so the band limiting of the transmitted signal will result in some distortions in the receiver output correlogram. In particular, the band-limiting produces an excess delay of about 17 ns (or about 5 m); these band-limited signals cannot resolve the fine structure of indoor multipath delays. Further, because the receiver output baseband signal is band limited, the number of independent samples in the correlogram is also limited. For example, if the spread-spectrum signal bandwidth is limited to $\pm 1/T_c$ where $T_c$ is the chip period, then the Nyquist sampling rate is two samples per chip. As the band-limited correlogram leading edge has a rise time of about one chip, the number of samples on the leading edge is limited to two. Although these samples can be interpolated, this procedure does not generate any new information. Consequently, the algorithms that determine the TOA should be limited to use only two samples. This limitation also reduces the amount of signal processing required, an important consideration for simple low-powered nodes.

The definition of TOA for position determination needs further explanation. In traditional systems where the peak of the correlogram is used to define the TOA, the signal processing results in the time alignment of the local code with that in the received signal to synchronise the two codes. In the case of a leading edge algorithm there is no obvious time reference, so the code synchronisation cannot be achieved. However, such synchronisation is not essential for position determination. In particular, the TOA relates to, but does not measure, the propagation time from the transmitter to the receiver. This is because the measured TOA includes delays in the transmitter and the receiver.
and an unknown time offset in the clock in the transmitter. Generally speaking, these ‘nuisance’ parameters are unknown to the receiver, so the TOA at the receiver output results in measurements of pseudo-range rather than range. A pseudo-range is the true range plus an offset which is common to all receiving nodes. Therefore if any displacement from the peak to the leading edge is common to all receivers, this displacement can be arbitrarily chosen to suit the leading edge algorithm, without any need to reference the peak of the correlogram. Accordingly, an algorithm can choose any convenient time reference, but note that this must be common to all the nodes measuring the TOA of the transmissions.

3 Band-limited correlogram characteristics

Because the spread-spectrum signal is band limited in all real systems, it is important to determine the band-limiting effect on the receiver output correlogram. While actual systems constrain the bandwidth of operation with appropriate analogue filters, the technique used in the following analysis is to derive the spectrum of the infinite bandwidth correlogram, and then null the spectrum above the required bandwidth. Although this technique is not based on realisable physical filters, the analysis captures the essential features of band limiting. Further, typical implementations are often based on the digital generation of the $pn$-code where this method of spectral nulling can be applied, so the analysis technique is appropriate. Given a signal $s(t)$, the autocorrelation function (or correlogram) is given by

$$c(\tau) = \int s(t) s(t + \tau) \, dt$$

where $F^{-1}\{ s(f) \}$ is the inverse Fourier transform, and $s(f)$ is the Fourier transform of the signal $s(t)$. For the case of a band-limited $pn$-code, the signal spectrum can be expressed as

$$|S(f)|^2 = |PN(f)|^2 \Pi_{2B}(f)$$

where $\Pi_{2B}(f)$ is a rectangular weighting function of bandwidth $\pm B$ and $PN(f)$ is the spectrum of the infinite bandwidth $pn$-code. By taking the inverse Fourier transform on both sides of (2), the band-limited correlogram is given by

$$c(\tau) = \hat{c}_0(\tau) \ast \left[ 2B \text{sinc}(2B\tau) \right]$$

where $\hat{c}_0(\tau)$ is the infinite bandwidth correlogram, and ‘$\ast$’ represents the convolution operation. The nominal infinite bandwidth correlogram is triangular in shape, with unit amplitude and a width of $\pm 1$ chip, so it is convenient to normalise the infinite bandwidth correlogram to unit amplitude, and the correlation time to the chip period.

The convolution can then be expressed as

$$c(\tau) = 2B \int_{-1}^{1} \hat{c}_0(\tau) \text{sinc}[2B(\tau - \tau)] \, d\tau$$

where $\beta$ is the bandwidth normalised to the chip rate, so that $\beta = B/T_0$ where $T_0$ is the chip period. For example, typically $\beta = 1$ for spread-spectrum systems which limit the signal to the main lobe. For GPS C/A receivers, typically $\beta = 10$. Note that in this context the ‘bandwidth’ is defined at the $-3$ dB points, while the actual RF channel is defined over a wider band. Commonly, the radio signal is limited at the first spectral null of the ideal infinite bandwidth signal, so the radio bandwidth is twice the bandwidth defined by the $\beta$ parameter.

The effect of band limiting can be observed from Fig. 1. Firstly, the sharp intersections of the straight-line segments associated with the triangular shape are smoothed out, with the shape becoming progressively smoother as the bandwidth decreases. In particular, the peak of the correlogram becomes a broad smooth curve with zero derivative at the peak. Thus, detecting the peak by the derivative is at least inaccurate for the band-limited case. In fact, as the accuracy of determining a position on any curve is proportional to the derivative of the function, band-limiting results in poor performance of peak detection methods, especially as the bandwidth becomes increasingly limited. Secondly, the amplitude progressively decreases as the bandwidth is reduced. Finally, the shape of the correlogram beyond $\pm 1$ chip becomes distorted, with the exact shape dependent on the exact nature of the filtering. For the case where $\beta = 1$, the magnitude of the sidelobes is small, and for $\beta = 0.75$ the sidelobes are only slightly larger. However, when $\beta = 0.5$ the sidelobes are considerable. The presence of sidelobes when peak detection algorithms are used is of little importance, but when leading edge algorithms are used the presence of sidelobes has a significant impact on the performance.

![Figure 1 Effect of bandlimiting the correlogram](image)

Note that the $B$ parameter in the figure is the $\beta$ parameter defined in the above text, and that the magnitude of the correlogram is plotted. The nominal correlogram shape is triangular, of unit amplitude and a width of $\pm 1$ chip.
The β parameter has a significant effect on the performance of the TOA estimation. Generally speaking, it should be selected to be as low as possible, since this maximises the chip rate for a given radio bandwidth. However, as discussed above, if the β parameter is made too small the distortions in the correlogram result in performance degradation, particularly in a multipath environment. In this paper, the numerical analysis is based on a conservative choice of β = 1.

By noting that the peak of the correlation corresponds to the signal amplitude, the band-limited peak amplitude can be expressed as

\[ P(\beta) = 2 \int_0^\beta \sin^2 x \, dx \]  

Another useful parameter that will be used later is the reduction in the relative amplitude, which is defined as

\[ \delta = 1 - P(\beta) = 1 - 2 \int_0^\beta \sin^2 x \, dx \]  

It will be shown that the multipath errors are closely related to this parameter.

4 Peak detection algorithm performance

Although the peak detection algorithm with band-limited spread-spectrum signal is not well suited for heavily multipath environments, it has superior performance in Gaussian channels. This section presents the analytical models of the peak tracking algorithm, which will be used for later comparison with leading edge detection algorithms.

4.1 Simplified correlogram models

To ease the analytical analysis, we first introduce two analytical models, the hyperbolic model and the Gaussian model, to approximate the complex correlogram presented in Section 3. The main reason for selecting the hyperbolic and Gaussian models lies in that the correlogram as shown in Fig. 1 resembles well with the shapes of the two models, particularly around the peak. By choosing the model parameters properly, the two models can have a good match with the correlogram.

With correlation time normalised to the chip period, the correlation function can be approximated by the hyperbolic function as

\[ H(\tau) = A \left[ 1 - \sqrt{\left( \frac{\tau - \Delta\tau}{\tau_0} \right)^2 + \delta^2} \right] \]  

where \( \tau_0 \) is the chip period, A and Δ\( \tau \) are two other parameters to be determined from the least-squares fitting procedure and \( \delta \) is defined in (6). Note that with a zero offset peak, that is, \( \Delta\tau = 0 \), the magnitude of the correlogram at \( \tau = 0 \) is \( A[1 - \delta] \). On the other hand, when \( \tau \) is far from the peak the function becomes \( A[1 \pm \tau/\tau_0] \), or the triangular shape.

The Gaussian model is defined by

\[ G(\tau) = A(1 - \delta) e^{-(\tau-\Delta\tau/a)^2} \]  

where a is the Gaussian shape parameter which is related to the chip period, and for a given band-limiting parameter β it can be determined by a least-squares fit, A and Δ\( \tau \) are two other parameters to be determined from the least-squares fitting procedure, and \( \delta \) is the known amplitude reduction. Fig. 2 shows a comparison between the ‘exact’ correlogram shape and the two models for the bandwidth parameter \( \beta = 1 \). In general the match is good in both cases, but the Gaussian fit is better.

4.2 Performance of peak tracking with random noise

The peak tracking algorithm is based on calculating the correlation at two points (termed \( P_1 \) and \( P_2 \)), one leading and one lagging relative to the peak by \( \tau \) chips. If the sampling clock in the receiver is misaligned by \( \Delta\tau \) from the nominal zero position, then the difference between the correlogram amplitudes at these two points when using the hyperbolic model is

\[ Z_1 - Z_2 = A \left[ 1 - \sqrt{\left( \frac{\tau + \Delta\tau}{\tau_0} \right)^2 + \delta^2} \right] - A \left[ 1 - \sqrt{\left( \frac{-\tau + \Delta\tau}{\tau_0} \right)^2 + \delta^2} \right] \]  

If the sampling points are chosen such that they are well

![Figure 2 Comparison between model and exact correlogram for bandwidth parameter \( \beta = 1 \)](image)
It can be seen that the differential amplitude is approximately linearly related to the timing misalignment, and this fact can be used to correct the misalignment. The differential signal can be used in a feedback loop to maintain tracking with a nominal zero displacement.

When the amplitudes at \( P_1 \) and \( P_2 \) are corrupted with noise \( n_1 \) and \( n_2 \), respectively, the differential amplitude becomes

\[
Z_1 - Z_2 = -2A \left[ \frac{(\Delta t/\tau_0)}{\sqrt{(\Delta t/\tau_0)^2 + \delta^2}} \right] (\Delta t/\tau_0)^2
\]

\[
\simeq -2A (\Delta t/\tau_0)
\]

The tracking loop tries to null the differential amplitude, so that the normalised tracking error is given by

\[
v = \frac{\Delta t}{\tau_0} = \frac{1}{2} \left( \frac{n_1 - n_2}{A} \right)
\]

Since the noise is assumed to have zero mean, the mean tracking error will be zero, whereas the variance of the tracking error is

\[
\text{var}(v) = E[v^2] = \frac{1}{4} E[n_1^2 + n_2^2 - 2n_1n_2] / A^2
\]

Accordingly, the standard deviation (STD) in the tracking error can be written as

\[
\sigma_v = \frac{1}{\sqrt{2}} \sqrt{1 - \rho} \text{ chips}
\]

where \( \gamma = A^2 / E[n^2] \) is the SNR and \( \rho \) is the correlation coefficient between the two noise components. The correlation coefficient can be estimated based on the properties of the receiver correlator. If noise samples are separated in time by more than about the reciprocal of the bandwidth of the signal, they will be largely uncorrelated (\( \rho = 0 \)). That is, for the case where the bandwidth is \( 2/\tau_0 \), the noise will be uncorrelated if the samples are separated by half a chip or more. Therefore if the points \( P_1 \) and \( P_2 \) are at \( \pm 0.25 \) chip or greater relative to the peak the STD of the tracking error is

\[
\sigma_v = K \sqrt{\frac{1}{\gamma}}
\]

where \( K = 1/\sqrt{2} \). This form of STD of the tracking error is general to all tracking algorithms, but the constant \( K \) is particular to each algorithm. As a result, the noise tracking performance of various algorithms can be compared by the associated algorithm constant \( K \).

### 4.3 Performance of peak detection in a multipath channel

The analysis of the multipath performance requires a model of the multipath interference signal. While this signal will in practice be complex consisting of many interference sources, a common approach is to calculate the tracking error because of a single in-phase interfering source. Such simplified calculations only provide a rough guide to the real situation, but it does provide a method for comparing various tracking algorithms in the presence of multipath signals.

The multipath interference is defined by two parameters, namely the relative amplitude \( (m) \) and the multipath delay \( (\tau_m) \). As the amplitude is defined relative to the direct signal, without loss of generality, the direct signal amplitude can be set to unity. For convenience of analysis, the shape near the peak will be approximated by a Gaussian function; the hyperbolic approximation is not used here because of the complexity of the expressions. The combined correlogram is then defined by

\[
e(t) = e^{-\left(t/\sigma^2\right)} + me^{-\left(t-\tau_m/\sigma^2\right)}
\]

The peak is determined by differentiating (16) and equating it to zero, resulting in

\[
t = m(\tau_m - t)e^{2\tau_m - \tau_m^2 / \sigma^2}
\]

where the solution for \( t \) determines the location of the peak of the correlogram relative to the peak without the multipath interference, which appears at \( t = 0 \). This method can be used for both numerical calculations as well as deriving approximate analytical expressions. As no closed-form analytical solution to (17) is available, an iterative approach is taken. Because (17) can be expressed in the form \( t_{n+1} = f(t_n) \), a series of analytical solutions can be derived, with the initial solution termed the zeroth order, the next the first order and so on.

An approximate analytical solution to (17) can be determined for the case of small multipath delay \( (\tau_m \ll \sigma) \). First, the equation can be expressed in a
normalised form as
\[ t = M(ax - t) e^{2ax/a} \]  
(18)

where
\[ M = me^{-x^2}, \quad x = \frac{\tau_m}{a} \]

In the case where \( x \) and hence \( t \) are small, the exponential function can be approximated by a series expansion, so the equation becomes
\[ t = M \left[ ax + (2x^2 - 1)t + \frac{2}{a} x(x^2 - 1)t^2 \right] \]  
(19)

where only the first three terms of the exponential function have been included and the resulting third-order term of \( t \) is dropped. When \( x \) is small (or when \( x \to 1 \)) the magnitude of the \( t^2 \) term is much smaller than that of the \( t \) term. However, when \( x^2 \to 0.5 \), the magnitude of the \( t^2 \) term is much greater than that of the \( t \) term. Therefore provided that \( x^2 \ll 0.5 \), one can ignore the second-order term of \( t \), which yields
\[ t_0 = \frac{amx}{m(1 - 2x^2) + e^2} \]  
(20)

For small \( x \), that is, \( \tau_m << a \), the solution can be further approximated by
\[ t_0 \approx \frac{m(ax)}{m + 1} = \frac{m}{m + 1} \tau_m \]  
(21)

It is clear that the initial slope of the error function is \( m/\left(m + 1\right) \) with respect to the multipath delay. For large \( x \), that is, \( \tau_m \gg a \) (even though this violates the original assumption that \( x \ll 0.5 \)), (20) becomes
\[ t_0 \approx \frac{m(ax)}{e^2} = m\tau_m e^{-(\tau_m/a)^2} \]  
(22)

Although the solution defined by (22) was derived originally based on the assumption of small \( x \), it can be shown that the same expression is produced from (17) under the assumption of a large \( x \). Also, this is the zeroth-order solution of (17) by setting \( t \) on the right-hand side to zero. Therefore the zeroth-order solution given by (22) is useful over wider ranges of multipath parameters than originally assumed in the above derivation. Fig. 3 shows the analytical zeroth- and first-order solutions of (17), with the first-order solution given by
\[ t_1 = m(\tau_m - t_0) e^{2\tau_m t_0 - \tau_m^2/a^2} \]

Finally, it can be shown that when \( \tau_m = a/\sqrt{2} \), the peak estimation error by (22) is maximum, given by
\[ e_{\max} = \frac{ma}{\sqrt{2e}} \]  
(23)

For the particular case considered previously where bandwidth parameter \( \beta = 1 \) and \( a = 0.634 \) chips, the maximum tracking error (in chips) is 0.27 \( m \).

Although the peak detection algorithm is reasonably satisfactory in a mildly multipath environment, the performance deteriorates rapidly as the multipath signals become more severe. For example, if the multipath delay exceeds the chip period, the peak signal can be in error by more than one chip from the leading edge. Therefore for better performance in multipath environments a leading edge algorithm is necessary. Nevertheless, the above expressions for the performance of a peak detection algorithm can be used as a reference for the leading edge algorithm discussed in the following section.

5 Leading edge algorithms

For more accurate estimation of the TOA, particularly when the multipath is severe, some form of a leading edge algorithm is necessary. As outlined in the discussion in Section 2, the band limiting of the spread-spectrum signal typically results in just two independent samples on the leading edge of the correlogram. As the time position of the leading edge is a function of its shape and is not related to the correlogram amplitude, simple methods such as defining when the correlogram signal crosses a threshold level are not satisfactory. Such a threshold typically would...
be a function of the SNR. Note that the position determination requires the measurement of a pseudo-range with a constant offset in all receiving nodes that have different SNR. Therefore such a simple threshold strategy results in poor performance because of the variations in the definition of the pseudo-range in each node. Also note that the leading edge of the band-limited correlogram has a skirt whose derivative is small, resulting in large variations in the threshold point as the SNR varies. A better strategy is to use a method which is independent of the correlogram amplitude, and is solely a function of the shape of the correlogram. As there are only two independent samples on the leading edge, the algorithm should be based on using two samples spaced approximately half a chip apart, so that the samples are largely independent. Further, as the susceptibility to multipath effects increases as the position on the correlogram increases in correlation time, these two sample points should be as close to the leading edge of the correlogram as possible. Additionally, the SNR and the leading edge skirts of the band-limited correlogram (see Fig. 1) also places limits on the lowest position of the leading point. However, within these constraints, any shape-based algorithm would be appropriate, although some algorithms may be superior to others. In this section, the performance of one such algorithm is analysed, both in the presence of random noise and a single multipath signal.

The leading edge algorithm to be analysed is based on estimating the position of the zero amplitude of the correlogram by linearly projecting the two points \( P_1 \) and \( P_2 \) on the leading edge. As shown in Fig. 4, the first point (\( P_1 \)) is at the leading edge at an amplitude defined by a threshold parameter \( \alpha \), with \( \alpha A \) being the threshold amplitude and \( A \) being the correlogram amplitude. The \( \alpha \) parameter is selected so that the \( P_1 \) point is sufficiently out of the noise, and above the curvature distortion associated with band limiting. A typical value is \( \alpha = 0.15 \), but this point can be adjusted for the desired design limiting SNR. The second point \( P_2 \) is defined by a separation parameter \( \tau \) from point \( P_1 \); typical values are in the range of 0.25–0.5 chips. Small separation values reduce the effects of multipath, but increase the susceptibility to noise. The basis of the algorithm is then to define the location of the signal epoch by projecting \( P_2 \) through \( P_1 \) until it intercepts the \( \tau \)-axis of the correlogram.

The intersection of this straight line with the \( \tau \)-axis is defined as the signal epoch for this algorithm so that the TOA of the signal is one chip period greater than the signal epoch. The epoch estimation error will be the same as the TOA estimation error. The performance of the algorithm with both random noise and multipath interference is analysed in the following sections.

### 5.1 Random noise performance of the projection algorithm

The analysis under random noise conditions assumes a simplified theoretical correlogram, namely a triangular shape for the leading edge with a rise time of one chip. As shown in Section 3, the effect of band limiting is to round off the sharp intersections of the triangular shape, but away from these areas the basic triangular shape remains largely unaltered. As the two points are located within this minimally distorted region, the assumption of using the ideal triangular shape has minimal effect on the derived performance.

The following analysis assumes that the correlogram is corrupted by random noise, with zero mean and known noise power. The noise at the two points \( P_1 \) and \( P_2 \) are not assumed to be statistically independent, but the correlation coefficient is assumed to be a function of the separation \( \tau \). This assumption means that potentially the separation time \( \tau \) can be chosen less than half a chip apart, in which case the noise will be partially correlated. The simplified geometry of the algorithm with the noise corruption is shown in Fig. 4. Both points are corrupted by noise \( n_1 \) and \( n_2 \), respectively, which are assumed to have identical statistics with zero mean and variance \( \sigma_n^2 \). The epoch is estimated by extrapolating the straight line joining these noise corrupted points to intersect the \( \tau \)-axis. The error in the estimated epoch is then \( \epsilon \) chips.

The first step in the analysis is to obtain an expression for the epoch error in terms of the algorithm parameters \((\alpha, \tau)\) and the SNR \( (\gamma) \), defined as \( A^2/\sigma_n^2 \). Simple geometric calculations show that the epoch error is given by

\[
\epsilon = \alpha - \tau \left( \frac{\alpha A + n_1}{\alpha A + n_2 - n_1} \right)
\]

\[
= - \alpha \left( \frac{\tau (\Delta n/A)}{1 - 1/\tau (\Delta n/A)} \right)
\]

where \( \Delta n = n_1 - n_2 \). From the denominator, it can be observed clearly that the error has a singularity when \( \Delta n/A = \tau \). Thus, for the algorithm to provide reasonably accurate results, the SNR must obey the condition \( \gamma > 2/\tau^2 \), otherwise large epoch measurement errors can occur when the noise is sufficiently large so that the singularity condition occurs (or nearly occurs). As typically \( \tau \) is about half a chip, this condition implies that the SNR should be much greater than 8 (or 9 dB); as the typical SNR will exceed 20 dB, this restriction is not a practical
problem. However, reducing $\tau$ to (say) 0.1 chips means the unsatisfactory requirement that the SNR must be much greater than 23 dB. Further increasing the parameter $\tau$ beyond about 0.7 is impossible, as the pulse width is one chip. Thus, the practical range for $\tau$ is about 0.3–0.7 chips.

Providing the SNR is sufficiently large, (24) can be simplified by using the first few terms of the series expansion, resulting in

$$e \approx -\left[\frac{n_1}{A} + \frac{\alpha}{\tau} \left(\frac{\Delta n}{A}\right) + \frac{1}{\tau} \left(\frac{n_1}{A}\right) \left(\frac{\Delta n}{A}\right)\right]$$

Taking the expectation on both sides, the mean epoch error is obtained as

$$E[e] \approx -\frac{1}{\tau} \left(\frac{\sigma_n^2(1-\rho)}{\gamma A^2}\right)$$

$$= -\frac{1-\rho}{\tau \gamma}$$

(26)

where $\rho$ is the correlation coefficient between $n_1$ and $n_2$. Thus, the algorithm results in a biased estimate of the epoch location provided that the correlation coefficient is not unity. For an SNR of say 20 dB and a spacing parameter $\tau$ of half a chip, the worst-case mean error is 0.04 chips, which is small but significant. However, as the SNR can be measured, this bias error can be corrected if needed. For most practical situations, the mean error can be considered zero. Then, the variance of the epoch error is given by

$$E[e^2] \approx \frac{1}{\gamma} \left[1 + 2\left(\frac{\alpha}{\tau}\right) + 2\left(\frac{\alpha}{\tau}\right)^2\right]$$

(27)

where higher order terms of the SNR have been dropped. Typically, if $\alpha = 0.15$ and $\tau = 0.5$ chips, $\text{var}(e) = 1.78/\gamma$.

For comparison, the corresponding expression for the peak parameter is $\text{var}(e) = 0.5/\gamma$. Thus the leading edge algorithm has worst-noise performance, but the aim of the leading edge algorithm is to improve the multipath performance rather than noise performance. From (27), it can be observed that the requirement to minimise the error in the epoch estimate in the presence of noise is to make the threshold level $\alpha$ as small as possible, and the separation of the two points $\tau$ as large as possible; these constraints will minimise both the bias error and the variance of the error. In practice, however, the STD of the epoch error is only weakly dependent on these parameters.

5.2 Multipath performance of the projection algorithm

The analysis of the multipath effect is again based on the simplified triangular correlogram used for the noise analysis under the assumption of a single multipath interference path. Further, the phase of the multipath signal is assumed to be the same as the direct signal. The multipath signal is assumed to be $m$ times the amplitude of the direct signal ($m \leq 1$). Although it is recognised that real multipath conditions will involve many interfering signals, the single-interferer theoretical analysis provides the basis for the analysis of the superposition of multiple interferers. However, such analysis is not pursued in the paper. For design purposes the aim is to obtain the single-multipath performance, and this can be a good guide to the case of multiple multipath interference sources.

The geometry of the multipath correlogram is shown in Fig. 5 for the case where the multipath delay is less than the threshold delay. The multipath signal is delayed by $\tau_m$ chips relative to the direct signal. The direct signal has amplitude $A$ and the multipath signal has an amplitude $mA$. The epoch is determined by the extrapolation of the line joining points $P_2$ and $P_1$ to the time axis.

The epoch error because of the multipath signals $\Delta A_1$ and $\Delta A_2$, respectively, at points $P_1$ and $P_2$ is given by

$$e = \left(\frac{A_1}{A_2 - A_1} - \frac{A_1 + \Delta A_1}{A_2 + \Delta A_2} - \frac{A_1 + \Delta A_2}{A_1 + \Delta A_1}\right)\tau$$

(28)

Three cases can be identified, namely when the multipath delay is less than the threshold delay, when the multipath delay is between the delays of point $P_1$ and $P_2$, and when the delay is greater than the delay of point 2. In the last case, there is clearly no corruption of the data at the measurement points, so that there is no multipath-related epoch measurement error. For the case where the multipath delay is less than the threshold delay, it can be shown that (28) becomes

$$e = \frac{m}{m+1} \tau_m$$

(29)

If the multipath ratio parameter $m$ is small, the measurement error in chips is approximately $m$ times the multipath delay. Note that in this case the algorithm parameters, such as the threshold delay and the separation delay, do not affect the measurement error; the error is solely a function of the multipath signal characteristics including delay and multipath ratio. If (29) is compared with the corresponding error expression for the peak

![Figure 5 Multipath delay ($\tau_m$) less than the threshold delay ($\tau$)]
The threshold delay $t_r$ is 0.15 chips, and the separation parameter $t$ is 0.25 chips. Note that the errors are constrained to multipath delays of the measurement points, the epoch measurement error of (28) can be evaluated to give

$$e = \frac{mt_r t_i}{\tau + mt_r} \quad (\tau_i = \tau + \tau_m > 0) \quad (30)$$

Therefore the delay is again proportional to the multipath ratio $m$. However, the error is also proportional to the threshold delay $t_r$ and the multipath delay relative to $P_2$ ($\tau_i$). Note that $\tau_i$ is constrained to be less than the algorithm parameter $t$, so again the maximum error is constrained by a parameter of the algorithm, rather than the multipath delay.

Fig. 6 shows the epoch error as a function of the multipath delay. Observe that the error is maximum when the multipath delay is equal to the threshold delay ($\tau_m = \tau_r$). In particular

$$e \leq \frac{m}{m+1} \tau_i \simeq m\tau_i \quad \text{(31)}$$

Again comparing (31) with the corresponding expression for the peak algorithm (21), it is seen that the maximum error is proportional to the multipath ratio $m$, but the maximum error for the leading edge algorithm is also constrained by an algorithm parameter, whereas for the peak algorithm the maximum error is determined only by the shape of the correlogram and hence the bandwidth. Thus, unlike the peak algorithm where the epoch errors because of multipath signals are related directly to the multipath signal parameters including the multipath delay and multipath ratio, the leading edge projection algorithm is more constrained, so that the maximum error is essentially given by the threshold delay parameter.

To minimise the epoch error for a given multipath signal, the threshold delay should be as small as possible. In practice, the threshold value is related to the SNR and the signal bandwidth. With band limiting of the signal, the correlogram near the threshold curves significantly making the projection measurements inaccurate. Further, if the threshold is set too low, noise causes ‘false’ peaks resulting in large epoch errors. Therefore in practice the threshold is typically limited such that $\tau_i \geq 0.15$ for the band-limiting parameter $\beta = 1$. For other bandwidths the threshold must be adjusted accordingly. For example, if the bandwidth parameter is reduced to $\beta = 0.75$ (the minimum acceptable value), the threshold should be increased to 0.2 as the band-limited correlogram has significant sidelobes which can be confused for a (small) direct signal if the threshold is not increased.

While the leading edge algorithm is not intended for a wide bandwidth, the parameters $\beta = 10, \tau = 0.25$ and $\alpha = 0.05$ are appropriate for the GPS in C/A mode, so it is interesting to compare the leading edge algorithm performance with a typical GPS multipath mitigation technique. With the above parameters, the nominal peak error is 0.017 chips, and the extent of the errors is limited to about 0.3 chips of multipath delay. For comparison, a 'narrow correlator' GPS receiver [30] has a peak error of about 0.03 chips, but this error is spread over multipath delays up to 1 chip. As the leading edge algorithm has both lower peak errors and is more constrained in the range of multipath delays which cause errors, the leading edge algorithm described in this paper has superior performance with parameters appropriate to GPS.

Now consider a numerical example for comparison with the performance of the peak tracking algorithm. With a conservative threshold delay of 0.15 chips (or $\alpha = 0.15$) and a multipath parameter $m = 0.5$, the peak error is 0.05 chips. For comparison, the corresponding peak algorithm error is 0.135 chips, or 2.7 times greater. Thus, the leading edge algorithm has much reduced errors because of multipath but, as shown previously, the noise errors will be 1.9 times greater. Thus, there is a tradeoff of worst-noise performance for improved multipath performance.

6 Performance summary of the algorithms

This section provides a concise summary of the respective performances of the algorithms, thus allowing easy comparison. The basis of the analysis is a band-limited spread-spectrum signal, which after correlation results in a band-limited correlogram. With typical band limiting of a baseband bandwidth of twice the chip rate (defined by the parameter $\beta = 1$), the correlogram can be completely...
defined by sampling points half a chip apart, although interpolation is possible. Therefore in all the algorithms, the number of points is limited to two.

The multipath interference performance of the algorithms is based on the simple case of a single multipath interference of relative amplitude \( m \) and delay \( \tau_m \). In this case there is no simple formula to summarise the multipath performance, except when the multipath delay is small all the algorithms exhibit a tracking error of the form \( e = (m/(m+1))\tau_m \). However, the maximum error characteristics of the peak and leading edge algorithm are very different. For the peak tracking algorithm (with a bandwidth parameter \( \beta = 1 \)) the maximum error is 0.27\( m \) chips, essentially independent of the only algorithm parameter (the separation of the sampling points). However, the leading edge algorithm has the characteristics that the maximum error is limited by the parameters to an error given by \( e = m\tau_m/\tau + m\tau_m \leq (m/(m+1))\tau_m \), so that the algorithm threshold parameter \( \tau_m \) can be chosen to significantly restrict the maximum errors. Although in theory choosing a very small threshold parameter can essentially eliminate the multipath errors, the effect of band limiting and receiver noise places a minimum usable limit on the threshold parameter. Nevertheless, with a conservative limit of \( \tau_m = 0.15 \), the maximum error with \( m = 0.5 \) is 0.05 chips, compared with 0.135 chips for the peak algorithm. Thus, the leading edge algorithm is about three times better than the peak algorithm when the signal is band limited.

The performance with random noise is given in terms of the mean and STD of the tracking error. In the case of the mean error, it is shown that for all the algorithms the mean is essentially unbiased (zero mean error). In the case of the STD, all the algorithms result in a STD of the form \( \sigma_e = K/\sqrt{\gamma} \), where \( \gamma \) is the SNR, and \( K \) is a constant unique to each algorithm. The values of \( K \) are 0.71 for the peak algorithm and 1.3 for the leading edge projection algorithm. Clearly, the peak tracking algorithm has the smaller tracking errors in the case of random noise.

The relative multipath performance of the detection algorithms with a band-limited signal can be best judged by normalising the results by the signal bandwidth. Using this method the performance of diverse systems can be compared on an equitable basis. Because the measurement timing errors are related to the bandwidth used by the system and the relative amplitude of the multipath interference, a multipath figure of merit (FOM) is defined as the dimensionless ratio

\[
\text{FOM} = \frac{\text{Maximum TOA error}/m}{(\text{RF bandwidth})^{-1}}
\]

Because the nominal time resolution is related to the bandwidth of the RF signal, the FOM defines how efficiently the algorithm utilises the available bandwidth in mitigating the effects of multipath interference. Using the definition of bandwidth, multipath interference amplitude and the spread-spectrum chip rate, the FOM becomes

\[
\text{FOM} = \frac{2e_{\text{max}}\beta}{m} = \frac{2e_{\text{max}}B\tau_0}{m} \quad (32)
\]

where the maximum TOA estimation error because of multipath interference is \( e_{\text{max}} \), the normalised RF signal bandwidth is \( 2\beta \) and \( \tau_0 \) is the chip period. The smaller the FOM is, the better the performance. Although the FOM is derived on the basis of a single interference source, it serves as a useful indicator of the performance of any algorithm for multipath interference in general. Note that when the FOM is used to measure the noise performance of various TOA estimation algorithms, the multipath amplitude can be simply set at unity, that is

\[
\text{FOM} = 2e_{\text{max}}B\tau_0 \quad (33)
\]

Equation (32) can be applied to calculate the FOM for the peak and leading edge algorithms subject to multipath interference. From (21), the multipath FOM for the peak algorithm is given by

\[
\text{FOM(peak)} \simeq \frac{2a\beta}{\sqrt{2}\gamma} = 0.54 \quad (\beta = 1, \quad a = 0.634) \quad (34)
\]

Similarly, the multipath FOM for the leading edge projection algorithm when the multipath delay lies between the delays of the measurement points is

\[
\text{FOM(edge)} = \frac{2\beta\tau_0}{m + 1} = 0.20 \quad (35)
\]

where the multipath amplitude of \( m = 0.5 \) has been used, and the threshold delay parameter is \( \tau_m = 0.15 \) chips. As noted previously, the leading edge algorithm has multipath errors about three times less than the peak detection algorithm and this is reflected in the FOM.

It is interesting to compare these results with the performance of GPS, which has a RF bandwidth of 20 MHz, and modulated at 1 Mchip per second in the civilian mode, so \( \beta = 10 \). The best-case ‘narrow correlator’ GPS receivers have a reported performance [30] of about 9 m error for a relative multipath \( (m) \) of 0.5, which corresponds to FOM = 1.2. Thus, in the civilian C/A mode of operation, GPS receivers utilise the available RF bandwidth rather poorly.

7 Conclusions

In this paper, we have analysed practical peak and leading edge algorithms for estimating TOA under bandwidth constraints. The TOA measurement errors are analysed in detail when either measurement noise or multipath
interference exists. Simplified expressions of the TOA error statistics are derived so the performance of the two methods can be readily compared. From these analytical expressions of the epoch tracking error, simple dimensionless FOM were derived to summarise the performance of the algorithms, both for random noise and multipath interference. These analytical expressions allow system designers to quickly assess the overall performance of positioning systems without the need for simulations. Although not pursued in this paper, the analytical expressions for TOA errors from a single multipath interference source could be used to develop analytical expressions for more complex multiple multipath interference by superposition.

8 References


[29] Maxim Integrated Products. MAX5865, ultra-low-power, high-dynamics performance 40 Msps Analog Front End