Fine Time Aiding in Unsynchronised Cellular Systems: the Benefits for GPS Integration in Mobile Phones

Peter Duffett-Smith, Cambridge Positioning Systems Ltd.
Tony Pratt, Orbstar Consultants

BIOGRAPHIES

Peter Duffett-Smith is the founding director of Cambridge Positioning Systems Limited and is its Chief Scientific Officer. He is also Reader in Experimental Radio Physics at the Cavendish Laboratory, University of Cambridge. His PhD is in Radio Astronomy and, with others, he is the inventor of the Matrix positioning method and E-GPS technologies mentioned here.

Dr Tony Pratt graduated with a B.Sc. and Ph.D. in Electrical and Electronic Engineering from Birmingham University, UK. He was on the teaching staff at Loughborough University, UK from 1967 to 1980. He held visiting professorships at Yale University; IIT, New Delhi and University of Copenhagen. In 1980, he joined Navstar Ltd, as Technical Director. In 1995, with Peep plc, he was involved in the formation of Tollstar Ltd, a 5 company consortium developing of Electronic Road Tolling. He left Peep in 1997 and joined Navstar Systems Ltd as Technical Consultant. Subsequently he was Technical Director (GPS) with Parthus. He is also Special Professor at the IESSG at the University of Nottingham, UK. He is a Consultant to the UK Government in the development of Galileo Satellite System. He is a Member of the EC Signal Task Force and sub-groups and was closely involved with the EU-US Agreement on GPS Galileo Cooperation. He is Senior Consultant with the GPS Telematics Group at QinetiQ Ltd, UK.

ABSTRACT

GPS will play a central role in the provision of location based services using mobile cellular telephones. Many commercially-available handsets ( terminals) already include integrated GPS components, mostly for use on CDMA and CDMA 2000 networks in which the transmitted signals are synchronised to GPS time. This synchronisation brings advantage in that it is relatively straightforward to provide the GPS receiver in the mobile terminal with assistance in the form of GPS time accurate to within approximately 10 µs. By a somewhat complex route, the time aiding extends the usability of the GPS service into buildings and other shielded environments. However, on GSM and W-CDMA networks, which dominate globally, the provision of time aiding is more difficult because the network transmitters are asynchronous by design. There is no concept of ‘network time’, and the transmitted signals have no pre-determined relationship with each other or with GPS time.

Duffett-Smith et al. [1,2] have proposed and demonstrated a method of providing fine time aiding (FTA) in GSM and W-CDMA networks, called Enhanced GPS (E-GPS), which uses the Matrix method [3-6] of locating a mobile terminal from the network signals received by the terminal. In E-GPS, the network signals are used as a ‘receptacle’ for GPS time after calibration. In ref. [1], they reported the first measurements made on GSM networks showing that FTA accuracies of 1 µsec or better were sustainable even over several hours. This conclusion is further supported in a companion paper to this one [7] which demonstrates that, in some circumstances, FTA at the sub 2 µs level is possible over elapsed time periods exceeding one day.

In this paper, we examine and quantify the benefits which FTA brings in terms of the reduction in the complexity of a GPS implementation. We show how FTA combined with a precise knowledge of the local oscillator offset render it unnecessary to use massive parallel correlator hardware designs. On the contrary, we show that excellent in-building performance may be achieved using as few as two complex correlators per GPS satellite channel. These benefits come mainly from three factors:

(a) a reduction in the code phase search window by a factor of more than 500;
(b) frequency stabilisation of the local oscillator using the network signals; and
(c) the use of a stronger statistical test in the signal detection algorithm for a given false-alarm rate.

The changes in the statistical tests requirements in (c) provide an effective gain in receiver sensitivity of up to 6 dB whatever number of correlators is used. The further benefits of FTA can be ‘traded off’ between cost saving (simpler GPS implementations), time saving (time to first acquisition), and sensitivity increase (penetration into in-building environments etc.). FTA at 2 μs accuracy provides the simplest and lowest-cost GPS receiver with an overall performance gain approaching 23 dB, giving it a performance in speed and sensitivity similar to that achieved using much more expensive parallel correlator designs.

INTRODUCTION

The principle of providing assistance to a mobile GPS receiver is well known and examples may be seen dating before 1986 (e.g. ICD GPS-150, first version). The assistance comes in several forms: satellite information (almanac and ephemeris), position information, and time information. We are concerned here with the provision of time information, in particular Fine Time Aiding (FTA) by which the GPS time at the terminal can be extracted from network signals within an accuracy of 2 μs. The standardised method of providing this information is by means of a message, sent from a GPS server in the network to the mobile terminal, indicating that a particular signature in a network signal receivable in the near future by the terminal corresponds precisely to a particular GPS time.

There are several problems with this approach. The first is that messages take a relatively long time to set up and send. Typically, the terminal sends a message to the server (several seconds delay) requesting the GPS time assistance. The server responds with the requested information (several more seconds), and only then can the GPS receiver begin its satellite acquisition phase. It will be appreciated that such assistance is only needed when the satellite signals are weak – e.g. inside buildings and other shielded spaces, so that a long acquisition period is anticipated. Several more seconds of integration may be required. The overall effect can therefore be to insert a long delay between position request and position response which may be too much for the application.

A second problem with this approach is that the server in the network does not know the precise position of the mobile terminal. A coarse correction may be made using Timing Advance (TA) information, if available, or Cell Identification (Cell-ID) information. In either case, there can be several microseconds of uncertainty in the GPS time provided.

A third problem is that expensive network-based GPS equipment is required to provide this FTA service. Few, if any, GSM networks can have been so equipped so far.

For these, and other, reasons, Duffett-Smith et al. [1,2] have proposed an alternative method of providing FTA to a mobile terminal. The method is based around the Matrix positioning system which is able to locate the mobile terminal using the measured relative receive time offsets of the signals from surrounding base stations of the network. The calculation provides not only terminal positions, but also a list of relative transmission time offsets of the base stations. When coupled with the calculated position of the handset and sufficient stability of the network timing signals, an initial calibration of the receipt of the signals from one base-station against GPS time can be carried around within the terminal and used, at a later time, to infer GPS time from the receipt of the signals from the same or another base station. This method, in effect, uses the unsynchronised, but stable, network signals as a remote repository of accurate GPS time.

Duffett-Smith and Tarlow [1] have demonstrated FTA accuracies of about 1 μs. Pratt et al. [7] have provided additional support in the form of Allan standard deviation curves of GSM signals received by a terminal which show that 2 μs accuracy is available over elapsed times of many hours or more than a day. That the calibration of the network signals is carried around inside the terminal ensures that FTA is available without delay when needed, and this method does not require the support of a network-based GPS server.

GPS EQUIPMENT DESIGN TRADE-OFF

What performance and design advantages does the availability of FTA accurate to 2 μs bring? The answer, of course, is multi-dimensional, and depends in particular on what GPS receiver configuration is being considered. The simplest useable configuration is probably one in which there are just two complex correlators per satellite channel, bearing in mind the probable need to track a satellite once acquired. Such a configuration will have the lowest cost, lowest silicon ‘real-estate’ footprint (the baseband processing could also be implemented in software), and may have the lowest power consumption, but it will be practical without FTA only in strong signal conditions. On the other hand, configurations with thousands of correlators can do parallel searching over all possible code-phase offsets, so need not be supplied with FTA. However, these are expensive, power-hungry, and have larger silicon footprints.

With any complexity of GPS configuration, whether there are just two or two thousand correlators per satellite
channel, FTA will always provide up to 6 dB of real sensitivity gain. This is because the satellite detection strategy can be made more aggressive for any given false-alarm rate as the search window in which a cross-correlation peak might be attributed to a satellite signal is much narrower. The statistical likelihood of a noise peak crossing a detection threshold within the window is therefore smaller. We quantify this below.

The other benefits of FTA include, for a given GPS receiver complexity, increasing the signal integration time and reducing the re-acquisition time. These are also examined below.

**PERFORMANCE GAIN**

One of the key attributes for understanding the performance gain from fine time aiding is the re-acquisition budget. The GPS re-acquisition mode can be entered when a GPS location solution (or ‘fix’) has previously been determined with a current ephemeris, or at least with one which is nearly current. The definition can be made precise: in the context of FTA, a fix will have been obtained within the previous few hours. As explained above, the FTA information is obtained from previously-calibrated modulation signatures (such as synchronisation bursts) in the base-station signals. The calibration in relative time offset and time offset rate is used to provide future calibrated GPS time to an accuracy of not worse than 2 \( \mu s \), and can be relied upon for many hours even though the mobile terminal may have moved and changed serving base station in the intervening period.

The code search budget during reacquisition is given in Table 1.

<table>
<thead>
<tr>
<th>code uncertainty</th>
<th>rms value</th>
<th>unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>receiver clock error (FTA)</td>
<td>2.0</td>
<td>( \mu s )</td>
</tr>
<tr>
<td>horizontal location error (Matrix)</td>
<td>100</td>
<td>m</td>
</tr>
<tr>
<td>satellite clock error</td>
<td>&lt; 4</td>
<td>ns</td>
</tr>
<tr>
<td>satellite position error</td>
<td>&lt; 0.1</td>
<td>m</td>
</tr>
<tr>
<td>HDOP</td>
<td>1.6</td>
<td></td>
</tr>
</tbody>
</table>

**Required code search space** 2.01 chips

Table 1 indicates the crucial role which FTA can play in containing the reacquisition search envelope for the correct code phase. There are several other important side-effects including:

(a) a reduction in the search space leading to a significant reduction in the complexity of the GPS base-band processor;
(b) for a given GPS base-band complexity and fixed observation interval, a proportional increase in the coherent or incoherent signal integration time, improving GPS sensitivity;
(c) for given base-band complexity, a reduction in the re-acquisition time (observation interval); and an improvement in the test statistics, so that there is a lower probability of false alarm or higher probability of detection, or both.

The proportion of each one of these is a performance cost trade-off established when the GPS receiver design is frozen. We provide indicative gains for case (b) since the reduction in re-acquisition time or base-band complexity is simply proportional to the reduction in code search space. By inspection of Table 1, this is evidently about 500:1 in comparison with an un-aided GPS receiver, performing a blind search (with an accumulated 1 ms time or location uncertainty). This is an appropriate response if no position aiding is available and the receiver has been moved to a new, unknown, location. Partial time or location aiding would reduce the improvement ratio, but it is usually the receiver clock which then contributes the largest uncertainty (with an undisciplined quartz oscillator). Fine time aiding solves this problem.

**SENSITIVITY IMPROVEMENT**

We now examine the sensitivity improvement which can be attached to GPS reception using fine time aiding of the receiver clock. In order to make valid comparisons, a number of receiver variables have been fixed in what follows. In practice, the performance gain may be taken in various ways and this makes comparison more difficult. The case studied here is somewhat artificial but is representative of what is possible.

The features of the GPS receiver, fixed for the comparison are:

1. the base-band complexity (i.e. the number of complex correlators available to each satellite channel);
2. the observation time interval (i.e. the allowed re-acquisition time); and
3. the detection statistics (i.e. the probability of detection, \( P_d \), and probability of false alarm, \( P_f \)).
As mentioned previously, one consequence of the reduction in the search window in code phase is a similar reduction in the overall false alarm probability, or a change in the detection threshold to maintain a given $P_f$. We introduce the statistical preliminaries in the next section which permit the computation of both $P_f$ and $P_d$.

**STATISTICAL PRELIMINARIES**

We consider the output of a complex filter matched to the wanted GPS signal (i.e. one which uses a replica of the code sequence for the chosen satellite, a replica of the received carrier signal, matched in frequency and phase within certain limits, and integrated coherently for a time, $\Delta T$). Most of the statistical expressions can be derived from standard forms (see for example: ref. [8]).

**PROBABILITY OF FALSE ALARM**

When there is no satellite signal present, the outputs $(x, y)$ of the in-phase and quadrature-phase channels of a typical correlation receiver, are random noise with Gaussian probability density functions (PDF), having means of zero and variances of $\sigma^2$ in each channel. The variance is a result of thermal and receiver noise, and its value is affected by the band-limiting action of the correlation receiver, the gain of the RF processor and the integration time. The joint PDF, $P(x, y)$ for $x$ and $y$ is therefore given by:

$$P(x, y) = \frac{1}{2\pi\sigma^2} \cdot \exp\left\{-\frac{1}{2\sigma^2}(x^2 + y^2)\right\}, \quad (1)$$

where we have assumed that $x$ and $y$ are statistically independent.

In common with many receiver designs, we take it that the receiver computes the test statistic $z = (x^2 + y^2)$ to determine the presence or absence of a signal by comparison with a threshold. The PDF for $z$ is computed by a reversible Cartesian to polar transformation $(x, y) \leftrightarrow (z, \theta)$, where

$$z = (x^2 + y^2) \quad \text{and} \quad \theta = \tan^{-1}(y/x). \quad (2)$$

The distribution of $z$ follows after integration over the $\theta$ domain:

$$P(z) = \frac{1}{2\sigma} \cdot \exp\left\{-\frac{z}{2\sigma^2}\right\}, \quad \text{with} \quad \bar{z} = 2\sigma^2 \quad \text{and} \quad \sigma_z = 2\sigma^2. \quad (3)$$

Attempts to improve GPS receiver sensitivity are limited by the duration of coherent integration because of the data modulation. A number of different strategies have been proposed but this is not the subject of this paper. For now we will take it that, because of FTA, the times of the data-bit transitions are known accurately in receiver time for the constellation which was visible at the time of the original calibration (the clock bias and clock rate of each satellite is required). This implies that coherent integration for $\Delta T = 20$ ms is allowed, though this might be lengthened with some clever algorithms. In order to improve sensitivity, incoherent integration for a number, $N$, of $\Delta T$ intervals is used with a test statistic, $u$, which is the accumulated sum of the energy from each coherent integration, $z_i$, weighted by the variance, $\sigma_i$, of each contribution:

$$u = \sum_{i=1}^{N} \frac{z_i}{\sigma_i^2}. \quad (4)$$

In practice, each $z_i$ contribution is usually given the same weight. We can assume that each of the $z_i$ random variables is statistically independent of the others. The PDF for $u$ is therefore given by

$$P(u) = \frac{u^{N-1}}{2^N \Gamma(N)} \exp\left\{-\frac{u}{2}\right\}, \quad (5)$$

where $\Gamma(N)$ is the Gamma function for $N$. It is not the purpose of the paper to discuss the multiple alternative strategies for GPS receiver processing (for example to improve sensitivity). The paper does, nevertheless, use a reasonable methodology for a suitable processing model. The results are therefore indicative of the performance which can be attained in general.

A false signal detection (false alarm) occurs when the test statistic, $u$, exceeds a threshold value, $t$. The probability, $P(u \geq t)$, of this event is:

$$P(u \geq t) = \sum_i P(u) \, du = P_f. \quad (6)$$

There is a closed form infinite series expression for $P_f$ (see for example in ref. [9]) but this does not assist the process of determining the values of $t$ for a required value of $P_f$. Equation (6) establishes the probability with which the threshold is exceeded in any searched cell.

In a typical search arrangement, a number, $k$, of cells will be searched. The probability that the threshold has been reached or exceeded in one or more cells, corresponding to one or more false detections, is therefore given by the Binomial distribution:

$$P_f = 1 - (1 - P_f)^k. \quad (7)$$
In the case where \( P_i \) is much less than 1 (as is the usual case), equation (7) can be approximated by using just the first two terms of the binomial expansion, giving

\[
P_f = k P_i
\]  

(8)

Plots of \( P_i \) vs \( \sigma \) are shown in Figure 1 for various values of \( N \).

![Figure 1: The probability, \( P_i \), of the noise crossing a threshold, \( \sigma \), in a single cell for various accumulations, \( N \), of coherent integrations of 20 ms.](image)

(In Figure 1, \( N \) has been selected as a power of 2 since the performance changes for smaller increments are not significant.)

The curves in Figure 1 show how a higher threshold, \( \sigma \), can be used for a given false alarm probability, \( P_f \). For example, FTA with an accuracy of 2 \( \mu \)s provides for a search window which is about 500 times narrower than one which encompasses the entire code-phase space. If we select, say, a probability of false-alarm of \( 10^{-2} \), then the curve for \( N = 1 \) in Figure 1 implies that a threshold value of about 10 \( \sigma \) is required with FTA (i.e. for \( k = 4 \) in Equation 7). The corresponding \( P_f \) without FTA (\( k = 500 \)) is \( 5.10^{-4} \), implying that the threshold value must be set to about 25 \( \sigma \). The corresponding thresholds in the case of \( N = 64 \) are about 175 and 210 \( \sigma \) respectively. In each case, of course, the lower threshold implies a higher probability of detection of a signal. If the deflection in a channel is simply proportional to the signal power, \( s \) (in units of \( \sigma^2 \)), then the examples just given for \( N = 1 \) and \( N = 64 \) imply increased probabilities of detection of about 25/10 (about 4 dB) and 210/175 (about 0.8 dB) respectively. However, more precise calculations are required (especially for \( N > 1 \)) as follows.

**PROBABILITY OF DETECTION**

When a signal is present, the signal energy is split between the I and Q channels of the code-phase aligned correlator, the split depending upon the phase of the carrier replica with respect to the signal. The frequency difference (\( \Delta \omega \)) between replica and received signal may be non-zero, providing it satisfies the condition that the phase-change over the coherent integration period is small, i.e. \( \Delta \omega \Delta T \ll \pi/2 \).

The maximum likelihood (ML) detection of the signal requires the formation of a test statistic \( z = (I^* + \bar{Q}^*) \), whereby the signal energy is concentrated in the measurement \( z \). After the \( i \)th period of coherent integration, each of the correlator outputs, \((x, y)\), is assumed to hold the same signal energy. We then apply the same non-coherent accumulation process as we applied above to the noise-only case. The PDF of the output is an independent, jointly-normal distribution with a non-zero mean, \( \lambda \). As before, we add together \( N \) values of \( z \) to form the accumulated output. This has a chi-squared distribution which is non-central with \( 2N \) degrees of freedom. There is a unique rotation on the variables in \( 2N \)-space which aligns the vector to the distribution mean with just one of the co-ordinate axes:

\[
\tilde{z}_i = \frac{z_i}{\sigma_i^2}.
\]  

(9)

The mean, \( \lambda \), along this chosen axis (say the \( 2N \)th axis) is then given by

\[
\lambda = \frac{\lambda_i}{\sigma_i^2},
\]  

(10)

where each \( \lambda_i \) is the mean of each \( z \). This is precisely the process of measuring the accumulated energy from each coherent integration interval, weighting in accordance with the associated noise energy, and forming the incoherent sum. This is the processing required by the ML optimum processor. After some manipulation (see ref. [8] page 238), we obtain the PDF, \( P(u|s) \), for \( u \) (equation (4)) given a signal level \( s \) (in units of \( \sigma^2 \)):

\[
P(u|s) = \frac{u^{(N-1)}}{2^N \sqrt{\pi}} \exp \left(-\frac{1}{2}(u + \lambda)\right) \sum_{r=0}^\infty \frac{\lambda_r u^r}{r!(2r)!} \frac{\Gamma(r+0.5)}{\Gamma(N+r)}
\]  

(11)

The probability of detection for this known signal is the cumulative distribution of \( P(u|s) \) from a threshold, \( t \) (just as in equation (6)). The value of \( t \) is chosen to provide the required \( P_d \):

\[
P_d = \int_{-\infty}^t P(u|s) \cdot du.
\]  

(12)

It is assumed in this analysis that the signal is only present in one known cell (a specific code alignment). Examples of curves of \( P_d \) versus signal-to-noise ratio are given in Figure 2. These and are known as Receiver Operating
to determine the size of this effect, we consider each of the 10 ms coherent integrations to have a loss of signal-to-noise ratio of 3 dB in comparison with a 20-ms coherent integration. Of the 10 ms integrations, on average three quarters of them do not contain a data bit transition whilst one quarter do. Of these that do contain data bit transitions, the worst case effective signal is zero (if the data bit transition is in the centre of the integration interval).

The incoherent addition of the two 10 ms coherent integrations recovers some of the lost signal power, but there is still a residual loss of approximately 2.4 dB. This is made up partly of the loss in signal energy (in the interval containing the data-bit transition) and partly in the change from a single 20-ms coherent integration to having two 10-ms coherent integrations being added together incoherently.

The foregoing illustrates that the use of FTA assists in extending the period of coherent integration, and also in reducing the search code space thereby allowing reduced detection thresholds. Smaller signal powers then provide the same probability of detection. The overall improvement in detection sensitivity from these two factors is about 6.4 dB for the case where $N = 1$ and about 5.1 dB for the case where $N = 64$.

**INCOHERENT INTEGRATION OPPORTUNITIES**

A second benefit may be taken from FTA when considering a GPS receiver fixed in complexity (i.e. having a defined number of complex correlators) and with a fixed observation interval. This is the basis for making a fair comparison between the performance of a receiver with and without FTA. The concept involved is that the channel slots released by the limited code search space can be used to increase the period of signal accumulation prior to detection. This does not apply to a receiver having an arbitrarily large number of correlators so that any task can have any number of correlators assigned to it. However, such a receiver is expensive. For the sake of comparison, we will consider the simplest case in which the receiver has just two complex correlators per satellite channel. Without FTA, these two correlators must be assigned to search over 500 code-phase offsets serially. Only one five hundredth of the time available can then be spent in an integration at each offset. With FTA, the two correlators can be positioned at the correct code-phase offset, and can spend all the time integrating the signal at that position, achieving a higher sensitivity.

As an illustration of this effect, we consider the cumulative probability distribution for the probability of detection from equation (11) above with $N = 64$. This corresponds to an observation interval of 1.28 s. The ROC curves corresponding to this number of incoherent integrations are shown in Figure 3.
random. Signal propagation in low signal level cases is almost always dominated by multi-path. This gives rise to large signal level variations, with a Rayleigh magnitude distribution, and a coherence time dependent upon the receiver motion dynamics (propagation channel delay spread). In most cases, the coherence time will be much less than the integration time used in the above example. Nevertheless, the proposed calculation methodology will provide a near optimum use of the available signal energy.

**FINE TIME AIDING TRADE-OFF**

The previous calculations imply that a ‘performance benefit’ of about 23 dB is available with FTA for trading off between complexity of design (and hence cost), time to signal acquisition, and sensitivity. This is illustrated in the somewhat naïve but illustrative Table 2 below where we have made the assumption that the cost of the GPS design is just $1 as it has been assumed that the base-band processing has been implemented in software. The figures correspond to a total time available of 10.24 s.

<table>
<thead>
<tr>
<th>correlators</th>
<th>cost/$</th>
<th>Sensitivity/dBW</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>with FTA</td>
</tr>
<tr>
<td>2 (s/w)</td>
<td>1.0</td>
<td>–192</td>
</tr>
<tr>
<td>32 (h/w)</td>
<td>2.0</td>
<td>–192</td>
</tr>
<tr>
<td>256 (h/w)</td>
<td>2.5</td>
<td>–192</td>
</tr>
<tr>
<td>2048 (h/w)</td>
<td>6.0</td>
<td>–192</td>
</tr>
</tbody>
</table>

Table 2, illustrating how the performance benefits of FTA may be traded off between the complexity of the GPS receiver implementation and sensitivity which may be achieved with and without FTA

**CONCLUSIONS**

- The availability of FTA accurate to 2 μs brings with it a performance benefit equivalent to about 23 dB in sensitivity. This may be traded, when the receiver design is frozen, between the complexity (number of correlators), the sensitivity, and the time to signal acquisition.
- Even for designs with 2048 correlators per satellite channel, FTA still brings a sensitivity increase of about 6 dB.
- The simplest design, having just 2 complex correlators per channel and therefore one which may be implemented in software, can achieve a sensitivity of –192 dBW with FTA in about 10 s.
REFERENCES