UltraWideBand Indoor Positioning Systems and their Use in Emergencies

S. J. Ingram, D. Harmer, and M. Quinlan
Thales Research and Technology (UK) Ltd., Worton Drive, Reading, RG2 0SB, UK

Abstract – **Reliable and accurate indoor positioning for moving users requires a local replacement for satellite navigation. An UltraWideBand (UWB) system is particularly suitable for such local systems, including temporary installations supporting emergency services inside large buildings.**

The requirements for emergencies will be very variable, but will generally include: good radio penetration through structures, the rapid set-up of a stand-alone system, tolerance of high levels of reflection, and high accuracy. The accuracy should be better than 1 m, as sometimes it matters which side of a door you are, and locations should be in 3 dimensions. Support for robots as well as people would call for still better accuracy. Rapid set-up implies very little surveying of the fixed terminals, and positioning relative to the mobile and fixed terminals.

A radio system that measures ranges between fixed and mobile terminals matches these requirements, but the requirements for accuracy and for dealing with multipath need a bandwidth of more than 1 GHz. Thus UWB is the preferred solution, as it has the specific advantage of high accuracy, even in the presence of severe multipath.

This paper presents the features and system design options for UWB positioning systems, and shows how they match the indoor location demands of emergency services.

The main features that are covered are: the deployment of terminals (how many, and where), the minimum requirements for fixed terminal surveying, integration (hybridisation) with GNSS, and solving for position inside the network.

The main system design options are: whether the mobile terminals are transceivers or solely receivers, the UWB signal design and frequency span, and the use of the same signal for communications.

The paper includes results from a demonstration UWB indoor positioning system being built at TRT (UK).

I. INTRODUCTION

If the general public take to satellite positioning (GNSS) in the way they have to mobile phones, then we will soon be relying on it, and on location-based services, in our everyday lives. An indoor extension of the same high-quality (and free!) service will then be very desirable. This paper deals with indoor positioning systems which fulfil this need.

The emergency services are already becoming major users of GPS, both directly and via the "E-911" capability of some US mobile phones. For them, the extension of a reliable and accurate positioning service into buildings is more than just desirable, it is a potential life-saver.

Work on the two UWB positioning demonstration systems is supported by the British National Space Centre and by the European Space Agency. Pulse UWB terminals have been supplied by Time Domain Corporation. This support and assistance is gratefully acknowledged.

The terrorist threat is an urgent concern now, but our emergency services have always had to go into dangerous buildings and industrial structures after accidents. As well as individual incidents like gas explosions or aircraft crashes, there have always been earthquakes and massive destruction in wars. At the very least we would like to keep track of those who enter disaster zones, and accurate positioning can also support other aids to make their work more effective and safer.

In this paper we discuss the requirements for indoor positioning systems. We will cover both a general-purpose system installed as part of a building's infrastructure, and a rapidly deployed ad-hoc system for use in emergencies. We then consider how UWB positioning systems can be designed to meet these different requirements.

A. Terminology

We use both "positioning" and "location" to mean the measurement of position. Fixed terminals are commonly called "pseudolites", but we have avoided this term as it suggests imitating a GNSS signal. We will refer to mobile terminals, carried by users who are to be told their position, fixed terminals, which do not have "users", and also reference terminals, which are those used in finding the positions of other terminals, and may be fixed or mobile.

II. INDOOR POSITIONING REQUIREMENTS

In this section we set out the general requirements for indoor positioning systems. If these are to be widely adopted, with any one user's terminal working in all equipped buildings, there must be a single standard. Such an open standard should support the majority of potential applications, even if most installations do not need them all.

A. What is "Indoors?"

We start from the assumption that, quite soon, the combined GNSS service, available free to everyone, will offer a 1-3 m positioning accuracy even in most city streets. There will still be some "urban canyons" and other blind spots, but for most users the most noticeable gaps in this service will be all those large indoor spaces which are used much like outdoors – shopping malls, airports, large stores and warehouses, railway stations (some underground), underground passageways, and others.

There are also many large buildings, both public and private, where precise location could be really useful: schools and colleges, public offices, workplaces of all descriptions… you can add to the list yourself.

0-7803-8416-4/04/$20.00 ©2004 IEEE.

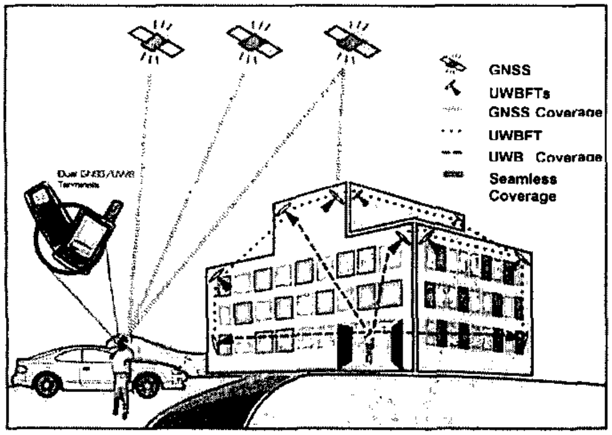

Fig. 1. A UWB positioning system as building infrastructure

Fig. 1 shows the general arrangement of such a system. Some of these indoor spaces have only one floor level; others may have many. In some instances height above the floor does not matter, and may be quite limited (e.g. spanning only 2-3 m). On the other hand, knowing "which floor" is usually very important, and any ambiguity about this would be seen as a gross error.

Other spaces not normally called "buildings", such as ships or mines, should also be considered. Finally, there are some outdoor spaces where GNSS service levels are poor; such as factories and chemical plants with a large amount of metalwork overhead and all around. These are also potential sites for such a positioning system.

B. Accuracy

We have said that we expect the open-access service of GNSS collectively to give an accuracy of 1-3 m in most places outdoors. Indoors we will need rather better accuracy than this, partly because the natural scale of such spaces (set by rooms and furniture) is smaller. The accuracy requirements of a range of potential applications have been assessed, as shown in Fig. 2. From this assessment, it can be seen that an accuracy of 0.3 m will serve the majority of applications.

GPS was never intended to work indoors, so it is remarkable how well indoor GPS techniques can perform. All indoor GNSS techniques rely on integrating the signal for longer than 1 ms and, for the highest sensitivity, for much longer. The pedestrian user moves in an irregular and unpredictable manner. This places severe limitations on the use of coherent integration, which is the only way of integrating up from very low signal levels. In addition, signals coming through a roof or outer wall are as likely to be indirect (by a deviated and longer path) as direct and attenuated. This detracts from the accuracy of positioning. For these two reasons, we do not expect to see metre-class performance from indoor GNSS in most buildings, nor coverage at 90% and over in indoor spaces as a whole.

Positioning systems in all these places could be installed as fixed infrastructure, supporting an extension of GNSS services. In a major emergency, such systems cannot be relied upon, and neither can other services such as lighting. There is thus a need for temporary system to support intervention after such emergencies, mainly (but not exclusively) by the emergency services (fire, police, and medical).

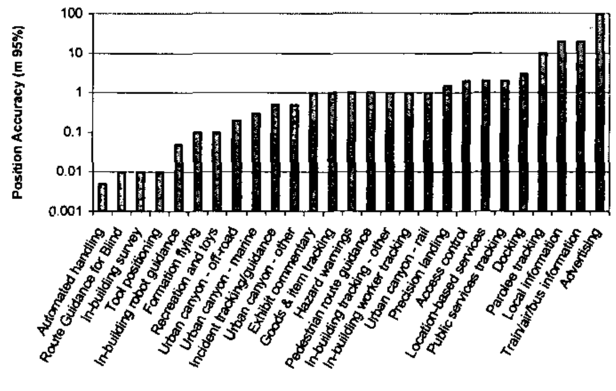

Fig. 2. Accuracy requirements of potential applications

C. Operation

In a large building, and especially one with public access, there will already be a building services team to operate and maintain the heating, power, security, communications, and other infrastructure. A positioning system would naturally fit in as another responsibility for this team. Emergency planning and safety equipment are also among their responsibilities.

An emergency deployment is very different. Apart from the use which may be made of pre-installed but compatible fixed terminals, everything must be brought by the emergency services in their (usually road) vehicles. While there are clearly dangers in operating a mobile control room close to a damaged building, on-site control of the system has operational advantages (Fig. 3).

D. Setting up

For a building services manager, when asked to install this new system, the biggest question will be how much effort do I need to install and operate it? And can I use my existing workers, or do I need more? Even if specialist contractors are used for installation, anything which simplifies their task and saves their time will reduce costs. As a high cost of installation is likely to be the biggest hurdle to widespread

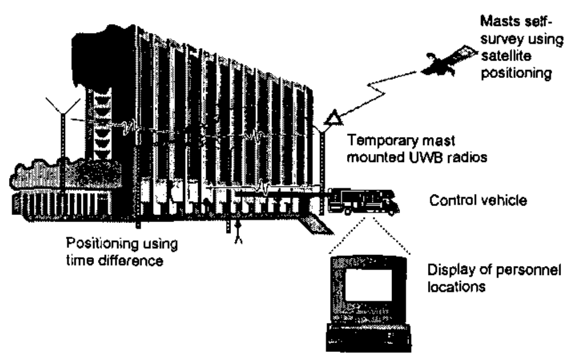

Fig. 3. A UWB positioning system as an emergency deployment

adoption of such systems, this is a high priority for the system design.

For an emergency application, a similar consideration applies, though here it is not so much cost that counts as the lack of time and the urgent priority for people to do something more important. Thus both cases must exploit self-calibration and self-survey to the maximum. If human intervention is called for, the system should identify what is needed so it can be done quickly and simply.

Speed is absolutely vital to the successful evacuation of the occupants. As soon as the emergency services arrive on site, they should have some coverage if possible.

III. Operating Principles

This paper concentrates on positioning by (1) radio and (2) measuring distances between terminals. We will briefly discuss the reasons for this choice.

A. *Alternative Principles*

Anything measurable, if it changes as the object to be located moves, is a possible basis for position estimation. Light and sound are commonly used in short-range positioning, and human senses rely on them (and, at very short range, on touch). However, these forms of "radiation" do not penetrate most materials without being seriously distorted.

The amplitudes of radio waves can be used for positioning, be comparing signals from several sources. This makes use of the small-scale local variations due to blockage and reflection to improve its accuracy. However, its accuracy is still moderate (worse than 1 m), and it relies on calibration measurements, and on the environment being stable. It is thus not suitable for the emergency application. Whether it is suitable for crowded public access buildings is an open question.

B. *Ranging*

Radio signals are very good for range measurement, having a well-defined and high speed in air, and being easy to time. The principles of secondary (transponder) radar, and indeed of GPS, are well established.

Radio signals can penetrate some solid materials well, depending on the radio frequency as well as the material. Building materials can be placed in two categories: metallic and non-metallic. Metals can be regarded as opaque and reflective at all frequencies, while most other building materials are reasonably dry insulators of moderate dielectric constant. They may be thin and uniform (glass), thick and fairly uniform (brick, concrete), or very non-uniform (reinforced concrete). There will be many irregular metallic intrusions, such as steel frames, pipes, ducts, window frames, and other exceptions too (e.g. water in pools and cascades).

While buildings do vary a lot, in most cases radio signals will travel through at least one wall and still be usable.

C. *Uplink versus Downlink Ranging*

There are two pre-existing location technologies that have quite separate origins but are converging: those found in GPS and RFID tags. GPS is a strictly downlink-based system, so it is scaleable (can serve infinitely many users) and provides information but no associated service (such as computation). Locally-based services may add to this, of course, and may charge (where GPS is free).

RFID tags are constrained by a severe cost and size limit, and are essentially transmit-only devices (though they may be externally triggered). Locating them is the function of fixed receivers, though in the simplest case this is just a matter of knowing a tag is close enough to be received at all. With a longer detection range it becomes possible to locate tags more accurately over a large area with a network of receivers.

Both of these principles are being pursued in indoor positioning systems under development, and both may be used in the same system. The systems described in this paper have a network of reference transceivers from which mobiles can locate themselves. These can locate transmit-only tags as well, provided the air interface is designed to accommodate such transmissions. To maintain scalability, the detection range and rate of such transmissions must be limited.

D. *Multipath and its Effects*

The indoor environment is usually cluttered, rather than resembling the ideal of "free space". This clutter (in a general, rather than strictly radar, sense) comprises not only the building, but furniture, stock (in shops and warehouses), vehicles, and of course people. Some of the clutter moves about, though most does not. There are two dominant effects of all this clutter: it obstructs direct signals, and it reflects signals via indirect paths that are longer than the direct one. Signals may also diffract and scatter in other ways, always resulting in a longer path than the direct one. Passing through a barrier (e.g. a wall) will add a small delay due to the dielectric constant of its materials.

E. *Bandwidth*

The considerations in this section lead to ranging using radio signals as the preferred operating principle. The choice of UWB over any narrower bandwidth signal is the result of performance considerations, which are the subject of the next section.

IV. Performance and Design Features

The performance of the system depends on the positions of the fixed terminals it can receive, just as that of GPS depends on the constellation and local shadowing. Since the local multipath environment is always bad by GPS standards, and often very bad, there is a premium on making sure there are enough good signals to allow any bad ones to be detected and rejected.

A. *UWB Signal Types and RF*

UWB refers to signals which spread across many frequency bands that are allocated to different purposes. Thus its use always involves co-existence and electromagnetic compatibility (EMC) considerations. Its use has been permitted in the USA since the Federal Communications Commission (FCC) issued its first rules in 2002 [1], but is still being considered elsewhere in the world and within the ITU.

The FCC's rules for UWB permit operation between 3.1 GHz and 10.6 GHz for communications and other purposes, without licensing. Some radars are permitted at lower frequencies, subject to licensing and other restrictions. Proposals in Europe and in the ITU have followed the same approach, so we will also consider only the spectrum above 3.1 GHz. While emergency services might claim enough priority to override the concerns about interference, they will also be relying on radio systems operating at 1-3 GHz, and so should avoid it anyway.

No significant use of UWB has yet made, but products are being developed, mainly for very short-range communications. IEEE standards group 802.15 is leading the way in wireless personal area networks (WPANs), and various signal types have been proposed there. By February 2004, the choice has narrowed to two contenders: multi-band orthogonal frequency-division multiplex (OFDM) and high-speed direct sequence (DS).

In both of these, the spectral region at 5.15-5.875 GHz is left unused, as it has other uses (wireless LAN, Hiperlan, other data communications, as well as industrial, scientific and medical (ISM)) which are seen as incompatible with WPANs. For similar reasons, they are not compatible with indoor positioning either. Thus UWB applications must look at 3.2-5.15 GHz or 5.875-10.6 GHz. However, spectrum is continually being re-allocated to new uses, so what is now available may become unusable in the future – UWB will never be a primary user of spectrum.

Since range performance is of particular importance for positioning, the band below 5.15 GHz is preferable because losses through obstructions are lower than at higher frequencies. A bandwidth of about 2 GHz still permits a basic range resolution of better than 0.3 m.

Both OFDM and DS occupy their UWB sub-band with a nearly continuous transmission, and both use sub-bands about 500 MHz wide. This is a result of the way the FCC rules were written, insisting on a minimum occupied bandwidth of 500 MHz. Sub-bands (or "multi-band") are favoured as a way of mitigating interference. Short-pulse UWB with time modulation (TM, also known as pulse position modulation) also falls within the rules, but it cannot so easily adopt sub-bands.

Two other contenders for UWB signals do not fit the rules so well, frequency hopping and sweeping (or "chirp"). These are similar, but hopping is far more flexible and sweeping has few proponents. Frequency hopping has the ability to tailor its use of the spectrum so as to avoid interference, while still filling the band well enough.

Most other UWB applications exploit their great bandwidth as a way of reducing power spectral density (PSD) to ensure co-existence with established users while carrying data at a useful rate. Positioning, however, like radar, needs a large and ideally contiguous RF bandwidth to give it good resolution. Range measurement in the presence of multipath involves estimating the impulse response of the path faithfully enough for the direct path delay to be measured alone, usually as a leading edge.

If the overall occupied spectrum of the signal is not completely filled, but is gapped or a comb, the impulse response estimate is distorted or aliased. However, methods do exist which can compensate for the loss of parts of the spectrum, provided the basic range resolution of the signal is good. Frequency hopping is thus well suited to positioning systems, if its use is accepted by the regulatory authorities – perhaps over a restricted part of the overall UWB spectrum.

B. Range Performance and Signal Design

The FCC's rules set a limit on the transmitted PSD which, when combined with the overall bandwidth, fixes the power level. The type of signal and the receiver design make only a small difference to this. The receiver integration time is the only other important factor, and this time must be long in order to maximise range. It is constrained by two dominant effects:

- Accuracy of frequency standards
- Required position update rate

For pedestrian applications, the frequency standard is usually the limiting factor rather than the update rate. The integration time that can be achieved for a given reference accuracy is dependent on the type of UWB chosen.. Here we consider two types of UWB:

- "Classical" pulse UWB with time modulation
- Frequency hopped direct sequence

In the case of classical pulse UWB, the integration can continue until the pulses have drifted out of coherence. For a pulse duration of 0.25 ns and a 10ppm reference we can coherently integrate over 25 μs, giving an update rate of 40 kHz. The range will be independent of the pulse rate, given the same average power. The range in free-space conditions will be about 100 m with no interference present and this has been verified using pulse equipment.

In the case of frequency hopped direct sequence, let us assume a direct sequence chip rate of 10 Mcps and the same 10 ppm reference. The integration time now can be 10 ms for chip timing coherence. Frequency coherence can be recovered by Fourier Transform in the same interval. If we receive 10 transmissions, the update rate is 10 Hz, which is close to what is needed.

A lower DS chip rate results in a longer integration time; however, there is another constraint. If we assume the FCC's general UWB rules apply to frequency hopping, then the requirement to measure power in 1ms means the frequency hopper must visit all frequencies within that time. Also, to avoid ranging ambiguity and multipath weaknesses we must avoid leaving gaps in our transmitted spectrum. Further, there are advantages in multipath resistance to having several 10's of chips on each frequency hop. This results in a practical lower limit to the DS rate of about 1 Mcps.

The regulations also limit the peak power, and consequently the DS rate is (practically) forced to above 1 Mcps. At around 100 Mcps and above, the spread bandwidth is more readily jammed by a few narrow-band transmissions, and current consumption becomes a more serious issue. A chip rate of about 20 Mcps is considered to be a good compromise given the likely regulatory position and available hardware

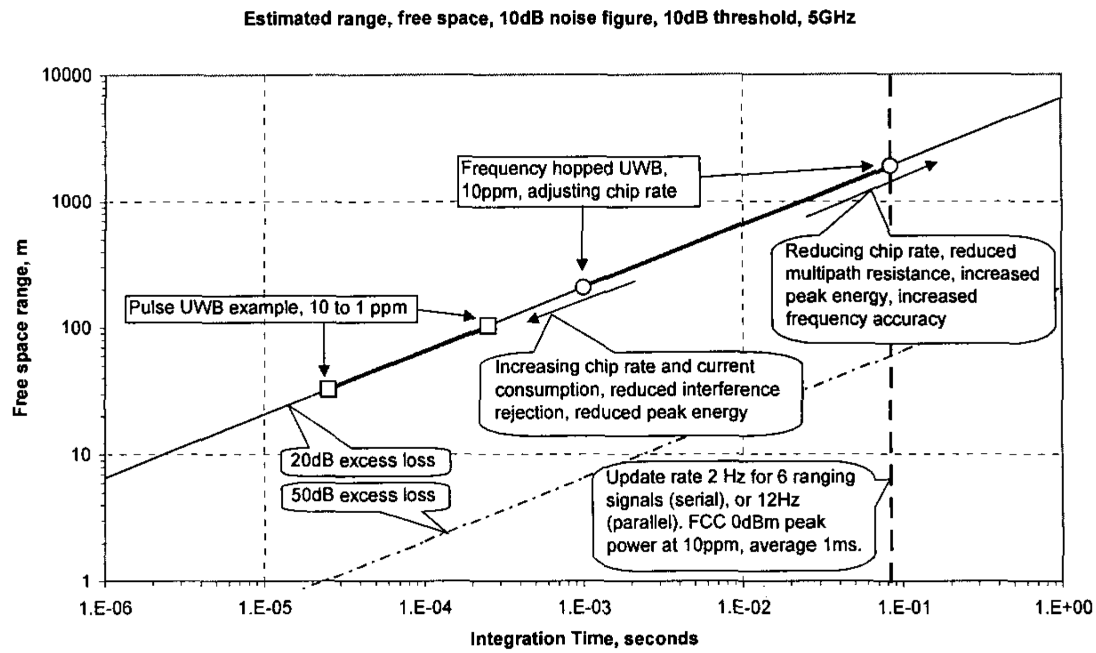

Fig. 4. The Constraints on Range and Integration Time

capability. A free-space range in excess of 1 km should be possible with this rate (or 100 m with a 20 dB penetration loss margin), and is hoped to verify this for the frequency hopped system in the near future.

Fig. 4 shows the relationship between integration time and range for systems transmitting at the FCC limit of PSD, and some of the limits for each of the UWB types. The expected fixed terminal density in 3D is about 1/1000th that of a classical pulse UWB system, with a consequent reduction in cost and in aggregated radiated power of that order. Also, by making the hopping patterns of the fixed terminals orthogonal, the aggregated radiated power for the system can be kept spectrally flat (the peak powers do not sum).

C. *Interference and EMC*

We have seen that interference is always a consideration for UWB, as it operates in other users' spectrum. If a UWB system causes interference, despite meeting regulatory limits, it will usually be because its victim is very close. In such a case the UWB system will have to adapt its transmissions. If the UWB system suffers interference it will need to adapt in its receivers to suppress this. Even if the interferer is illegal, it may not be possible to get it shut down (or not quickly).

The extreme case of "close" is another radio built into the same housing as a UWB transceiver. We expect this to be the usual arrangement, as there are already too many small gadgets to carry about, and emergency services personnel have more of them than most of us. This compatibility requirement is easiest to meet if the UWB terminal is solely a receiver, if its collocated transmitters are only on part of the time, and if positioning is still possible using signals received between transmissions. Otherwise the compatibility must be ensured by the co-ordinated design of the different terminals sharing the housing.

The key to UWB's ability share spectrum with its existing users is its high processing gain. A signal with a moderate bandwidth is spread over several gigahertz, so moderate power corresponds to a lower power spectral density. This appears as noise in a receiver not designed for its form of coding, and should be at a tolerably low level as that is how the rules for maximum PSD were derived.

On reception, this power is integrated together by a matching de-spreading operation. This process turns the signals of other systems into noise in the decoded, narrow-band, signal. Clearly, the degree of suppression in either direction is directly related to the bandwidth spreading ratio, and so inversely related to the effective bandwidth (or data bit rate) of the system.

The compromise struck between the demands of data rate and interference suppression will depend on the application. A WPAN with a very short range can have a date rate of over 100 Mb/s, while a positioning system needs a range of tens of metres but only a few kb/s of data.

The description above holds for both DS and TM coding, if "spreading" is understood to apply in both time and frequency domains, and to other systems if coding is applied at a higher level. Some forms of UWB can make explicit provision to omit particular frequencies on transmission and excise them on reception. OFDM and frequency hopping can both do this. In other cases the modulation may be used to do the same thing, notably with TM coding. A rake receiver uses sets of samples with fixed time intervals to implement a finite impulse response filter, and TM can still be applied to these bursts of samples. A corresponding technique, with short pulse bursts replacing single pulses, is possible on transmit. However, the filter which results in each case distorts the pulse shape and its spectrum, and this limits the usefulness of this technique.

D. Deployment Needed for Accuracy and Coverage

The requirement for positioning is the same as GPS in theory: a minimum of four good pseudoranges well spread out in direction. To detect a bad measurement and reject it we need at least two more. One important difference is that the positioning service must operate close to the reference terminals – far closer than is the case for satellites.

The fact that the reference terminals are so close is important because it introduces ambiguous solutions (see IV.F) and because the geometry changes very rapidly with position. For example, consider an installation within a single storey of a building. A good geometry for 3D positioning has a reference terminal nearly overhead, but it can only be 3 m above the floor or 2 m above the mobile terminal. As soon as the mobile has moved a few metres to one side, it starts to need another reference terminal above it. This would lead to a very high density of fixed terminals.

Simulations have been used to determine how many fixed terminals will be needed, and have shown that one per 200 m^2 will suffice for most large buildings [2].

While it is difficult to define clear rules for placing fixed terminals in any building, the problem becomes much simpler when considering a specific building. However, some general guidelines can be given.

- Since most of the solid matter in a building is closer to the floor than to the ceiling, place fixed terminals high up on the walls.
- Pick sites with good clear sight-lines if possible.
- While the loss through floors can be very high, the distance to the next floor up is quite short, which helps with 3D coverage.
- Good height measurement may only be needed in part of a building – and often it is not needed where it cannot be easily supported.

For an emergency installation, the considerations are substantially different. Here we must not rely on the constraints of floor, doors, elevators etc., and will operate as far as possible from outside the building. Some of the fixed terminals should be located absolutely – preferably from GNSS. To cover buildings of several floors with 3D coverage we need to place some fixed terminals as high up as we can.

This leads to an obvious basic piece of equipment: a tall mast with fixed terminals at various heights along it and a collocated GNSS receiver at the top (and perhaps lower down). This may be attached to a vehicle, or arrive on one and be erected free-standing. With three such masts around a building, provided it is not too large, further reference terminals – fixed or mobile – can locate themselves and extend coverage inwards.

Taller buildings inevitably cause greater problems. However, such buildings are now likely to have well-developed emergency plans and suitable equipment installed. The UWB positioning system could be installed so as to maximise its usefulness in an emergency. For example, including back-up power sources, and ensuring fixed terminals know where they are, would permit many to keep operating usefully if they avoid immediate destruction. If some are dislodged then they, or mobiles using them as reference terminals, will be able to deduce this by standard error-detection processes.

E. Synchronisation, and Active Versus Passive Mobiles

A GPS receiver must know the transmission times, relative to a common clock, of the signals from all the satellites. This is done by giving each satellite a very accurate clock, correcting this clock from ground-based observations, and then transmitting a short-term clock error estimate as well. This permits the mobile terminal to solve for its position and its own clock time directly.

The same principle can be used in an indoor positioning system, which has the same requirement on clock accuracy. While there is no scope for the kind of monitoring stations used for GPS, there are all the other fixed and mobile terminals which can be said to monitor the transmissions. The accuracy of synchronisation required must be commensurate with the positioning accuracy, allowing for the geometry (dilution of precision or its equivalent). For a target accuracy of 0.3 m, we need to synchronise to better than 0.3 ns.

If two transceivers are exchanging messages, and are suitably equipped for timing these messages, then they will simultaneously measure both clock offset and time of flight. The time of flight will include some internal delays within the transceivers, and any difference between and return path delays must be known. However, there are ways of separately measuring these if there are enough other terminals seen by each of them.

If all terminals are transceivers, mobile as well as fixed, then the ranging and synchronisation are the same process. Each transceiver indicates in its messages how sure it is of its own position and time, and which others it is relying on to derive these. On receipt, this information is used to prevent any "collective self-delusion" of a group of terminals relying on each other.

For large numbers of mobile terminals, there is a signal congestion problem if they all transmit. Making the mobile passive ensures scalability to any number of mobiles. Now, however, only pseudoranges can be measured, and the transmitters must all be accurately synchronised.

There is a simpler alternative, which is for the receiver to process signals from two reference terminals as a synchronised pair. This pair is performing a round-trip range and time measurement at the same time, which synchronises them. Subsequent pairings then include all reference terminals within range, and the measurements are processed as path-length delay differences (often called time of arrival difference or TDOA).

F. Solving for Position

The solution methods for TDOA are specific to that case, and even to details of the clock error model and sequence of

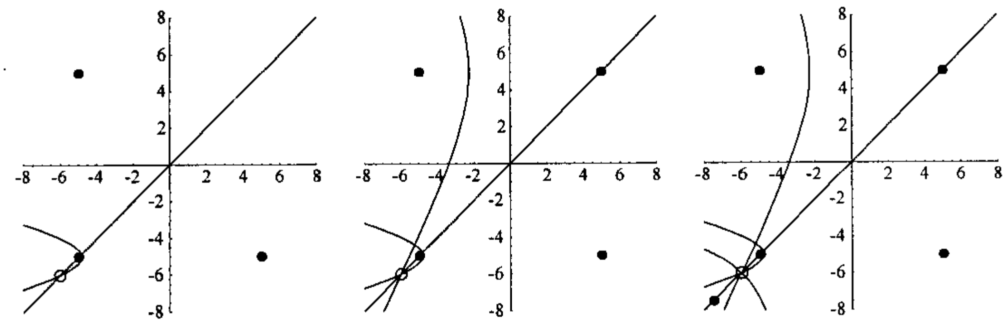

Fig. 5. Ambiguities from intersecting hyperbolae in 2D for 3, 4 and 5 reference terminals

transmission around the network. However, one approach is to convert the measurements to a set of likely pseudoranges.

Solving for position from pseudoranges is familiar from GPS practice, and can be approached by the same techniques. Linearising the geometrical equations for pseudorange around an estimated position leads to a matrix form which can be solved to minimise the least square error.

For GPS, the satellites are so far away that ambiguous solutions are rare, even if the initial estimated position is as much as 100 km from the true position. For this indoor network this is not true: the fixed terminals are more like 10 m away, so the tolerable error in the initial guess may be only millimetres in some cases. Ambiguities are possible, and may be quite close to the true answer.

To see how this can happen, a 2D example is shown in Fig. 5, where the dots are reference terminals and the circle is the mobile. In a pseudo-range system, the solution is equivalent to a minimum set of range differences solved as simultaneous equations. Starting with only three reference terminals, there are two loci, which are hyperbolae, one for each range difference. Close to a reference terminal, one locus is so curved it crosses the other twice. Adding each further reference terminal adds one more hyperbola, which should resolve the ambiguity. The second graph shows how this may not fully resolve the ambiguity but, in the presence of noise, a local minimum may give the wrong solution.

Such ambiguities are more likely with pseudo-range than with range measurements, and at the edge of the volume occupied by the reference terminals rather than in the middle.

G. 2D or 3D Solution and Constraints

Our world is three dimensional, but the three dimensions are not at all equal. We move about freely in two of them, but gravity pins us to the nearest horizontal surface downwards, and we only move vertically in certain specific places and ways. It is possible to exploit this set of constraints to improve system performance or to reduce fixed terminal numbers and so cost.

Solving for position in 3D calls for a good 3D disposition of reference terminals around the mobile terminal. This may be difficult to provide in some buildings, and it always costs extra compared to the simple installation of all fixed terminals at ceiling height.

While vertical movement is usually restricted to stairs and elevators, the exceptions may be important in some cases. The emergency application is an important exception.

V. Design Options

The choice of operating principles and main features of the system design determine the system architecture and air interface. There are several options available which will adapt this basic system to specific applications, which we will consider now.

A. Communications

The messages that are exchanged for ranging and synchronisation purposes can contain data. Some of this capacity is needed by the positioning system itself, being the equivalent of a GPS navigation message. However, some extra capacity can be provided for other uses. One can look at this combination another way, and say that data messages are being used for positioning.

There is a performance trade-off involved here, as extra data means longer messages or lower processing gain. Since the total time for a set of all transmissions is fixed by the dynamic response of the positioning service, once all this time is taken up then more data means less gain. Since the transmitted power level is strictly controlled, less gain means less range.

Military users may have a particular need to use UWB communications for its covertness, but for the target users of this system there is no such need. We can assume that they will have adequate communication channels already available to them.

Priority for use of the data capacity goes to the internal housekeeping of the positioning system. For the emergency application this is a very important factor, as fixed wiring is not available, and other communications bands are likely to be saturated. After this, the next priority is for a back-up communications link for users, possibly incorporating a "panic button" alarm facility.

B. Self-Calibration and Self-Survey

The operation of the system requires the exact locations of the reference transmitters to be known. While the source of this information is not fundamental to the operation of the system, its automatic derivation is highly desirable and may be essential for some applications to be feasible.

The locations of the reference transmitters can be determined in several ways. Some at least must be derived from external information; either maps or plans, or from a collocated GNSS receiver. However, once a few are so located, we would like the rest of the transceivers to be located automatically.

The minimum requirement for externally derived locations is usually three fixed terminals: one defines a point, two a vector, and a third one defines a second vector. Provided these three are not nearly collinear, they fully define a Cartesian co-ordinate frame. However, the direction of the third axis is decided by the handedness of the frame, which is an arbitrary choice. Providing a fourth location of a terminal well away from the plane of the other three solves the problem, as would any other way of breaking the symmetry (e.g. for three known points at the same height, that all the rest are higher up).

While the main objective is to find the terminals' locations, in doing so we usually will also determine the internal delays between the timing point in the circuitry and the antenna's effective centre.

C. *Relative Positioning by Mobile Terminals*

In the emergency application, there may not be time to position all the fixed terminals needed for full coverage, perhaps because no-one can get into the right places in advance. Coverage can be extended by placing extra fixed terminals as personnel move further into the building, and relying on their being able to locate themselves from the existing network. However, we can also use mobile terminals as part of the reference network.

The conditions for this to work are:

- Mobiles are operated as transceivers
- Error build-up can be controlled
- Errors due to terminal motion are small.

The error build-up problem is analogous to that in land surveying, where each stage of triangulation relies on previously surveyed marks. It happens with self-surveying of fixed terminals, but some of these can be sited to offer good clear paths to each other. That ensures both good range accuracy and long range, so few hops are needed to span the network.

Mobile terminals are less favourably placed, with higher errors and shorter hops. To offset this, there may be many more of them, and as they move the propagation and multipath effects vary. Both of these factors can be exploited to improve the accuracy.

The errors due to terminal motion arise from the finite time that it takes to perform all the ranging measurements. Defining the sequence of pair-wise ranging transactions will be a complicated task, and is likely to involve sub-nets with orthogonal signals and each terminal belonging to several sub-nets. This is similar to the Management of an ad-hoc communications network.

D. *Access Control and Encryption*

In a system for general public use, there may be a need for access to be controlled as part of the charging mechanism. However, this does depend on which business model is adopted, and free provision is likely to be common, and may become predominant. If access control is protecting a small fee per user, it does not need to be very strong.

It may well be that mobile phone companies are involved in the collection of these fees, either as providers of the service, or just as collectors. After all, the UWB positioning terminal is likely to be part of a mobile phone, which also contains a GNSS receiver. The phone can supply a route for controlled distribution of the keys or codes or some other information (such as reference terminal locations).

For closed professional user groups, or the emergency services, a stronger access control may be needed. In particular, a way of disabling lost or stolen terminals, even if they have been tampered with (for example so they do not report their position) might be thought important.

The techniques of access control are sufficiently well developed, whether from communications or GNSS signals, that suitable ones can easily be found.

E. *Strong Protection from Interference*

If crime and terrorism are part of the scenario in which the positioning system is needed, then malicious interference or jamming cannot be ruled out. However, the localised nature of the system and its coverage, and the presence of the emergency services themselves, will make this difficult.

One key enabling factor in defeating jamming is to make every terminal measure such interference and pass on its measurements to a central point. This will allow the jammer to be found and quickly put out of action.

A second way of approaching the problem is to provide interference suppression in each receiver, as discussed in section IV.C.

F. *System Control Topologies*

Some aspects of the system require co-ordination by a single central control function, while others are better left to each terminal to decide or negotiate with its neighbours. As a general rule a top-down controlled arrangement is easier to implement, and would be better for a fixed public system. If there are any questions of regulatory compliance to deal with, then deterministic behaviour is preferable.

On the other hand, for rapid intervention in a disaster area, with a fast-changing set of reference terminals, the ideas of ad-hoc network building are more appropriate.

G. *Vehicle-Mounted Terminals.*

"Pedestrian" is only a temporary state, and people get in and out of vehicles very frequently. Emergency services personnel carry a lot of kit with them, but still have far more in their vehicles. If a GNSS-based terminal in the vehicle is used to get to the scene, it will need to be co-ordinated with the handset or terminal which its user "wears" when out of the vehicle. This will involve the vehicle having a UWB terminal, which may offer communication with the user. However, in the "network-enabled" future we should expect there to be other communications channels for this.

H. Hybridisation with GNSS

For large public buildings, and a service that extends GNSS indoors, we have already identified that both systems will have their worst performance close to the outer boundaries of the system. For GNSS, it is the shadowing by the building itself, plus reflections off others, which drastically reduces the number of satellites in view as the building is approached. Canopies and colonnades make this worse. For the indoor system, the geometry for positioning becomes less favourable as the user moves from well inside the network to a position where all the reference transmitters are to one side of it. In addition, fewer signals in total will be received if there are no reference terminals outside the buildings.

If there are enough signals for one system to yield a solution and to reject corrupted measurements, then the other's signals can be added as a form of augmentation. However, this is a rather restrictive condition. It would be better, in the sense of restoring coverage to more of these perimeter situations, to perform a joint solution from both sets of signals: a full hybrid solution.

In such a case it is no longer possible to solve for local clock error against each system's clock, so the clock offset between the indoor systems and the GNSS system(s) must be provided, and to the equivalent of the pseudorange accuracy. Any pair of collocated GNSS and UWB terminals, a long as both have good signal reception and favourable geometry, can provide this synchronisation. It is also possible to use the GNSS receiver as a reference for local relative operation, providing high accuracy without depending on differential corrections from an external service.

VI. Demonstrations

TRT (UK) has implemented two proof-of-concept demonstration indoor positioning systems. One uses "classical" short-pulse UWB, with time modulation coding to whiten the power spectrum and discriminate against interference. The other uses a frequency-hopped direct sequence UWB signal designed specifically for positioning.

At the time of writing, some initial results are available from the short-pulse system. A screenshot, recorded before full calibration and without true position for comparison, is shown in Fig. 6.

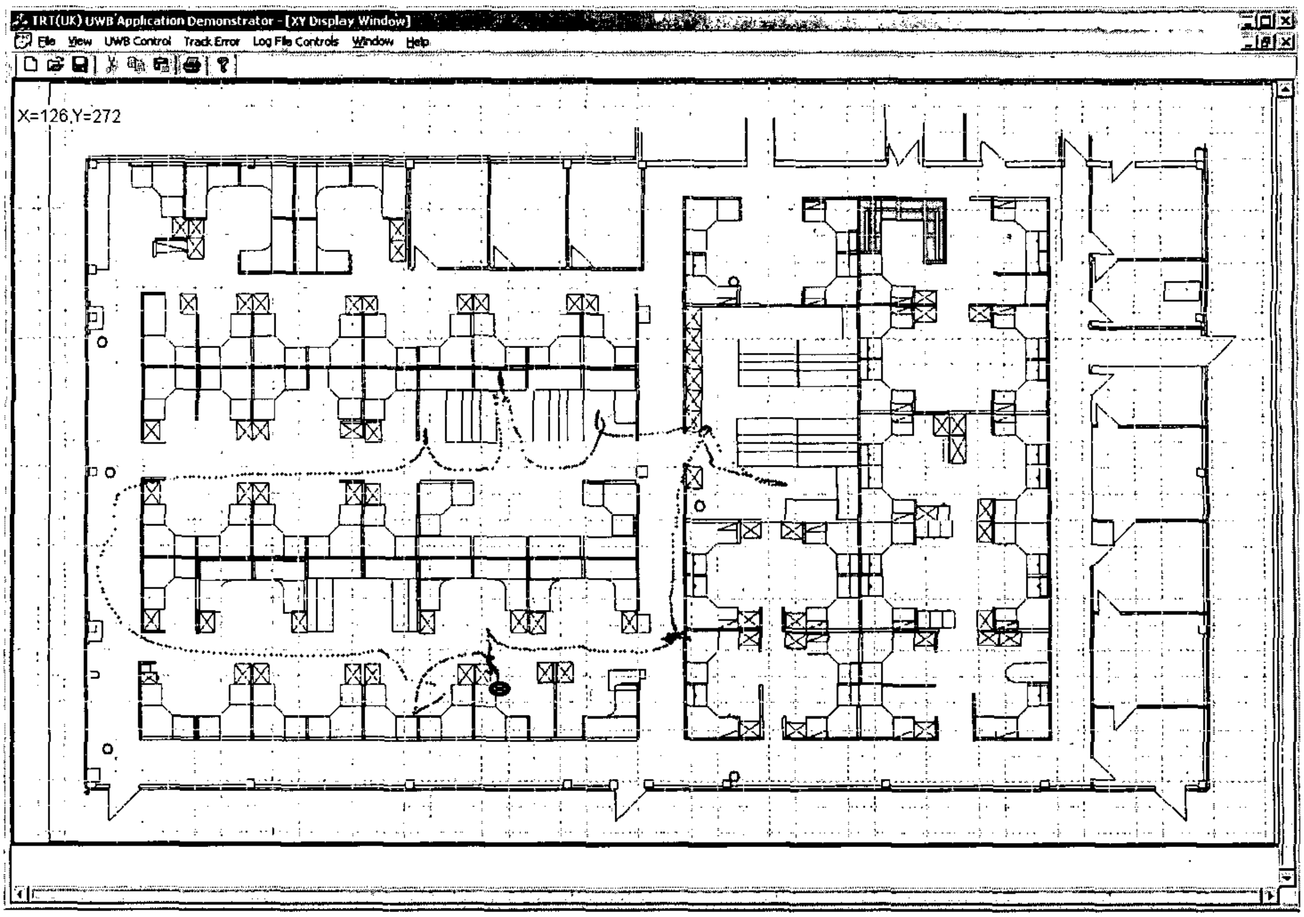

Fig. 6. Sample position tracking display

VII. CONCLUSIONS

Indoor positioning is a demanding application, for which UWB ranging offers a particularly good solution. Two kinds of use have been identified; as permanent building infrastructure, and rapidly deployed in emergencies. Different forms of UWB system are found to be appropriate in these two cases.

REFERENCES

[1] FCC 02-248; "In the matter of Revision of Part 15 of the Commission's Rules Regarding Ultra-Wideband Transmission Systems, FIRST REPORT AND ORDER", ET Docket 98-153, Released: April 22, 2002.

[2] T. B. Welch et al., "The Effects of the Human Body on UWB Signal Propagation in the Indoor Environment," *IEEE J. Select. Areas Commun.*, vol. 20, pp. 1778-1782, December 2002.

[3] Stuart Ingram, Jean-Paul Sicard, Ewan Frazer, and Dave Harmer, "Ultra Wide Band Positioning as an Indoor Extension of GNSS", *GNSS 2003 The European Navigation Conference*, Graz, 22-25 June 2003.