Chapter IV
RF Ranging Methods and Performance Limits for Sensor Localization

Steven Lanzisera
University of California, Berkeley, USA

Kristofer S.J. Pister
University of California, Berkeley, USA

ABSTRACT

Localization or geolocation of wireless sensors usually requires accurate estimates of the distance between nodes in the network. RF ranging techniques can provide these estimates through a variety of methods some of which are well suited to wireless sensor networks. Noise and multipath channels fundamentally limit the accuracy of range estimation, and a number of other implementation related phenomena further impact accuracy. This chapter explores these effects and selected mitigation techniques in the context of low power wireless systems.

INTRODUCTION

In this chapter we will discuss techniques for estimating the range between wireless sensor nodes using radio frequency (RF) measurements. Localization is a two part process that can roughly be divided into a phase where the relationships between nodes are estimated (range or angle) and a phase where these relationships are used to estimate locations of the devices. RF ranging, one of the options for the first phase, will be the topic of this chapter. In particular, RF time of flight methods where RF propagation time is estimated will be considered in depth. Other ranging methods (ultrasonic, sonic, light) have been proposed and tested but they are all limited from widespread adoption. Ultrasonic and sonic signals have limited range and do not pass through obstacles well when compared to RF signals. Acoustic systems also require the addition of speakers and microphones that are cumbersome for most applications. Light based systems require line of sight and are typically directional. Radios are pervasive in WSNs, and adding an accurate ranging feature would enable location aware networks in ways that are not possible using other technologies (Pahlavan, 2002).

Ranging accuracy is limited by noise, multipath channel effects, clock synchronization, clock frequency accuracy, and sampling artifacts. Fundamental performance limits exist due to these error sources, and these limits will be discussed qualitatively and mathematically. Signal bandwidth is an important factor when considering performance limits, and the impact of varying bandwidth will be shown.

Ranging methods will be discussed in the context of how well they meet application requirements for accuracy, energy consumption, latency, and useful range, and these requirements will be based on sample wireless sensor network applications. The major commercial application is asset tracking and management in factories, hospitals and other large spaces, and some commercial systems are available for these applications. Other applications including network configuration will be considered.

A number of RF based ranging systems have been proposed and implemented. The most common is the Global Positioning System (GPS), but others including cellular phone based systems are also widespread. Currently, ultra-wideband techniques are starting to be demonstrated along with more advanced narrowband techniques. The methods used and performance capabilities and limitations in selected systems will be discussed.

APPLICATION REQUIREMENTS

The requirements of a localization system are dependent on the application. This section will discuss a few applications to determine requirements on accuracy, latency, useful range, and infrastructure complexity of a ranging system. The accuracy requirement is defined to be the maximum error between true and estimated position that is acceptable for some percent of all estimates. For example, if 80% of estimates must be accurate to within 2 m, then 20% of measurements can have larger error. It is important to understand that localization is probabilistic in that the environment among other factors randomly degrades the accuracy of a measurement. Latency is the time it takes from when a request for a location update is made to when the update is presented to the user for a single device in the network. The range requirement is roughly how large of a sphere must one make around any node to find at least 4 other nodes or infrastructure points in 3D and 3 infrastructure points in 2D. Infrastructure requirements impact the cost of a network, and this impact can be considered qualitatively.

Relationship between Range Accuracy and Location Accuracy

Location accuracy requirements are in terms of difference from estimated location to true location as opposed to range accuracy. Localization algorithms and network geometries differ in how ranging accuracy translates to location accuracy, and many range based localization methods are presented in this book. In order to address the link between location and range accuracy, we apply a common method of range based location estimation: the maximum likelihood estimate (MLE) of the location based on a set of range estimates. The MLE of the location is found by calculating probability density function (PDF) of the location based on each range estimate, multiplying the

PDFs together for each range estimate, and finding the point where the resulting joint probability is maximized. Consider the case where the PDF of the location given a range estimate is given by $f(r_{est}|r_{true})$. If n independent range estimates (r_{est1}, r_{est2},$_{\ldots}$,r_{estn}) are used to find the MLE of the location, then the joint probability distribution of the location is given by the product of the individual PDFs,

$$f(\{r_{est_i}\}\,|\,l) = \prod_i f(r_{est_i}\,|\,r_{true_i})$$

where l is the location. When $f(\{r_{est_i}\}\,|\,l)$ is maximized, the corresponding location is the MLE. Figures 1 and 2 show the results of a random simulation of one simple 2D case when $f(r_{est}|r_{true})$ is normally distributed with parameters ($\mu=r_{est}$, σ). In Figure 1 the cumulative distribution function (CDF) of the location error normalized to the root mean square (RMS) ranging error is plotted when there are 3, 4 and 5 reference points. In Figure 2 the CDF of the location error normalized to the worst case ranging error is plotted. When more than 3 reference nodes are available, performance improves significantly especially when compared to the worst case ranging error. From this simulation two conclusions result: 1) increasing the density of nodes with known location is important for improving accuracy; 2) ranging accuracy and location accuracy are very similar. Although the location accuracy can be better or worse than the ranging accuracy depending on the conditions and localization algorithm used, we will assume that location error is equal to the RMS ranging error for simplicity.

Asset Tracking

In the hospital environment, equipment, staff and patients could all be tagged to increase the efficiency and safety of the healthcare environment. There are many cases in which hospitals own many extra pieces of equipment in hopes of ensuring that the appropriate items can be located and used quickly. Despite this preventive measure much time is often wasted searching for equipment. Because wasted time is so costly in terms of both dollars and care, this environment would benefit significantly by location aware devices. Everything must be monitored occasionally without tight latency requirements,

Figure 1. CDF of location error normalized by the RMS ranging error

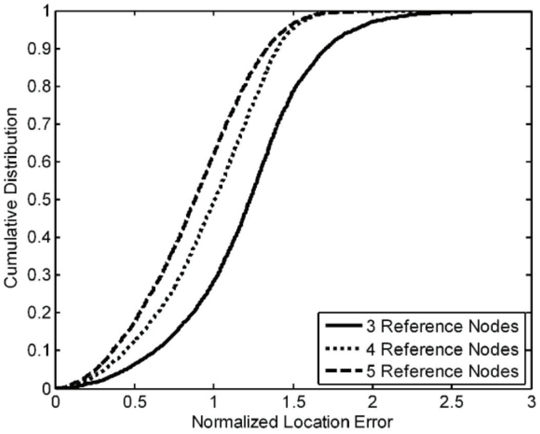
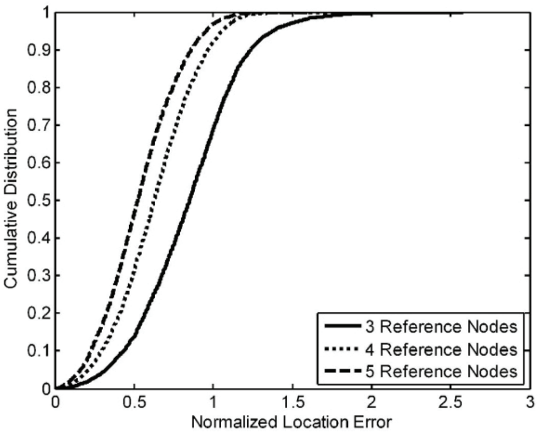


Figure 2. CDF of location error normalized by the worst case ranging error



but short latency updates of specific items are required. Accuracy must be good enough to ensure that the correct room is shown almost all of the time. Given that a typical hospital room is about 4 by 7 m, accuracy of better than 1.5m ensures the correct room is indicated 80% of the time. Alarms or query targets must be localized within a few seconds, and only one or two devices may be in a room. In order to ensure enough connectivity for localization, a range of 15 m is required.

Large Data Collection Network Configuration

A primary cost of deploying a large scale wireless sensor network is the installation of nodes and recording node locations. Localization systems can reduce this cost by determining the locations of devices after deployment. Latency requirements are minimal in that it is acceptable for the initial network configuration to take hours to complete. The scale of many industrial campuses requires long ranges in the hundreds of meters but accuracy requirements depend on the location of the device. Devices outdoors can be located with less accuracy and longer range, and indoor devices are more densely populated and require the typical 1.5 m accuracy.

Security

Security systems such as radio frequency identification (RFID) operated systems are commonly used to grant privileges (e.g. room and building access), and localization systems will be able to enhance these capabilities. If the correct person or people are in the correct rooms, privileges can be granted or revoked to ensure a secure environment for sensitive information, prison populations and many other situations. Latency must be on the scale of a second, and accuracy must ensure correct room identification (Anjum, 2005).

Summary of Specifications for Ranging Systems

Location accuracy, latency, range and infrastructure complexity are quite consistent across a broad spectrum of applications. For most networks, inter-node ranges are a few tens of meters and accuracy of 1.5 m with latency of a few seconds provides a robust solution. Much higher accuracy may be required in some applications not discussed here, but it isn't all that common. Infrastructure points, or nodes, can vary in cost by orders of magnitude depending on the ranging method used, and reducing the cost of these points is important to a successful location aware wireless sensor network.

Table 1. Summary of ranging specifications for typical indoor and outdoor sensor networks

Specification	Value	Conditions
Accuracy	1.5 m	80% of estimates indoors
	5 m	80% of estimates outdoors
Range	>15 m	Indoors, through walls
	100 m	Outdoors, line of sight
Latency	< 5 s	Including data relay across network
Infrastructure Cost	Low	

ACCURACY LIMITS

The achievable accuracy of ranging systems is limited by four primary factors which are noise, time synchronization, sampling artifacts, and multipath channel effects. These factors introduce random, time and spatially varying errors into the estimate resulting in reduced accuracy. Frequency accuracy between the devices involved in the measurement can also impact ranging system accuracy significantly. Each effect can dominate the error under different circumstances, and a system must be designed so that the combination of these effects does not degrade accuracy beyond useful limits. Because the introduced errors are stochastic, the errors can never be eliminated, but it is possible that measurement techniques can be used to mitigate these effects.

Noise

Noise and interference introduce unknown errors into measurements. The effect of white noise processes such as thermal and electronic noise is well understood and can be quantified. A range measurement degraded only by noise is limited in accuracy by the signal energy to noise ratio at the receiver and the occupied bandwidth.

A ranging system suffers in low signal to noise ratio (SNR) environments because the exact time of an event cannot be resolved precisely. In a simple example "edge detection" ranging system, the ranging signal is a step function sent by the transmitter at $t = 0$ and the receiver measures the time of the rising edge it observes. When this signal is received, the edge time may be detected slightly early or slightly late due to noise added to the signal. For RF measurements radio waves move at the speed of light ($c = 3\times10^8$ m/s) meaning that a distortion of just 10ns results in 3m of measurement error. The speed of this rising edge at the receiver is proportional to the bandwidth of the communications system, and wider bandwidth typically results in better performance. Because the noise amplitude increases as the square root of bandwidth and the signal transition speed increases linearly with bandwidth, a faster rising edge is more tolerant to noise. This qualitative understanding of how SNR and bandwidth affect the noise performance of ranging is useful, but a quantitative limit of ranging accuracy in a noisy environment is needed.

The mathematical expression that links SNR and bandwidth together to give a bound on ranging performance can be derived from the Cramér-Rao lower bound (CRB). The CRB can be calculated for any unbiased estimate of an unknown parameter. Van Trees (1968) discusses ranging as a parameter estimation problem studied in the context of radar and sonar applications, and the CRB under a variety of conditions has been calculated. For the prototype "edge detection" ranging system discussed earlier, the CRB can be used to calculate a lower bound for the variance of the estimate for the range, $\hat{r}$, as

$$\sigma_{\hat{r}}^2 \geq \frac{c^2}{(2\pi B)^2 \frac{E_s}{N_0}} \left(1 + \frac{1}{\frac{E_S}{N_0}} \right) \qquad (3)$$

where σ_{r}^2 is the variance of the range estimate, c is the speed of light, B is the occupied signal bandwidth in Hertz, and E_s/N_0 is the signal energy to noise density ratio. The SNR is related to E_s/N_0 in that

$$SNR = \frac{P_s}{P_n} = \frac{E_s}{N_0 t_s B} \qquad (4)$$

where P_s is the signal power, P_n is the noise power, t_s is the signal duration during which the bandwidth, B, is occupied. The concepts of occupied bandwidth and signal duration are important as illustrated by our step function example. The maximum bandwidth of the signal is set by the transmitter filter, and increasing the receiver's filter bandwidth does not increase the bandwidth used by the signal. Similarly, t_s is not simply the length of time that the signal was observed at the receiver but the length of time that the signal was observed when it was doing anything meaningful (such as changing in value). In the case of this step function, a small window of time contains nearly all of the useful information about the transition, and observing the signal for a longer period time contributes almost no additional information. In this example and in many common signals, the bandwidth and duration are tied together such that $t_sB \approx 1$. Therefore, the E_s/N_0 ratio is approximately equal to the SNR. By exchanging the locations of the factors in (4),

$$\frac{E_s}{N_0} = t_s B \cdot SNR \tag{5}$$

one advantage of having a t_sB product greater than unity becomes clear. Signals with this property would exhibit better noise performance at lower SNR values. One class of signals that exhibit this property are pseudorandom number (PN) sequences that result in long duration while retaining the same bandwidth as the constituent sub-symbols. These sub-symbols are called chips to differentiate them from bits (information) and symbols (collections of bits). Taking advantage of signals with $t_sB > 1$ improves noise performance, but it comes at the cost of increased signal processing. Often there is no other way to improve noise performance (i.e. fixed transmitter output power and receiver noise floor), and the signal processing cost is acceptable. For a fixed signal energy and noise density, increasing the bandwidth provides significant improvements in noise performance. This fact is one argument for increasing the bandwidth of RF based ranging systems, but the bandwidth required to achieve reasonable noise performance is not very large (Lanzisera, 2008).

One common example can be found in GPS. The C/A (course acquisition or civilian) signal in GPS uses a PN sequence modulated with binary phase shift keying (BPSK) at 1.023×10^6 chips/s. At a receiver on the ground, the observed SNR is typically -20dB, the bandwidth occupied is about 2MHz, and there are 1023 chips per symbol (Kaplan, 2005). This is all the information required to determine the best case noise performance of GPS. First we calculate E_s/N_0 through the application of (5):

$$\frac{E_s}{N_0} = \frac{1023}{1.023\times10^6}\cdot 2\times10^6 \cdot 10^{-2} = 20$$

Applying this result to (3)

$$\sigma^2_{\hat{r}_{GPS}} \geq \frac{(3\times10^8)^2}{(2\pi\cdot 2\times10^6)^2\cdot 20}\left(1+\frac{1}{20}\right) = (5.5m)^2$$

This accuracy is close to what GPS routinely provides, but this range estimate is updated at 1kHz in the above calculation, and the typical user uses systems that update at less than 10 Hz. This can be used to reduce the variance by a factor of 100 resulting in $\sigma^2_{\hat{r}_{GPS}} \geq (0.6m)^2$. GPS users are accustomed to accuracy of better than 5m (80% of trials) in open, flat terrain suggesting that the noise limit is not obtained or that other factors are reducing accuracy. In this case, approaching the CRB is possible because of the high value of E_s/N_0 and the signal design, but random atmospheric effects contribute the majority of the remaining error. The P (precise or military) GPS signal is broadcast at two different carrier frequencies so that these atmospheric effects can be estimated and removed which greatly enhances accuracy. It is also worth noting that 1+ E_s/N_0 term contributes very little to the CRB, and it is commonly ignored for E_s/N_0»1.

GPS provides a good reference for looking at other ranging systems because it is familiar and has some characteristics in common with communications systems, but it has significant differences as well. In typical wireless communications systems, the distances traveled are much less, and atmospheric effects are not significant. In addition, narrowband systems have signal SNR that is large such that, when coupled with processing gain, high values of E_s/N_0 result. These high values for E_s/N_0 allow the CRB to be nearly achieved in many systems, but the CRB is not a tight bound at low E_s/N_0 (Van Trees, 1968). If the desired error variance is not achievable directly, averages of multiple measurements will yield improved results. GPS occupies a 2MHz bandwidth which is comparable to the common IEEE 802.15.4 radios used in WSNs, but GPS signals are broadcast at a single carrier frequency. WSN radios are usually frequency agile, and information from different frequencies can be used to improve ranging performance (Lanzisera, 2006).

The CRB can also be improved through the use of additional bandwidth. Ultra wideband (UWB) technologies are being developed partially to provide accurate ranging capability to wireless systems. A UWB signal is defined to be a signal that either uses at least 500MHz or that occupies as much bandwidth as half the signal's center frequency. The use of 500MHz of bandwidth and an E_s/N_0 of -10dB yield a CRB of

$$\sigma_{\hat{r}}^2 \leq \frac{(3\times10^8)^2(1+\frac{1}{0.1})}{(2\pi\cdot500\times10^6)^2\cdot0.1}=(1m)^2$$

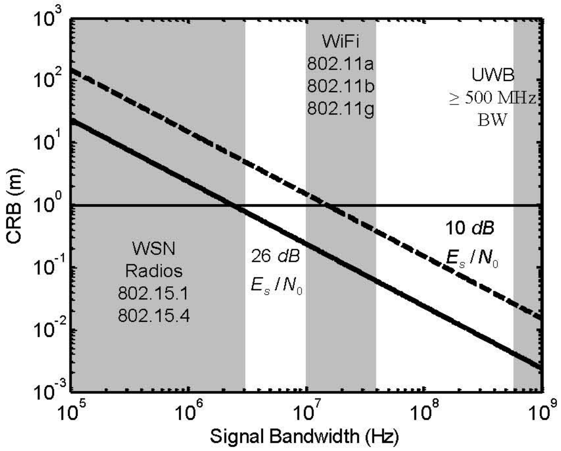
Although the CRB may not be achievable at this low value for E_s/N_0, small bounds are possible. This promise, along with superior performance in multipath environments (to be discussed later), has driven much interest in UWB for extremely accurate location systems.

Both bandwidth and E_s/N_0 play significant roles in determining noise limited performance. Figure 3 shows the CRB as a function of bandwidth for E_s/N_0 of 10 dB and 26 dB. Signals with t_sB products of 10 to over 1000 are commonly used enabling large E_s/N_0 in communication systems. It is interesting to note that that noise alone does not prevent 1 m accuracy for bandwidths of a few megahertz or more.

Time Synchronization

RF time of flight measurement systems must be able to estimate the time of transmission and arrival using a common time base for accurate measurements. When two wireless devices, A and B perform range estimation, the most straightforward method is for A to send a signal at $t = 0$ and for B to start a timer at $t = 0$ and stop it when it receives the signal sent by A. The value of the timer at B is equal to the time of flight (TOF). If the clocks are not perfectly time synchronized, however, and B's notion of $t = 0$ is offset in time from A's, then this offset, Δt, directly adds a bias to the measurement. Time synchronized wireless networks are typically synchronized to on the order of 1μs resulting in errors of up to 300m, but high power and expensive systems can achieve time synchronization of better than 10ns or 3m. This method is shown in Figure 4a.

Figure 3. Cramér Rao Lower Bound as a function of bandwidth for 10dB and 26dB E_s/N_0. Common radio standards used in wireless sensor networks such as IEEE 802.15.1 (Bluetooth), IEEE 802.15.4 (Zigbee and others), and wireless LAN (802.11a/b/g) are shown. Ultra-wideband (UWB) radios with more than 500MHz of bandwidth have excellent noise performance, but even a few megahertz of bandwidth can enable the 1.5m accuracy required for most applications.



If A and B have full duplex radios, that is, they can transmit and receive at the same time, then a two way or round trip measurement can be made. A sends a signal to B at a carrier frequency f_{c1} and B translates this signal to a different carrier frequency f_{c2} and retransmits that signal in real time. The signal is received back at A at f_{c2} such that A can compare the signal it is receiving from B to the signal it is sending to B. By measuring the delay between these two signals, the round trip TOF, $\hat{\tau}_{RT}$, is estimated, and the range estimate is $c \cdot \frac{\hat{\tau}_{RT}}{2}$. This method is shown in Figure 4b.

Most WSN nodes do not have full duplex radios because they are more complicated and expensive than half duplex transceivers. Many other wireless systems are half duplex as well (e.g. wireless LAN and GSM), and the round trip method can be adapted for these systems. A round trip method known as two way time transfer (TWTT) has been developed to improve time synchronization between wireless base stations after the first communications satellites were launched, and it provides both range estimation and improved time synchronization capability (Kirchner, 1991). This method, shown in Figure 4c, allows the time offset between A and B to be ignored. Both A and B are responsible for measuring a time delay accurately using a local clock. A must measure the time that it takes for the signal it sends to return to it, and B must measure the time that the signal spends at B accurately. If the time A sends the signal is t_{sA}, the time B receives the signal from A is t_{rB}, the time B replies to A is t_{sB}, the time A receives the signal is back from B is t_{rA} such that $t_{sA} < t_{rB} < t_{sB} < t_{rA}$ then A measures $t_A = t_{rA} - t_{sA}$ and B measures $t_B = t_{sB} - t_{rB}$. By combining these two measurements together both the time of flight (τ) and clock offset ($\Delta\hat{t}$) can be estimated.

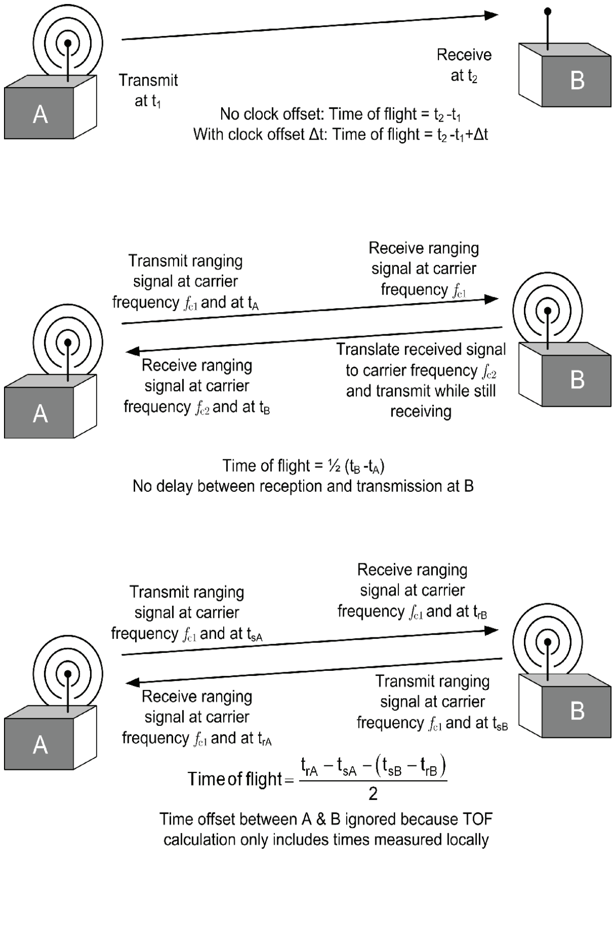
$$\Delta\hat{t} = \frac{1}{2(t_A + t_B)} \qquad (6)$$

$$\hat{\tau} = \frac{1}{2(t_A - t_B)} \tag{7}$$

This or related methods are used with less accurate hardware to provide the rough time synchronization common in wireless systems.

One problem with two way ranging is that the measurement takes place over a relatively long period of time such that if the reference clock frequencies at the two nodes are not identical, an unknown bias can be added to the signal. In WSN nodes, inexpensive crystals are used where the frequency difference from crystal to crystal may be 100ppm or more across commercial temperature ranges. This clock frequency offset (also called clock drift) error must be mitigated in some fashion (Lanzisera, 2006). Consider a system where the time spent sending a ranging signal is 100μs and the time spent changing from transmit to receive mode is 200μs, and the time spent receiving the returned ranging signal is 100μs. Over this 400μs time, a clock frequency mismatch of just 10ppm would result in about 1m of estimation error. The clock frequency offset can be measured, and then the clock frequency can either be corrected to match within bounds or the resulting error can be calculated and subtracted from the

Figure 4. Three methods of performing time of flight ranging measurements: (a) time of arrival which is susceptible to clock offset Δt; (b) full duplex two way ranging; (c) half duplex two way ranging called two way time transfer


A
Transmit at t_1
Receive at t_2
B
No clock offset: Time of flight = $t_2 - t_1$
With clock offset Δt: Time of flight = $t_2 - t_1 + \Delta t$
Transmit ranging signal at carrier frequency f_{c1} and at t_A
Receive ranging signal at carrier frequency f_{c1}
Receive ranging signal at carrier frequency f_{c2} and at t_B
Translate received signal to carrier frequency f_{c2} and transmit while still receiving
A
B
Time of flight = ½ $(t_B - t_A)$
No delay between reception and transmission at B
Transmit ranging signal at carrier frequency f_{c1} and at t_{sA}
Receive ranging signal at carrier frequency f_{c1} and at t_{rB}
Receive ranging signal at carrier frequency f_{c1} and at t_{rA}
Transmit ranging signal at carrier frequency f_{c1} and at t_{sB}
A
B
Time of flight = $\frac{t_{rA} - t_{sA} - (t_{sB} - t_{rB})}{2}$
Time offset between A & B ignored because TOF calculation only includes times measured locally

estimate later. Many methods have been used to measure frequency offsets in wireless systems, and we summarize one simple method here. This method is to run a counter over a long period of time to measure the offset. One node sends a start packet to the second node and starts a local timer, and the second node starts a local timer when it receives this packet. After waiting a sufficiently long time, the timer at the first node expires, and it sends a stop packet. The second node receives this stop packet, stops its timer, and compares the value left on the timer to the expected value (zero if the counter is counting down). This difference is a measure of the clock offset. The minimum time between packets, T_{wait} can be calculated as follows:

$$T_{wait} \geq \frac{1}{\Delta f_{xo}} \tag{8}$$

where Δ is the required matching , and f_{xo} is the crystal frequency. For a 20MHz crystal and a system requiring 10ppm accuracy, T_{wait} must be great than 5ms. This process is rather long but very simple, and other methods trade complexity for time savings.

Sampling Artifacts

Ranging systems estimate the time of arrival of a signal and compare that time with the time the signal was transmitted to calculate the time of flight and thus the range. It is commonly assumed that ranging accuracy is limited to c/f_s where f_s is the receiver sampling rate (Richards, 2005). This limit is known as range binning, and it can impact resolution if steps are not taken to mitigate its impact. A common implementation is to estimate the time of arrival using a matched filter that is sampled at the signal bandwidth resulting in time resolution of $1/B$. This sampling adds error to the estimate because the estimate space is divided up into range bins that are c/B wide. The error associated with this process is uniformly distributed inside the range bin. By using the variance of the uniform distribution, the impact of sampling can be found (Hoel, 1971).

$$\sigma^2_{sample} = \frac{1}{12 \cdot f_s^2} \tag{9}$$

In the case of the GPS example, with sampling at $1/B$ the variance due to sampling can be calculated.

$$\sigma^2_{sample} = \frac{1}{12 \cdot (2 \times 10^6)^2} = (144ns)^2$$

This results in a range resolution of 43 m. In GPS, this coarse estimate is filtered (averaged) to improve the resolution, and a feedback loop can be used to null out the sampling error while the receiver tracks the satellites (Kaplan, 2005). Using just averaging, over 1500 measurements are required to achieve a variance of $(1\text{m})^2$. These methods are not realistic for many wireless sensor network applications where extremely low power consumption and therefore duty cycle is required. An accurate range estimate must be made in a short period of time. To reduce the sampling error, the signal can be over sampled. Figure 5 shows the CRB for a 2 MHz bandwidth signal with E_s/N_0 of 26 dB, the standard deviation of the range error due to sampling, and the combined effect of both error sources as a function of sampling frequency. This plot shows that with a 2 MHz bandwidth, the required sampling rate to ensure that the error is not dominated by sampling is over 70 MHz. It is clear that one must sample very fast to have the error dominated by the CRB rather than sampling. As the CRB improves due to increased bandwidth, the sampling speed required remains higher than twice the signal bandwidth down to E_s/N_0 of about 3 dB.

If the signal is sampled above Nyquist (f_s>2B), then the entire information content of the signal is captured in the sampling process (Oppenheim, 1975). Therefore, it is possible to extract better time resolution than σ_{sample}. In Figure 6, a signal is shown along with dots representing the samples of that signal that is band limited to a 2 MHz bandwidth. This signal is sampled at 10 Msps which is above the Nyquist rate of 4 Msps, but the sample rate still is far too low to achieve the CRB. The range bins are 100ns (30m) wide in this case where as the CRB from Figure 3 is only 3.5ns (1.1m) demonstrating a dramatic resolution reduction. Looking at the time of the zero crossing, it is clear that even a linear interpolation between the two adjacent samples would improve the estimate of that zero crossing location significantly. A major challenge is that many systems would need to perform this interpolation in real time increasing system complexity and power consumption beyond reasonable limits.

A round trip time of flight method known as code modulus synchronization (CMS) that takes advantage of Nyquist sampling has demonstrated its ability to approach the CRB while maintaining low sampling rates. CMS emulates a full duplex ranging system where the repeating node is retransmitting the signal that it is receiving from the first node without any delay. In CMS, however, half duplex radios such as those used in wireless sensor networks are used so the delay between reception and retransmission must be managed carefully. CMS as implemented uses a short PN code modulating an RF carrier as the ranging signal. Figure 7 shows the basic operation of CMS for a time of flight of zero. When the time of flight is greater than zero the Node B and Node A Receive lines would each be circularly shifted to the right by an amount equal to the time of flight and twice the time of flight respectively. For example, a range of 9m would have a time of flight of 30ns. The second line would be shifted by 30ns and the Node A receive line would be shifted by 60ns but the chip period may be 500ns, and it is acceptable that the shifts are much smaller than the chip period. The first node, A, generates a local code that is synchronized with a local clock called the event clock that has the same period as the PN

Figure 5. A comparison of range binning due to sampling error and the Cramér Rao bound on noise limited ranging for a 2 MHz bandwidth with a E_s/N_0 of 26 dB. The sampling rate required is much higher than required by sampling theory to achieve noise limited resolution.

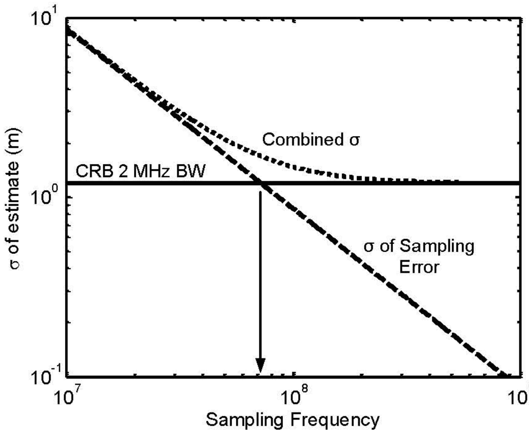
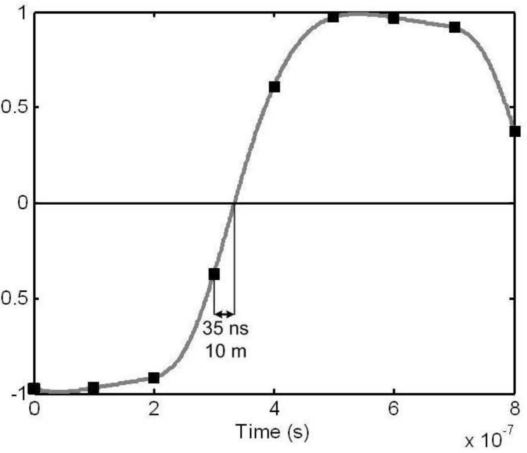
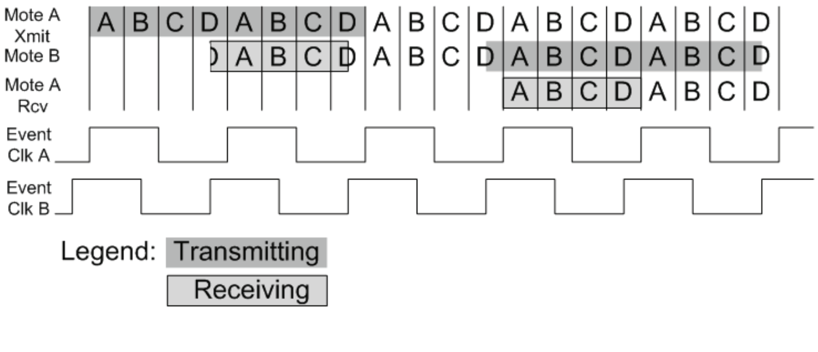


Figure 6. An above Nyquist sampled waveform is shown with the sample points marked in an example of sample based range binning. An interpolation between points enables time resolution of the zero crossing far better than 1/B and 1/f_s reducing the size of the range bins significantly.



code. This code is used to modulate the carrier and is transmitted to the second node, B. B has a local event clock with the same period as at A, but the phase of the clocks are offset. As a result, B knows the length of the incoming PN code. B samples and demodulates this signal, and exactly one circularly shifted copy of the code is stored in memory. The system can accumulate multiple copies of the code in order to improve SNR, but they are all exactly one copy of the code that is circularly shifted in exactly the same way as the other received copies. At this point, B has a local copy of the code that is an average of multiple receptions and that is circularly shifted due to the event clock phase offsets between A and B. After A has sent a predetermined number of code copies and B has received some of these copies, the transceivers switch states, and B is now the source of the code. Starting on its event clock rising edge, it transmits the circularly shifted code it received back to A. On the next rising edge of its

Figure 7. Code modulus synchronization (CMS) achieves noise limited ranging performance through interpolation of data points in an non-real time time of flight (TOF) estimation phase, and this figure shows the case where the TOF is zero. Non-zero TOF would result in sub-chip width circular shifts to the signals on the node B and node A receive signals. CMS is a two way ranging technique that emulates a full duplex ranging system (Figure 4b) to eliminate clock synchronization errors.

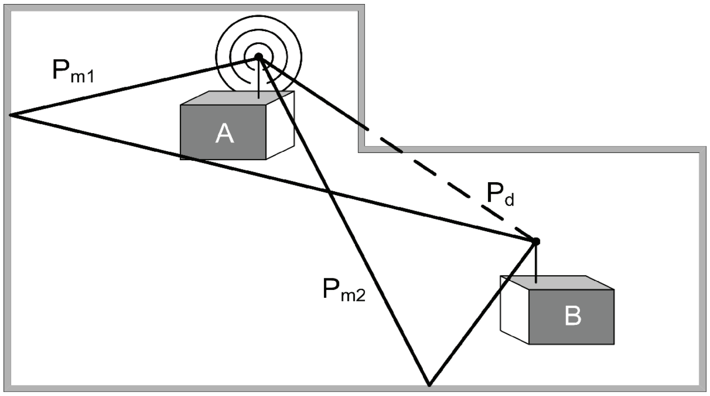


event clock, A starts to record exactly one copy of the code. Again, A can accumulate multiple copies to improve SNR. Because of the roundtrip nature of the system, the circular shift that occurred going from A to B is exactly undone going from B to A. After A has received and accumulated the desired number of code copies, the transceivers are shut off, and all of the real time processing is completed. A then computes the cross correlation between the code it recorded and the code that it sent, and zero code offset exists if the time of flight is zero. Because this system relies on sampling the signal at or above Nyquist, the received code can be interpolated to improve resolution up to the noise limit of the system. The correlation and code offset estimation are not done in real time enabling the computation to be done at any time using any method the user desires. This system can achieve the CRB in a single measurement as long as the sampling rate of the received code is above Nyquist, substantially improving over other two way ranging methods (Lanzisera, 2008).

Multipath Channel Effects

When a ranging system has been well designed, it often still fails to achieve the expected performance because the measurement is not taken in free space. In real environments the RF signals bounce off objects in the environment causing the signal to arrive at the receiving antenna through multiple paths as shown in Figure 8. In this figure, the direct path is obstructed by walls, but the other paths are not. This is common indoors, and it is likely that the non-direct paths have higher power than the direct path (Spencer, 2000). The communication environment is called the channel, and multipath channels not only vary by the type of environment (office building, residential or outdoors) but are specific to the geometry of the transmitter and receiver in that environment. The channel is often time varying resulting in a multipath environment that changes from one time to another. For narrowband radios like those common in wireless sensor networks, moving one transceiver by just a fraction of a wavelength (12cm at 2.4GHz) will cause the receiver to see what looks like an entirely new multipath environment because the paths will interfere constructively or destructively differently. The path length change is referenced to the wavelength of the RF making these small changes have large effects. The speed that the channel changes depends on how quickly objects are moving in that environment. Slower objects

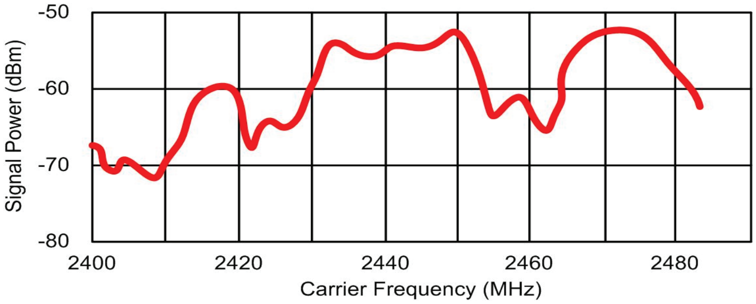
Figure 8. A multipath environment that exhibits a common condition. The direct path (P_d) which is to be estimated for ranging is obstructed and heavily attenuated while the reflected paths (P_{m1}, P_{m2}) have much higher signal power.



result in slower changes to the channel. This typically means that indoor channels change more slowly than outdoor channels, and the time it takes for the channel to change significantly is called the coherence time, t_c, of the channel. The value of t_c is roughly $c/(2fv)$ where c is the speed of light, f is the carrier frequency, and v is the speed of the fastest moving object in the environment. Recall that the wavelength of radio waves, λ, is c/f, and a more intuitive form of t_c is $\lambda/(2v)$ where it is clear that the time it takes to move a half wavelength corresponds to the coherence time (Tse, 2005). A series of measurements that take much less than t_c to complete can be used together as if the channel was time invariant over those measurements. This fact is useful when attempting to reduce the impact of multipath because multiple measurements taken at different frequencies can be used together. Because this interference effect is closely tied to the wavelength, changing carrier frequency even by 1% or less can dramatically affect the apparent multipath environment in narrowband systems. This can be easily observed by considering the received signal strength (RSS) profile across carrier frequency in an indoor environment as shown in Figure 9 (Werb 2005). At some carrier frequencies, the signal is in deep fade (destructive interference), while at others it has much higher signal strength (constructive interference). Without knowing the channel characteristics, knowledge of the RSS at one frequency tells you nothing about the RSS at another frequency. Wider bandwidth signals suffer less from this effect, and the bandwidth required to combat this is related to the time difference between the first and last significant path arrivals known as the delay spread, t_d. The coherence bandwidth, W_c, is approximately $1/(2\pi t_d)$ and it is the bandwidth over which the channel can be considered to be flat (either in deep fade or not, for example). If the bandwidth, B, is much larger than W_c the signal does not depend on carrier frequency to the same extent as a signal with a bandwidth less than W_c (Tse, 2005). Typical delay spreads for indoor channels are between 10ns and 100ns yielding coherence bandwidths between 1 MHz and 20 MHz. Outdoors, the delay spread can be up to microseconds, significantly reducing W_c . In RF ranging systems, the inter-path delay, $t_{\Delta p}$, is more important than the delay spread, however, because short inter-path delays can significantly impact ranging accuracy. Indoors, inter-path delays of 5ns to 10ns are very common and must be resolved if accuracy is to be better than $c \cdot t_{\Delta p}$ (Van Trees, 1968).

In a multipath environment, the receiver must somehow choose or estimate the direct path length and ignore the other paths. If a receiver can determine when only the first path arrives, then this will be the shortest distance and desired estimate. If the system is not able to resolve the individual paths, then the estimate is blurred by the multipath effects resulting in measurement error. In this case, if the receiver

Figure 9. Received signal strength verses frequency measured in a line of sight multipath channel with a 2MHz RF bandwidth. The significant changes in signal strength show that changing carrier frequency changes the apparent multipath environment significantly. Adapted from Werb (2005).



has an estimate of the channel impulse response, it can calculate the bias caused by the multipath channel and subtract the bias from its estimate. This leads to two classes of multipath mitigation methods: 1) resolving the direct path through increased bandwidth, or 2) estimating the channel response and using this information to improve or generate a range estimate.

In the first case, the ability to resolve the response of the multipath channel is directly linked to the bandwidth of the signal. Inter-path delays, $t_{\Delta p}$, separated by more than $1/B$ in time are resolvable and paths separated by less are generally not. To resolve paths that are separated by 1m or more, a bandwidth of at least 300MHz is required which shows a significant advantage of UWB systems. Using bandwidths in excess of 500MHz enables accuracy better than 1m in many cases, but this accuracy is not always achieved (Shah, 2005). Sometimes there is line of sight between the transmitter and receiver, or, in other cases, the direct path is attenuated somewhat by obstacles but still reaches the receiver with sufficient strength to be resolved, resulting in acceptable accuracy. When the direct path is too weak compared to other paths, however, a secondary path will be chosen to estimate the range resulting in an over estimate. In indoor environments, 10% to 20% of all measurements will fall into this category, but some environments are worse and a direct path is rarely available. True line of sight paths are not very common indoors, and most indoor channels will have a few strong paths spread across a few tens of nanoseconds (Spencer, 2000). Localization systems typically mitigate the severe cases of obstructed ranging by adding extra devices to "see" the obstructed areas and through localization algorithms that reject large ranging errors.

The second mitigation strategy relies on estimating the impact of the multipath environment on the range estimate and then subtracting off this error. This method is used when the signal bandwidth is too small to sufficiently resolve the multipath environment, and it is somewhat analogous to channel equalization. There are two critical steps to this method: 1) estimating the channel frequency/impulse response and 2) estimating the impact the channel has on the range estimate. Each step can be completed in different ways, and the solutions fall into the family of super-resolution algorithms. A super resolution algorithm is one that attempts to provide range resolution that is better than c/B (Dickey, 2001). If the impulse response can be estimated to include the static offset due to the time of flight, then the range can be estimated directly from the impulse response. If the impulse response is estimated with the first path always being at a delay of zero, then some other ranging method must also be used. One method to estimate the channel impulse response is to send a modulated signal that consists of a sequence of chips (Nefedov, 2000). Recall that the inter-path delay is a few nanoseconds compared to the chip duration of 100's of ns to µs, and the chip width used must typically be shorter in time than the features to be resolved. A super-resolution technique resolves features that would be too close together in time to be resolved normally. If the signal sent is x, the channel impulse response is h, and the received signal is y, then

$$y = x * h + \widetilde{n}$$

Where * denotes convolution, and $\tilde{n}$ is complex noise. This can be rewritten in the frequency domain.

$$Y(\omega) = X(\omega)H(\omega) + N(\omega)$$

If the signal to noise ratio is large, and the spectrum of the transmitted signal (including the transmitter frequency response) is known, then $H(\omega)$ can be approximated.

$$H(\omega) = \frac{Y(\omega)}{X(\omega)} + \frac{N(\omega)}{X(\omega)} \approx \frac{Y(\omega)}{X(\omega)}$$

This approximation is only valid in sufficiently high SNRs, and noise causes significant problems in super-resolution estimation methods. *Y*(ω) is calculated by taking the Fourier transform of the received signal, and *X*(ω) is a system parameter known a priori. Once *H*(ω) has been estimated, *h*(*t*) must be estimated. The inverse Fourier transform will solve this problem, but a number of substantially more complicated algorithms exist that provide better time resolution. Examples of such algorithms include Multiple Signal Classification (MUSIC) and matrix-pencil methods that have been developed for use in imaging and radar systems (Dharamdial, 2003; Song, 2004; Pahlavan, 2002). These algorithms achieve time resolution that is up to ten times better than the Fourier transform method when the SNR is high enough. Due to the narrowband nature of many radios used in WSNs (i.e. IEEE 802.15.4's 2MHz bandwidth), a resolution enhancement of even ten times may be insufficient to provide reasonable accuracy. Once the channel estimate has been made, an additional algorithm to estimate the impact of the estimated channel on a ranging measurement using TOF techniques (i.e. TWTT) can be used. Such an algorithm can include, to some degree, the effect of paths buried inside the estimated channel response, resulting in a good estimate of the range error. This error can be subtracted from the estimated range to achieve a better range estimate.

Summary of Performance Limits

In wireless sensor networks, the devices are resource and power limited, and efforts should be made to reduce the time the radio is active and reduce the amount of signal processing while preserving performance. The above discussions show that signal bandwidth is a system parameter of high importance. Increasing signal bandwidth improves noise and multipath performance linearly with bandwidth. The bandwidth required to achieve very fine resolution in a Gaussian white noise environment is far smaller than that required to achieve equivalent resolution in a typical indoor multipath environment, and the techniques to improve multipath performance are far more computationally intensive than those to combat noise. Many measurements in indoor environments will not have a resolvable direct path using any method or bandwidth, and the resulting range estimate will be highly inaccurate. Localization algorithms must deal gracefully with range measurements that are widely inaccurate some of the time. Methods to deal with other error sources such as synchronization and sampling exist and should be applied to minimize energy while maximizing performance. Although ultra-wideband systems are sure to provide high range resolution, the energy cost of data communication over an ultra-wideband radio remains very high compared to narrowband radios. Therefore, ranging methods that use small bandwidths are critical to many low power wireless networks, and methods to improve range accuracy given fixed, small bandwidths are an unsolved problem.

DEPLOYED SYSTEMS

The localization problem has seen widespread attention in the research community, and a number of RF range based methods have been proposed and implemented. Location information has significant value as represented by the E911 location requirements for cell phones (along with similar requirements around the globe) and the commercial tracking systems available. This section provides a brief survey of ranging techniques that could be applied to wireless sensor systems. Important characteristics of systems are the noise performance, suitability for indoor use, cost, and infrastructure requirements.

Received Signal Strength Range Estimation

The RF received signal strength (RSS) has been used as a surrogate for range measurement in many systems. In free space, the power of an RF signal can be calculated using the Friis transmission formula (Ulaby, 2004).

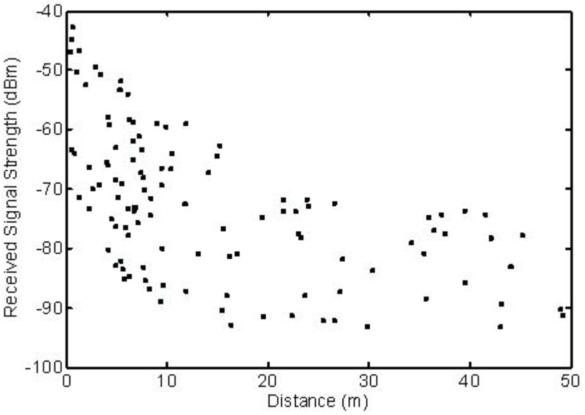
$$P_{rx} = \frac{P_{tx}}{(4\pi R/\lambda)^2}$$

The received power, P_{rx}, decreases as the range, R, squared, and there is a unique correspondence between RSS and range. Real environments have multipath channels, however, and the signal power does not behave predictably. Other models have been proposed such as the popular $1/R^{\alpha}$ model where the signal power now decreases as one over distance to the power α. Alpha is a fitting parameter that is environmentally dependent and is usually between 1 and 4. These models are unreliable because the power does not decrease monotonically with range and passing through walls causes sudden drops in power over short distances. The multipath environment causes areas of constructive and then deconstructive interference so that a user's location will be uncorrelated with signal power. This effect is frequency dependent as well, so a measurement at one carrier frequency will be uncorrelated with a measurement at another carrier frequency. Even if this multipath effect could be successfully mitigated, the effect of wall and obstacle attenuation prevents this method from providing reliable range estimates (Cheng, 2005). Figure 10 shows a plot of the received signal strength verses distance for an indoor environment. The same signal strength corresponds to more than half of the useful range of the radio, and this accuracy is typical for these systems. Range estimation error, when it can be quantified, is typically proportional to range such that short range measurements may be accurate within a few meters, and longer range measurements are less accurate. RSS measurement for ranging is often considered a near verses far technique that can provide some information regarding proximity but less about true range.

Most radios include a received signal strength indicator (RSSI), and the measurement is available to the user without any additional hardware or power costs which explains the technique's popularity. Because the RF ranging problem is challenging, RSS based techniques have received tremendous attention, and many RSS based localization systems have been proposed and implemented with varying degrees of success at turning poor range estimates into accurate location estimates.

Some RSS based systems use RSS "fingerprinting" techniques rather than RSS for ranging and achieve improved accuracy. The RSS at different carrier frequencies is recorded for many locations in the network through a site survey at the time of network deployment. In normal use the network tries to match the measured RSS of a mobile node with the fingerprint map it has stored to estimate location. Accuracy of these methods can be a few meters, but changes in the environment (an open door that was closed) can cause significant errors in location estimates (Lorincz, 2006).

Figure 10. Received signal strength plotted verses distance. A best fit does not capture the large deviation of data points showing that models with unique correspondence between range and signal power cannot provide reasonable accuracy.



Global Positioning System

The most widespread RF localization scheme is the global positioning system (GPS). A constellation of at least 24 GPS satellites orbit the earth and constantly transmit unique signals. User receivers take four or more of these signals and use them to estimate values for position and time. GPS has a coarse/acquisition signal (C/A code) that is used for civilian uses and to aid in the synchronization to the precise (P) code used by the military. The P code is encrypted to prevent general use, and is called the P(Y) code in the encrypted state. C/A code users generally enjoy location accuracy of better than 10 m on the ground with slightly less accuracy in the vertical direction as long as they have a clear view of the sky without any significant multipath effects. Because the C/A code only occupies about 2MHz of bandwidth at a single carrier frequency, multipath can greatly degrade performance. The received signal power on the ground is extremely low making it difficult to receive the signal when a clear view of the sky is unavailable (Kaplan, 2005). GPS receivers have become much lower power in recent years, but they still consume tens of milli-Joules for the first fix. The SiRFstarIII is a low power receiver with good low received signal power performance and consumes 50 mW typically while taking 5 s to provide a location after power on. With assistance from a cellular phone network, this fix can take 1 s, but now the cellular phone radio will dominate the energy consumption (SiRF, 2008). Given that many WSN applications require location updates at much lower rates than the 1Hz typical of GPS, GPS is far from power optimized. The price of the GPS units is also high due to the complexity associated with signal acquisition and processing.

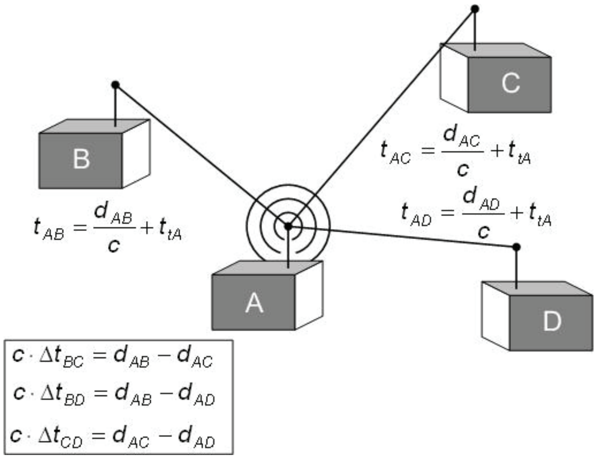
Time Difference of Arrival

Time difference of arrival (TDOA) is a powerful and commonly used technique that relies on time synchronized infrastructure to estimate the range of a mobile device. The most common scenario is where the mobile device transmits a signal that is simultaneously received by multiple base stations. These base stations estimate the time of arrival of the signal and compare the estimates among multiple stations to estimate the user location. For each receiver pair, a TDOA estimate can be made, and 3 pairs are required to determine a location. As seen in Figure 11, the time difference of arrival at the three base stations is a function only of the unknown distances. When the three measurements are made, a system of three equations with three unknowns results enabling the ranges to be calculated. The primary advantage of this system is that a mobile unit can be very simple because all of the complexity is at the base station. The disadvantage is that time synchronized infrastructure is required which increases cost and complexity of the overall network. Accuracy is linked not only to the environment but to the density of base stations thus requiring large numbers of expensive base stations to cover a network. As a rule of thumb, the density of base stations for TDOA must be four times the density required for data coverage. This technique enables the use of simple mobile devices that can periodically send a ranging signal to be detected by the always-on infrastructure providing a low energy location on schedule or on demand. This technique is not limited to a particular bandwidth or multipath mitigation scheme, and can provide highly accurate or poor performance indoors depending on the implementation. Three commercial systems will be discussed briefly below.

GSM Cellular Phone Networks

Time difference of arrival (TDOA) localization in GSM cellular phone networks has been standardized as part of GSM since 1999 (GSM 03.71 1999 and 2001). In the 1999 version, the handset sends network access packets that are received by three or more base stations which use TDOA to estimate the position of the handset. The 2001 version requires that the handset measure the time difference of arrival of signals sent from base stations. Location accuracy depends heavily on the number of base stations within communication range. In urban environments it is common to be able to communicate

Figure 11. Time difference of arrival (TDOA) uses time synchronized infrastructure nodes (B,C,D) to simultaneously measure the time of arrival of a signal transmitted by A. Because the transmission time, t_{tA}, is unknown, the time differences between arrivals, Δt, can be used to setup three equations to solve for the three unknown ranges (d_{AB}, d_{AC}, and d_{AD}).


C
B
$t_{AC} = \frac{d_{AC}}{c} + t_{tA}$
$t_{AD} = \frac{d_{AD}}{c} + t_{tA}$
$t_{AB} = \frac{d_{AB}}{c} + t_{tA}$
A
D
$c \cdot \Delta t_{BC} = d_{AB} - d_{AC}$
$c \cdot \Delta t_{BD} = d_{AB} - d_{AD}$
$c \cdot \Delta t_{CD} = d_{AC} - d_{AD}$

with more than three, and accuracy of better than 100m is common. When only two or less base stations are available, estimates can be many 100's of meters off resulting in an unreliable system. Due to the widespread coverage of cellular systems, cellular based localization can provide location accuracy to within a single building in areas where GPS access is denied such as indoors or in urban canyons. Room level accuracy is not possible with this technology because GSM is a very narrowband system with limited frequency diversity. Performance may improve as wider bandwidth 3G devices become more common and these methods are applied to the newer technologies. The power consumption and cost of cellular radios is very high when compared to the inexpensive and low power radios typically used in WSN (Sahi, 2002).

ANSI 371.1 RTLS (Wherenet, Inc)

ANSI 371.1 is a standard that specifies physical layer requirements and location accuracy for a real time location system (RTLS) that is based on the system developed and marketed by WhereNet, Inc. This is a 2.4 GHz direct sequence spread spectrum (DSSS) based system that consists of time synchronized base stations and low complexity tags that can be programmed to send a signal at regular intervals. The tag location is estimated using time difference of arrival, and the base stations are mounted on either the ceilings of manufacturing facilities or on tall posts for outdoor networks. The tags are programmed to send localization signals at regular intervals, and multi-year lifetimes are achievable when location updates occur every few minutes. The 60 MHz bandwidth localization signal is transmitted in the 2.4 GHz band and contains the tag's ID as well as a small payload that can be filled by user applications to transmit fault conditions or to send other brief messages. In one example deployment over a 280,000 m^2 outdoor facility, an access point was mounted to a post approximately every 90m to ensure localization accuracy of within 3m. This network was deployed just like a traditional wireless LAN starting with a site survey followed by access point installation. A full trial of the system was completed within 75 days of the start of the site survey showing both that deployments are quick on the scale of industrial automation but are also slow and expensive compared to what is expected in the WSN space. With a reported accuracy of 3m and a bandwidth of 60 MHz, it is clear that the accuracy is not limited by noise. Multipath effects common in the industrial environment cause significant accuracy degradation, and a fairly complex system with powerful, centralized base stations is required to provide reasonable performance (Wherenet, 2008)

Ubisense UWB Localization

Ubisense developed and markets an UWB based localization system that combines TDOA and angle of arrival (AoA) measurements to estimate tag location. The tags are equipped with 802.11b transceivers for data communication and proprietary UWB transmitters for localization. The tags are capable of operating at very low duty cycles to enable multi-year battery lifetimes and typically operate by sending UWB signals at regular intervals for localization. The base stations are complicated devices consisting of an array of UWB antennas that are used to estimate the angle of arrival of the UWB signal. These antennas are attached to UWB receivers that precisely estimate the time of arrival of the incoming signal before this information is passed to a central server where the location estimation occurs. In typical deployments, location accuracy of 15cm has been reported, and the UWB signal occupies 2GHz of RF bandwidth. The location accuracy is equal to c/B suggesting that the system bins incoming signals into 1/*B* bins allowing the direct (first) path to be resolved whenever there is a direct path signal. Just as with any ranging system, this excellent accuracy is not achieved when there is no direct path. The site survey process attempts to determine ideal base station positions to reduce the number of locations in which this occurs. Only two base stations must be within range of the tag due to the combination of AoA and TDOA resulting in a lower base station density than would be required otherwise. As with all systems relying on time synchronized and wired infrastructure, the installation process is protracted and expensive (Ubisense, 2008).

Radio Interferometric Positioning System

Radio interferometric positioning system (RIPS) is an idea that uses the effect of interference between RF signals that are closely spaced in frequency to estimate position. This technique is not strictly a ranging technique as discussed in this chapter because a large number of ranges between nodes are estimated simultaneously using many measurements across the network. Four nodes are needed to perform an interferometric measurement under this scheme as shown in Figure 12, and at least 6 nodes are required to achieve network localization. To take a measurement, four loosely time synchronized devices within range of one another negotiate an operation. Two of the devices transmit unmodulated carrier signals that are separated by a very small frequency offset of about 1kHz. The signals interfere at the receivers to generate a signal that has a time varying envelope at the difference frequency. This envelope can be measured by using the RSSI on the radio, and the relative phase difference, ϕ, between the envelopes at the two receivers is recorded. This phase contains information regarding the distance between the four nodes in that $\phi = 2\pi(d_{AB} - d_{BD} + d_{BC} - d_{AC})/\lambda_{carrier}$. Once enough measurements are collected to fully define the problem, all of the ranges and locations can be calculated simultaneously. This scheme does not require precise time synchronization or significant signal processing, but it does require radios with highly precise control over the transmitted RF carrier frequency limiting its applicability to a small subset of available radios. In an open, outdoor space, RIPS has achieved accuracy of a few centimeters over ranges of many tens of meters. The primary drawback to this system is that it intrinsically relies on the carrier phase to estimate range, and the carrier phase is a strong function of multipath propagation. As a result, the system is largely unusable indoors due to poor accuracy (Maroti, 2005).

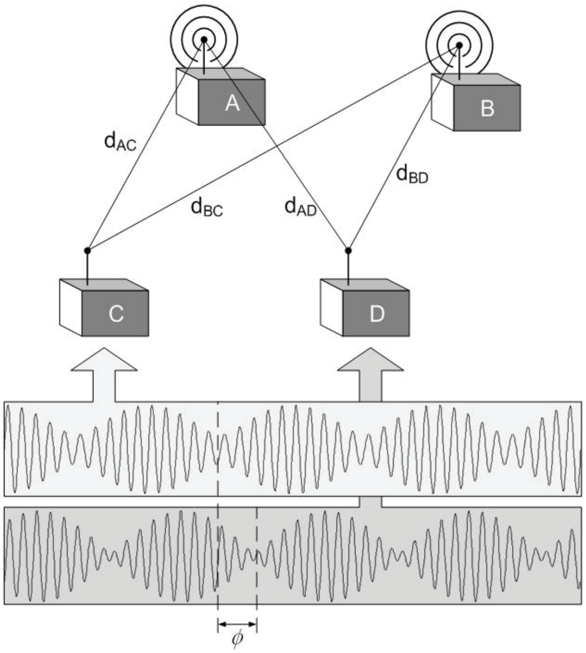
Two Way Ranging

In many WSN applications, time synchronized infrastructure is too expensive or time consuming to deploy. As discussed in the clock synchronization section, two way ranging techniques can be implemented to eliminate the effects of unknown time offset, and a few systems using these methods of two way ranging have been proposed.

Two Way Time Transfer

Two way time transfer (TWTT) was first proposed in the 1960's to provide better time synchronization between ground stations using the first communications satellite links, and the method was discussed earlier in this chapter. By performing many TWTT measurements over a period of time, time synchronization and time of flight estimates could be made accurate to within a few nanoseconds. This method has recently been proposed for use in UWB ranging systems where the E_s/N_0 values are low enough

Figure 12. The Radio Interferometric Positioning System uses the interference between two RF Signals closely spaced in frequency to generate a varying envelope that can be measured using the received signal strength indicator in a radio. The phase offset, ϕ ,between two envelopes at two nodes is a function of the unknown distances. Adapted from Maroti (2005).



such that matched filter sampling is sufficient to achieve the CRB. Some systems have been published showing accuracy of *c/B*, but no low power systems have been demonstrated. Recent work on low power UWB transceivers have reduced power consumption compared to their high data rate counterparts, but the receivers still consume a great deal of power and/or communicate over just a meter or two of range . Although the accuracy of UWB ranging systems is quite good, the power consumption and complexity of these devices are very high. Narrowband 802.15.4 radios can turn on their radio, transmit a full packet (150B total) tens of meters, receive an acknowledgement and shut down in about 5ms. The radio consumes about 20mW during this time for a total 100µJ of energy per packet. When everything is considered, it is not clear that a viable wireless sensor network UWB communication system will be developed in the near future that can compete with this low energy consumption and reasonable range.

ISO/IEC WD 24730-5 (Nanotron Technologies)

The ranging system implemented by Nanotron Technologies and standardized under ISO/IEC WD 24730-5 uses Chip Spread Spectrum (CSS) over an 80MHz bandwidth in the 2.4GHz band. CSS is a form of linear frequency modulation that can have t_sB>1, and chirp pulses have been widely used in radar because they exhibit excellent spectral occupancy and correlation properties (Richards, 2005). They also propose a two way ranging method called Symmetric Double Sided-Two Way Ranging (SDS-TWR) to combat the effects of frequency reference mismatch at the two cooperating nodes. SDS-TWR measures the round trip TOF between two nodes twice. A signal is sent from node A to B back to A, and then a signal is sent from B to A and back to B. The resulting estimates are averaged, and the effect of the reference clock frequency offset is eliminated because the two measurements have the same bias in magnitude but opposite in sign. This method is simple in implementation but it increases the required signal processing because the range must be estimated twice. In order to offset the costs associated with SDS-TWR, ranges can be taken as simple round trip measurements or even as TDOA measurements depending on how the system and infrastructure is configured. Nanotron reports location accuracy for a typical indoor office building to be 2m showing that accuracy is limited by multipath. Because roundtrip measurements are possible, fixed location nodes can be added cheaply to improve location accuracy in difficult areas. Available devices are reasonably low power while providing peer to peer, infrastructure free ranging capability with sufficient accuracy for many applications. The devices are also capable of data communication for a complete RF solution for WSNs (Nanotron, 2008).

Ranging System Comparison

The ranging systems and methods presented in this section compare to the performance and specification wish list of good accuracy, low energy consumption, low node cost, and no infrastructure requirements.

Commercially, limited options exist for RF range measurement in WSNs. Table 2 summarizes available options in both the commercial and research sphere, and the numerical information is provided to give a rough estimate of typical performance. In the evaluation of any system (ranging or localization) accuracy cannot be simply expressed as a single number because estimates are impacted by random environmental parameters resulting in estimates with random error as discussed in this chapter, yet this table provides an approximate comparison of techniques. Most available systems are extremely limited in accuracy or require significant and costly infrastructure, but it is clear that progress has been made in enabling real time localization in wireless networks. Significant research is ongoing in this area to develop adequate, low power ranging systems. This comparison should provide insight into the capabilities, limitations, and challenges of RF ranging systems and show that local area RF ranging is an open problem for research.

SUMMARY

This chapter has provided an overview of the important factors that influence RF ranging system accuracy. A number of techniques and systems designed to address these factors have been presented in order to provide an understanding of issues at hand.

Fundamental Limits

The fundamental limits to performance are the result of noise and finite bandwidth in multipath environments, but accuracy is almost always limited by multipath induced error rather than noise. Indoors, most systems are limited to resolving multiple paths that are spaced by less than $1/(2B)$ in time, but super resolution and sampling aware techniques can improve accuracy to better than $c/(2B)$ in range. Predicting error in multipath environments requires knowledge of the channel impulse response as

Table 2. List of various ranging techniques and approximate values for performance

Method	Class	Outdoor accuracy	Indoor accuracy	Noise performance	Energy Consumption	Node hardware cost	Infrastructure hardware cost
RSSI – uncalibrated	RSSI	5m	10m	Poor	Low	Minimal	None
RSSI - calibrated	RSSI	3m	3m	Poor	Moderate	Minimal	High
GPS	TOA	5m	-	Good	High	High	None
GSM TDOA	TDOA	20m	>100m	Good	High	High	High
Wherenet	TDOA	1m	3m	Good	Moderate	Low	High
RIPS	Interferometric	< 10cm	-	Good	Low	Minimal	None
UWB TWTT	TWTT	10 cm	1m	Good	High	Moderate	None
Nanotron	TWR	1m	2m	Good	Low	Low	None

well as the characteristics of the transmitter and receiver making high performance ranging systems challenging to design.

Narrowband vs. Wideband

Most research on RF ranging has relied on increasing bandwidth as much as possible to obtain reasonable ranging accuracy. Wide bandwidth is a good thing to have to improve accuracy in a variety of environments, but it is not the only solution. Multipath mitigation schemes for narrowband radios have been proposed that enable good accuracy, and it is likely that both wideband and narrowband systems will see broad application. In systems requiring the best accuracy, wider bandwidth is a good choice. For general applications, however, the additional energy costs may limit the application of UWB and other wideband systems. Narrowband radios will likely remain cheaper to design (and therefore purchase), cheaper in energy consumption per packet, and widespread in wireless sensor networks.

Conclusions

Ranging systems for low power systems have just started to be developed, and the field is open for new ideas and improvements. Significant work remains to provide the ideal ranging platform for wireless sensor networks, but some systems can provide reasonable performance for a number of applications. The future promises to provide truly location aware wireless networks, and RF ranging is critical for widespread use.

REFERENCES

Aiello, G. R., & Rogerson, G. D. (2003). Ultra-wideband wireless systems. *IEEE Microwave Magazine*, *4*(2), 36-47.

Anjum, F., Pandey, S., & Agrawal, P. (2005). Secure Localization in Sensor networks using transmission range variation. *Proceedings of the IEEE Mobile Adhoc and Sensor Systems Conference*.

Carter, M., Jin, H., Saunders, M., & Ye, Y. (2006). SpaseLoc: An adaptive subproblem algorithm for scalable wireless sensor network localization. *SIAM Journal on Optimization*. *17*(4), 1102-1128.

Cheng , Y., Chawathe, Y., LaMarca, A., & Krumm, J. (2005) Accuracy Characterization for Metropolitan-scale Wi-Fi Localization. *Proceedings of the Third International Conference on Mobile Systems, Applications, and Services*. (pp. 233-245).

Dharamdial, N., Adve, R., & Farha, R. (2003). Multipath Delay Estimations using Matrix Pencil. *Proceedings of the IEEE Wireless Communication and Networking Conference, 1*, 632-635.

Dickey, F., Romero, L., & Doerry, A. (2001). *Superresolution and Synthetic Aperture Radar. Sandia Report SAND2001-1532*. Retrieved March 14, 2008 from Department of Energy Scientific and Technical Information Bridge. Web site: http://www.osti.gov/bridge/purl.cover.jsp?purl=/782711-Y2uIQp/native/

Doherty, L., Pister, K. S. J., & El Ghaoui, L. (2001). Convex position estimation in wireless sensor networks. *Proceedings of the IEEE Conference on Computer Communications*, *3*, 1655-1663.

Hoel, P., & Stone, J. (1971). *Introduction to Probability Theory*. Boston: Houghton Mifflin.

Kaplan, E., & Hegarty, C. (2005). *Understanding GPS: Principles and Applications*. Norwood, MA: Artech House Publishers.

Kirchner, D. (1991). Two-way time transfer via communication satellites. *Proceedings of the IEEE*, *79*(7), 983-990.

Lanzisera, S., Lin, D., & Pister, K. (2006). RF Time of Flight Ranging for Wireless Sensor Network Localization. *Proceedings of the IEEE Workshop on Intelligent Solutions in Embedded Systems*.

Lanzisera, S., & Pister, K. (2008) Burst Mode Two-way Ranging with Cramér-Rao Bound Noise Performance. *Proceedings of the 2008 IEEE Global Communications Conference.*

Lorincz, K., & Welsh, M. (2006). MoteTrack: A Robust, Decentralized Approach to RF-Based Location Tracking. *Springer Personal and Ubiquitous Computing, Special Issue on Location and Context-Awareness*.

Maroti M., Kusy B., Balogh G., Volgyesi P., Molnar, Karoly, Dora S., & Ledeczi A. (2005). Radio Interferometric Positioning. *Proceedings of the ACM Conference on Embedded Networked Sensor Systems*.

Nanotron Technologies nanoLOC TRX Transceiver (NA5TR1) Datasheet Version 1.03 (2008). Retrieved March 14, 2008 from Nanotron Technologies Web site: http://www.nanotron.com/EN/docs/nanoLOC/DS_nanoLOC_TRX_NA5TR1.pdf

Nefedov, N., & Pukkila, M. (2000). Iterative channel estimation for gprs. *Proceedings of the 11th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, *2*, 999-1003.

Oppenheim, A., & Schafer, R. (1975). Digital Signal Processing. Englewood Cliffs, N.J.: Prentice-Hall.

Pahlavan, K., Xinrong L., & Makela, J. P. (2002) Indoor Geolocation Science and Technology. *IEEE Communications Magazine, 2002*(2), 112-118.

Richards, M. (2005). *Fundamentals of Radar Signal Processing*. New York: McGraw-Hill.

Sahai, P. (2002). Geolocation on Cellular Networks. In B. Sarikaya (Ed.) *Geographic location in the Internet*. (pp. 13-49). Boston: Kluwer Academic Publishers.

Shah, S., & Tewfik, A. (2005). Enhanced Position Location With UWB In Obstructed Los And NLOS Multipath Environments. *Proceedings of the XIII European Signal Processing Conference*.

SiRFStarIII Product Insert (2008). Retrieved March 14, 2008 from SiRF Technologies Web site: http://www.sirf.com/products/GSC3LPProductInsert.pdf

Song, L., Adve, R., & Hatzinakos, D. (2004). Matrix pencil positioning in wireless ad hoc sensor networks. *Proceedings of First European Workshop on Wireless Sensor Networks*, (pp. 18-27).

Spencer, Q., Jeffs, B., Jensen, M., Swindlehurst, A. (2000). Modeling the statistical time and angle of arrival characteristics of an indoor multipath channel. *IEEE Journal on Selected Areas in Communications*, *18*(3), 347-360.

Tse, D., & Viswanath, P. (2005). *Fundamentals of Wireless Communication*. Cambridge, UK: Cambridge University Press.

Ubisense Limited (2008). *Ubisense System Overview.* Retrieved July 29, 2008 from Ubisense Limited Web site: http://www.ubisense.net/media/pdf/Ubisense%20System%20Overview%20V1.1.pdf

Ulaby, F. (1999). *Fundamentals of Applied Electromagnetics*. Upper Saddle River, NJ: Prentice Hall.

Werb, J., Newman, M. Berry, V., & Lamb, S. (2005). Improved Quality of Service in IEEE 802.15.4 Mesh Networks. *Proceedings of International Workshop on Wireless and Industrial Automation*.

Wherenet, (2008). *NYK Logistics Case Study*. Retrieved July 29, 2008 from Wherenet Web site: http://www.wherenet.com/NYKLogisticsCaseStudy.shtml