Comparison of wireless clock synchronization algorithms for indoor location systems

Ciarán McElroy, Dries Neirynck, Michael McLaughlin
Decawave Limited,
Dublin, Ireland.
www.decawave.com

***Abstract*—The advent of time based location systems paves the way to the introduction of many exciting applications. However, they can only function correctly if the system has a common concept of time. Often it is not practical to synchronize all receivers using wires so alternative methods must be found. This paper outlines the use of a DW1000 ScenSor, an IEEE 802.15.4a transceiver, as a test platform. It then describes and compares several methods of wirelessly synchronizing all the sensors in the location system.**

Keywords—wireless clock synchronization; RTLS; UWB

I. INTRODUCTION

Decawave's DW1000 ScenSor is an IEEE 802.15.4a compliant transceiver aimed at wireless sensor networks (WSN) and real-time localization systems (RTLS). The chip enables customers to replace proprietary solutions based on discrete components with a standard-based integrated circuit. It builds on the advantages of ultra wideband (UWB), such as unlicensed operation, robustness in multipath environments, high precision ranging and low power transmission. It benefits from the low cost and small form factors of IC technology. The fully coherent receiver architecture ensures maximum communications range and positioning precision.

The chip implements a coherent 802.15.4a compliant UWB PHY, including all the modulation, demodulation and error correction required. Most of the MAC functionality has to be implemented by a host processor, which can communicate with the chip over a SPI connection. In order to support automatic acknowledgment of frames, address filtering and cyclic redundancy check (CRC) are implemented on the chip itself. The integrated circuit is manufactured using TSMC CMOS 90 nm technology.

The chip also contains the necessary timers and a leading edge detection algorithm to accurately timestamp transmit and receive messages. This allows it to be used in two-way ranging and RTLS applications.

In the next section, the main features of the standard on which the chip is based are described. This is followed by a description of an example application, time difference of arrival (TDOA) based RTLS. In order to be able to calculate time differences, accurate clock synchronization is necessary. Methods to implement this wirelessly are discussed and tested in the last sections.

II. THE TRANSMITTED SIGNAL

The 802.15.4a standard [1] was originally an extension of the IEEE 802.15.4 standard, which is commercialized as ZigBee [2]. It was merged into the main 802.15.4 standard as the UWB PHY [3] in 2011.

The goal of the amendment was to provide wireless sensor networks with increased range, lower power, multipath immunity, enhanced coexistence and precision ranging capability. Two additional physical layers were defined, one based on chirp spread spectrum and one based on UWB. The DW1000 uses UWB.

The standard defines the format of the waveforms to be transmitted and leaves a lot of flexibility to the implementers, especially when it comes to the receiver architecture. One of the goals of the task group was to produce a standard that could be implemented by either a coherent and non-coherent receiver architecture.

A. The frame format

A standard compliant frame is made up of three main parts: a synchronization header (SHR), followed by a PHY header (PHR) and a data field, see table I.

The synchronization header is transmitted first. Its first part, SYNC, consists of the repetition of a known preamble sequence to enable a receiver to detect the transmission and determine the channel impulse response. For each frequency band, a number of codes with minimal cross correlation were chosen. This allows for multiple networks operating at the same frequency.

One of the advantages of an impulse radio transceiver is that it is easy to implement a ternary modulation scheme. The standard defines a set of ternary preamble codes in order to support both coherent and non-coherent receivers. Ternary refers to the fact that the codes consist of positive going pulses, an absence of a pulse, and phase inverted pulses. Each preamble code element is transmitted as a single pulse, spaced a fixed distance apart. A baseband example is shown in Fig. 1.

TABLE I. IEEE 802.15.4A PHY FRAME FORMAT

SHR		PHR	Data Field
SYNC	SFD		

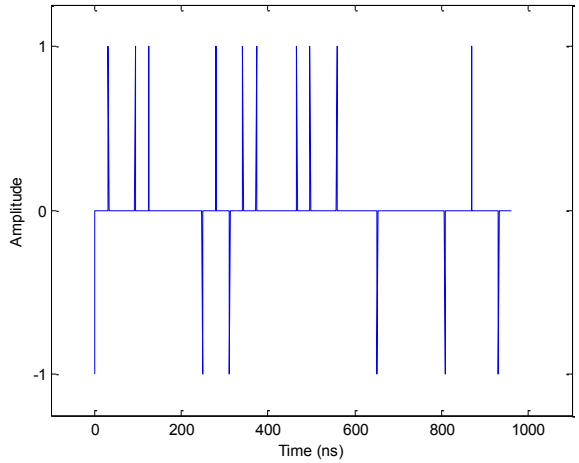

Fig. 1. Example preamble code

A non-coherent receiver needs to use the patterns of signal-silence to detect transmissions. A coherent receiver can also use the phase of the signal to obtain a further 6 dB performance gain from the correlation with the preamble sequence.

The codes are members of a set of codes known as Ipatov sequences, discovered by Valery Ipatov in 1980. Both the ternary code and its magnitude have perfect periodic autocorrelation. This means that once the presence of a transmission has been detected, the receiver can use the remainder of the preamble to reconstruct the channel impulse response for leading edge detection based ranging, see Fig. 2.

The second part of the SHR preamble is a start of frame delimiter (SFD), indicating to the receiver that the preamble is coming to an end and the transmission of the data part of the frame will follow. Like the SYNC preamble, the standard defines ternary SFD sequences. The pattern of absence and presence of the preamble symbol is intended to be used by non-coherent receivers. However, the silence periods provide no extra information to coherent receivers. Decawave have included optional nonstandard SFD sequences that always transmit a symbol and the SFD pattern is contained in a sequence of phase inversions of some of the preamble symbols. This provides a significant boost, of up to 8 dB, in the coherent receiver sensitivity, see Fig. 3. The available SFD sequences are given in table II (a 0 corresponds to a silence, a -1 corresponds to a phase inverted preamble symbol while a 1 corresponds to a normal preamble symbol).

The modulation format changes after the SHR. Pulses are no longer transmitted separately, but grouped in continuous bursts. The burst sequences and positions are determined by a pseudo-random spreading code in order to flatten the spectrum of the transmission and improve coexistence of networks. Coherent receivers can despread with the same pseudo-random code to improve their robustness to noise.

The information to be transmitted is encoded by a modulation scheme which is a combination of burst position modulation (BPM) and binary phase shift keying (BPSK). Each symbol contains two bits of information, one in the position of the burst, another in its phase. Before transmission, the PHR and data field pass through a systematic, rate ½ convolutional encoder. The systematic output is mapped to the position of the burst. Since both coherent and non-coherent receivers can detect the position, both are able to receive the packet. The parity bit is used to determine the phase of the burst. Coherent receivers achieve superior performance by exploiting this extra error correction information.

The part immediately after the SFD, the PHY header (PHR), informs the receiver about the length of the data field and the data rate used to transmit it. Since this information is crucial for successful decoding of the data, it is protected by a single error correcting, double error detecting (SECDED) Hamming code.

Finally, the data field is transmitted at the rate specified in the PHR. To help the receiver with error corrections, a systematic (63,55) Reed Solomon code over Galois field 6 is applied to the data field.

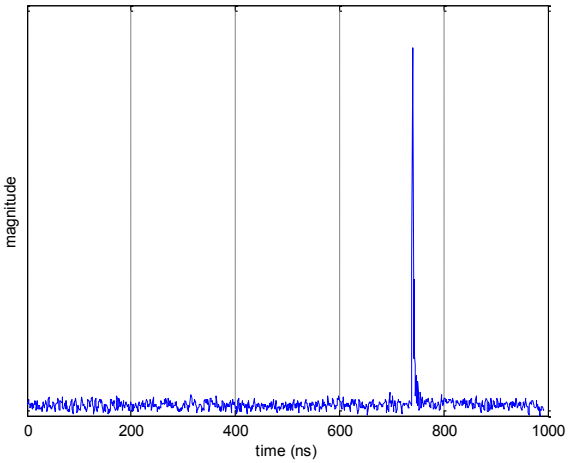

Fig. 2. Accumulating the cross-correlation for 1000 preamble symbols at an SNR of -10dB in an AWGN channel.

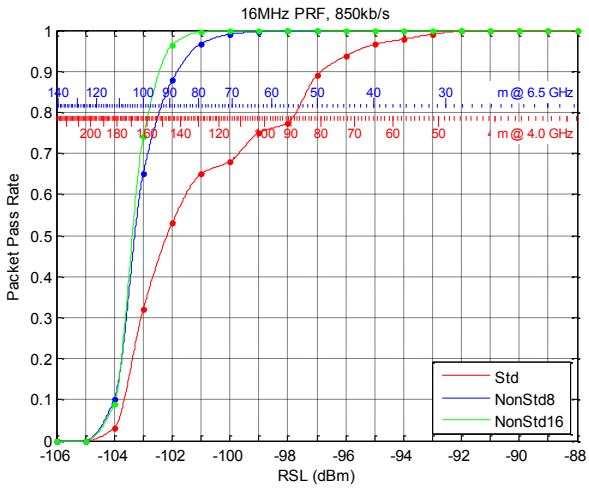

Fig. 3. Sensitivity plot comparing the standard SFD pattern with Decawave nonstndard SFD sequences for 850kb/s.

TABLE II. THE SFD SEQUENCES ON THE DW1000.

ID	Sequence
Standard Short	0 +1 0 -1 +1 0 0 -1
Standard Long	0 +1 0 -1 +1 0 0 -1 0 +1 0 -1 +1 0 0 -1 -1 0 0 +1 0 -1 0 +1 0 +1 0 0 0 -1 0 -1 0 -1 0 0 +1 0 -1 -1 0 -1 +1 0 0 0 0 +1 +1 0 0 -1 -1 -1 +1 -1 +1 +1 0 0 0 0 +1 +1
NonStd8	-1 -1 -1 -1 +1 -1 0 0
NonStd16	-1 -1 -1 -1 +1 -1 +1 -1 -1 +1 +1 -1 -1 +1 0 0
NonStd64	-1 -1 -1 -1 -1 -1 -1 +1 -1 +1 -1 -1 -1 -1 -1 -1 +1 -1 -1+1 -1+1 -1 -1 +1 -1 -1 +1 -1 -1 +1 -1 -1 -1 +1 +1 -1 -1 +1 +1 +1 -1 +1 -1 +1 -1 -1 -1 +1 -1 -1 +1 -1 -1 -1 -1 +1 +1 +1 0 0

III. LOCATION USING TIME DIFFERENCE OF ARRIVAL

Since the DW1000 allows accurate measurement of the time of arrival of the signal, it is ideally suited for determining the location of the transmitter.

The typical scenario considered for this work is the case where a transmitter (the tag) is to be located. There will be several (4 or more) receivers (the anchors) at known locations listening for the tag's transmission.

The tag needs to be as small, light and inexpensive as possible. So the battery powering it will be as small as possible and its clock will use an inexpensive crystal. The standard specifies that a crystal with a tolerance of ±20 ppm is allowed in the transmitter and receiver. If the anchors use a ±1 ppm temperature controlled crystal oscillator (TCXO) then the tolerance on the tag can be relaxed even further.

The time of arrival (TOA) at the anchors can be measured. The difference in the TOAs between any two anchors defines a hyperboloid that contains the source of the transmission. The intersection of at least 3 hyperboloids from at least 4 anchors is sufficient to locate the source in 3 dimensions. This is known as the time difference of arrival (TDOA) approach.

In reality measurement noise will mean that there is often no single point of intersection so a best fit solution should be used. There are a wide range of algorithms discussed in the literature to solve this multilateration problem [4][5]. The algorithm used in this work is the spherical intersection method outlined in [5]. It is an LMS based solution that can be used with just 4 anchors, although using more than 4 anchors does improve its performance. However it is not guaranteed to produce a valid answer. Sometimes the solution is a complex number. The results are discarded in this case.

The limitations placed on the tag, as discussed above, are not a problem for the TDOA approach. In contrast to time of flight based systems where the tag needs to negotiate separately with every anchor in range, once the anchors themselves can report the time at which the signal arrived at their location in a common time domain then there is sufficient information to solve the location problem in a single packet transmitted by the tag. This results in a considerable tag power saving.

The main sources of error for TDOA based location include:

- Channel impairments.

  The location system assumes that there is a line of sight between the tag and the anchor. If there is no line of sight then the signal that is received will have either gone through the obstacle or the reception is based on a reflection. This will result in a longer time of flight than for the line of sight path.

- Anchor location errors.

  The location that is provided is a location relative to the anchors. If they are not where they are reported to be then it will be wrong. Furthermore if the error in the anchor location is large enough to be inconsistent with the measured TDOA is may not be possible to solve the multilateration problem.

- Clock synchronization errors.

  Each anchor reports the TOA of the received packet. This information is only meaningful if all the TOAs can be converted to a common time domain.

The clock synchronization problem is the one that will be considered here.

IV. CLOCK SYNCHRONIZATION

Clock synchronization in WSN is a widely studied problem [6]. This paper focuses on the effects of clock synchronization on location performance.

Each anchor has its own independent clock source that is used to measure the TOA of a packet from the tag. Each of these sources will have frequency errors relative to each other, due to different tolerances on the clock oscillators. Furthermore the frequency errors are not constant [7]. They will vary with the local temperature variations at each anchor and also with random errors such as Allan variance [8]. Different startup times mean that time zero will be different for every anchor. This means that no two anchors will report the time of arrival relative to a common clock.

Fig. 4 demonstrates the problem. In this case one anchor is transmitting a packet every 150ms as measured by its clock. The plot is showing the variation in the TOA at another anchor. Both anchors are stationary. The anchors are using TCXOs with a ±1 ppm tolerance. In fact the measured frequency error is ~-0.2 ppm (a change of about -31ns over a 150ms period). Given that light moves at approximately 30cm per ns it is clear that even with such tight tolerances a clock synchronization algorithm is necessary to adjust the reported time of arrivals so that they are all reported relative to a common clock before the multilateration is performed.

It is not just the dominant frequency error of -0.2 ppm that needs to be compensated for. There are smaller variations happening continually. The relatively large transient at 350s can be caused by a small draft (e.g. opening a door). The synchronization algorithm must have a low enough latency to react to these rapid changes while simultaneously ignoring the measurement noise that is also apparent in the figure. The measurement noise from the DW1000 is Gaussian with a standard deviation in the 100 to 150ps range.

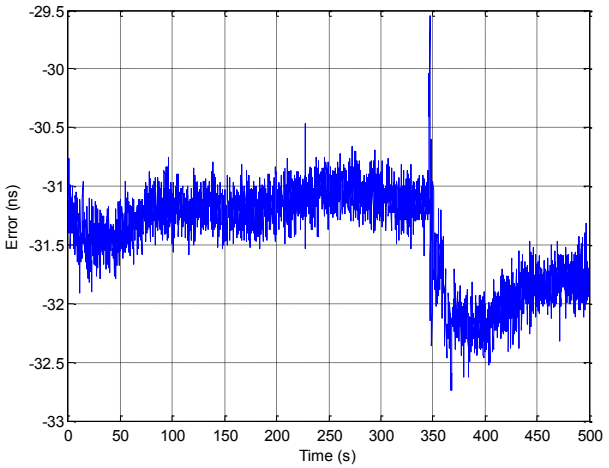

Fig. 4. Variation in the time of arrival of clock sync packets with a 150ms period.

A. Synchronization Algorithms

All the algorithms considered here use a one-way message dissemination approach of [6] where a single master periodically broadcasts a packet indicating its local time. This packet is received by each of the slave anchors and the time of arrival in each of the slave clock domains is recorded.

The period of the master broadcasts involves a trade-off between the clock synchronization performance and the amount of air time that is used by the broadcast packets. These packets are sharing the channel with the tags that are to be located. The shorter the master broadcast period is the more information about the anchor clocks but at the cost of a reduced tag capacity. This is especially significant for the lower data rate modes where a single packet could be several ms long.

Clock synchronization can be performed on the anchor itself or all the data can be returned to a central server and the calculations performed there.

1) Linear Interpolation

The linear interpolation simply buffers the TOA of all received packets. A tag blink is adjusted by linearly interpolating between the TOA of two successive clock synchronization packets. This leads to a delay since the location of the tag is only calculated after the next clock sync has been processed.

This algorithm assumes that all clock synchronization TOAs are correct so it makes no attempt to reject outliers or even to ignore the measurement noise. For this reason the algorithm can suffer from some very large errors and a method of detecting these must be used. So when the multilateration is performed the exact TDOA for this estimated location is computed. If this TDOA is very different from the measured TDOA then the estimated is rejected and the multilateration is considered to have failed.

2) PI Control

This algorithm uses the classical proportional-integral (PI) control loop. The error signal that drives the loop is the difference between the actual time of arrival and the expected time of arrival. The output of the loop filter is the amount the next increment of the expected time should differ from the nominal interval. Fig. 5(a) shows the complete loop. Only the k_p and k_i coefficients are used for the PI loop. Fig. 5(b) shows the algorithm to convert the tag blink's TOA from the slave anchor's clock domain into that of the master. The frequency offset is the integral value from the control loop.

The proportional and integral coefficients (k_p and k_i) are determined by an exhaustive search with recorded data. The coefficients that minimize the mean squared error between the expected TOA and the actual TOA are selected, see table III.

3) PID Control

The PI control algorithm needs to have a very wide bandwidth to meet the low latency requirement. This significantly reduces its noise immunity. A differentiator was added the loop filter to allow a reduction in the bandwidth while preserving the low latency. The rest of the control loop is unchanged (see Fig. 5(a)).

As before the coefficients are tuned with recorded data and they are shown in table IV.

4) PII Control

Another approach to reducing the latency of the algorithm is to increase the order of the loop filter. Introducing a double integral does this, see Fig. 5(a). Higher order loops are usually avoided because it can be difficult to ensure that they remain stable. However, if the coefficients are tuned using sufficient recorded data the risk is reduced. Table V shows the tuned coefficients.

This time the frequency offset used in the TOA adjustment algorithm is the sum of the two integral arms of the control loop.

TABLE III. COEFFICIENTS USED FOR THE PI CONTROL LOOP.

Sync. Period (ms)	k_p	k_i
150	6.667	0.667
300	3.333	0.333
450	2.222	0.222
600	1.667	0.167
750	1.333	0.133
900	1.111	0.111

TABLE IV. COEFFICIENTS USED FOR THE PID CONTROL LOOP.

Sync. Period (ms)	k_p	k_i	k_d
150	4.647	3.340	0.601
300	2.323	1.670	0.301
450	1.549	1.113	0.200
600	1.162	0.835	0.150
750	0.929	0.668	0.120
900	0.774	0.557	0.100

TABLE V. COEFFICIENTS USED FOR THE PII CONTROL LOOP.

Sync. Period (ms)	k_p	k_i	k_{ii}
150	3.5	0.9	0.2
300	2.5	0.7	0.2
450	1.7	0.7	0.2
600	1.5	0.6	0.2
750	1.3	0.5	0.2
900	1.1	0.5	0.2

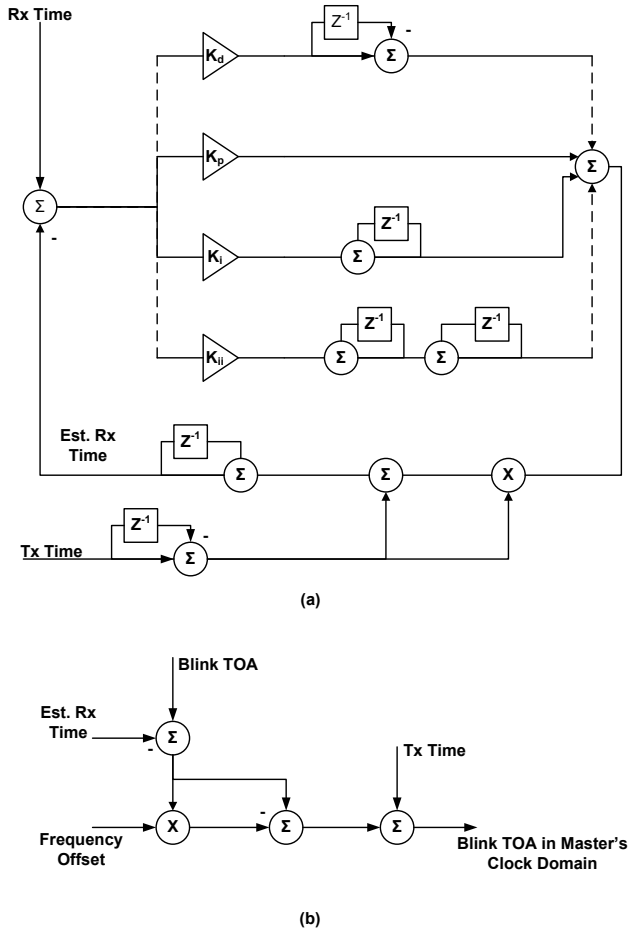

Fig. 5. Outline of the PI, PID and PII (a) control loops and (b) tag blink time adjustment.

5) Kalman Filter

A good introduction to Kalman filters is given in [9]. The filter is a recursive system that attempts to estimate the state $x \in \mathbb{R}^n$ of a system governed by

$$x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1} \quad (1)$$

With a measurement $z \in \mathbb{R}^m$ described by

$$z_k = Hx_k + v_k \quad (2)$$

Where w_k and v_k are Gaussian random variables that represent the process noise and the measurement noise respectively. The vector $\underline{u_k}$ is the control input.

This work uses a 2 dimension state vector (the time of arrival and the clock skew) and the state transition matrix, A is given by

$$A = \begin{bmatrix} 1 & dt \\ 0 & 1 \end{bmatrix} \quad (3)$$

There is no control input so the B term disappears.

Only two tuning parameters are required, namely the variance of the measurement noise (σ_v^2) and the process variance (σ_w^2). The DW1000 typically has a TOA variance of ~1.5×10^{-20}s^2. This corresponds to a standard deviation of ~3.5cm in a range estimate from a single packet. Allowing for clock jitter, temperature variations and device variations we allowed a measurement noise variance of 3×10^{-20}s^2. The process noise variance was tuned using the same recorded data as before and 5×10^{-20} seems to work well.

V. EXPERIMENTAL SETUP

Performance experiments were conducted in a room 6.5m X 6.5m X 2.7m, see Fig. 6. This is a relatively small area but the standard deviation of the DW1000's estimate of the time of arrival is unaffected by the range. There were no obstacles in the room so all the communications channels are line of sight.

The tag is placed at known locations inside the region bounded by the anchors. Location performance degrades outside this region but this degradation is not due to clock synchronization issues.

During the experiment the master is broadcasting the clock synchronization packets with a period of 150ms. All the information seen by each of the anchors is recorded in a log file. This file is then post processed using a Matlab program. Slower synchronization periods are obtained by simply decimating the synchronization information. This ensures that all the algorithms are working with exactly the same data.

The performance metric that is used to compare the algorithms is the R95xy and the R95 of the location estimate. That is the radius of the circle that contains 95% of all the estimates on the xy plane (i.e. the floor) and in all 3 dimensions respectively. This ignores any fixed bias in the location estimates. This is reasonable since the primary sources of any bias errors are issues such as errors in the anchor locations. Errors in the height of the tag are expected to be worse than in either of the other two dimensions since there is less vertical separation in the anchors.

Most of the data was gathered while the anchors were using TCXOs with a tolerance of ±1ppm as the clock source. A small amount of data using crystals with a ±20ppm tolerance was also collected for comparison purposes.

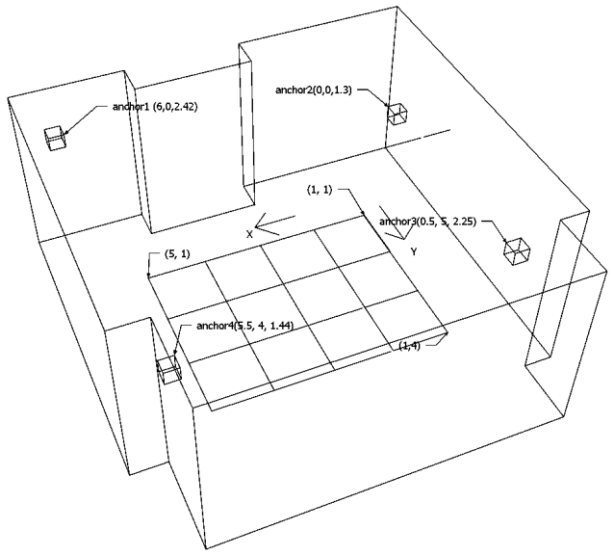

Fig. 6. Layout of the test area.

A. Results

TABLE VI. R95XY (CM)

Sync. Period (ms)	Linear Interpolation	PI	PID	PII	Kalman Filter
150	14.6	14.1	16.9	16.1	11.3
300	15.1	18.9	21.3	22.5	13.3
450	16.1	24.7	25.7	28.9	15.4
600	17.1	34.0	31.4	34.0	17.7
750	18.9	43.9	34.7	40.0	19.8
900	20.2	53.7	39.6	45.6	20.7

TABLE VII. R95 (CM)

Sync. Period (ms)	Linear Interpolation	PI	PID	PII	Kalman Filter
150	50.1	48.1	58.4	54.4	36.7
300	51.8	63.3	71.1	73.4	43.6
450	55.0	80.2	84.6	91.1	50.5
600	57.3	107.4	100.9	108.6	57.4
750	63.8	134.0	112.3	127.0	64.3
900	67.3	159.4	121.9	142.9	68.9

TABLE VIII. MULTILATERATION PASS RATE.

Sync. Period (ms)	Linear Interpolation	PI	PID	PII	Kalman Filter
150	97.8%	99.7%	99.7%	99.7%	99.7%
300	94.4%	99.5%	99.6%	99.5%	99.5%
450	91.0%	99.0%	99.3%	99.2%	99.4%
600	87.1%	98.4%	99.1%	98.9%	99.2%
750	83.3%	97.5%	98.8%	98.8%	99.2%
900	79.9%	96.3%	98.4%	98.4%	98.9%

VI. CONCLUSIONS

In terms of location accuracy both the linear interpolation and the Kalman filter are the clear winners. The performance degrades gracefully as the clock synchronization period is increased.

The linear interpolation method does suffer from the significant limitation that the multilateration pass rate drops significantly as the synchronization period increases. It also introduces a significant latency into the location estimates that may be unacceptable for many applications. The Kalman filter does not suffer from this limitation.

The DW1000 device is an implementation of the IEEE 802.15.4a UWB standard. It contains the hardware to accurately determine the time of arrival of an incoming signal. While proprietary SFD sequences increase the sensitivity of the receiver allowing anchors to be placed further apart. This makes it ideally suited for use in wirelessly synchronized indoor real-time location systems.

REFERENCES

[1] Standard IEEE 802.15.4a-2007, “Part 15.4: Wireless medium access control (MAC) and physical layer (PHY) specifications for low-rate wireless personal area networks (WPANs): Amendment to add alternate PHY, ”, March 2007

[2] ZigBee Alliance, “ZigBee specification,” ZigBee document 053474r06, version 1, 2006

[3] Standard IEEE 802.15.4-2011, “Part 15.4: Low-rate wireless personal area networks (LR-WPANs)”, September 2011.

[4] Petre Stoica and Jian Li, “Source Localization from Range-Difference Measurements”, IEEE Signal Processing Magazine, November 2006.

[5] Julius O. Smith and Jonathan S. Abel, "Closed-Form Least-Squares Source Location Estimation from Range-Difference Measuerments", IEEE Trans. Acoustics, Speech and Signal Processing, Vol. ASSP-35, pp. 1661-1669, Dec. 1987.

[6] Y.-C. Wu, Q. Chaudhari and E. Serpedin, “Clock synchronization of wireless sensor networks”, IEEE Signal Processing Magazine, vol. 28, pp 124-138, Jan. 2011.

[7] D. Dardari , A. Conti , U. J. Ferner , A. Giorgetti and M. Z. Win "Ranging with ultrawide bandwidth signals in multipath environments", Proc. IEEE, vol. 97, no. 2, pp.404 -426 2009.

[8] D. W. Allan "Time and frequency (time-domain) characterization, estimation and prediction of precision clocks and oscillators", IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. UFFC-34, no. 6, pp.647-654, Nov. 1987.

[9] G. Welch and G. Bishop, “An introduction to the Kalman filter”, 2000.