Design of the SX1280 Ranging Protocol and Result Processing

AN1200.50

Table of Contents

List of Figures

Part I Background:
Error considerations in RTToF Ranging

1. Introduction

All practical radio communication employs a protocol, this is the set of rules determining when a radio has to be ready to receive messages destined for it or transmit its own messages to another radio. Ranging operations are similar because the ranging transactions need to be coordinated, but when and how these operations are performed also affects the resulting ranging accuracy. In addition to this, the collected results need conversion into a physical range. Depending upon how we have performed the ranging, we may have additional metadata such as signal strength or frequency error that we can also use to augment the precision of the ranging measurement.

This document describes both the communications protocol and post processing performed by the SX1280 evaluation kit (EVK) when in 'Outdoor Ranging Mode'. (We do not cover user interface or operation of the kit, for this information please consult [1]). In addition to a brief overview of the basic principles of ranging, we walk you through both the protocol and the post processing of the ranging results, also explaining the rationale behind each step and alternative approaches.

2. Ranging Operation

The ranging distance measurement is based upon the round trip time of flight (RTToF) of a signal between a pair of radios. The basic principle is illustrated below. Here, one radio assumes the role of ranging Requester and another the role of ranging Responder.

The ranging Requester sends a ranging request to the Responder, which sends a synchronised response back to the Requester. The Requester measures and interpolates the time elapsed between the ranging request, $T_{Request}$, and response $T_{Response}$.

This measured time reported by the Requester is hence the round-trip time between the Requester and Responder. With all propagation occurring at the speed of light, the measured time is equivalent to the measured round-trip distance - with some additional timing errors (ε).

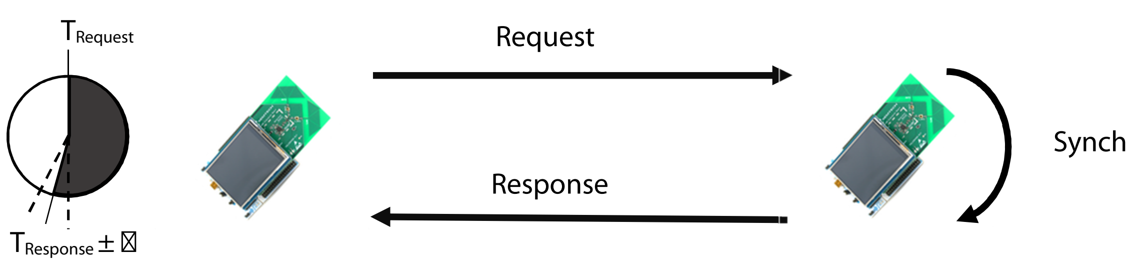


Figure 1. Conventional Requester-Responder Round Trip Time of Flight Measurement. This generates a single result, but also with an unknown error, ε.

When considering the post processing of the RTToF distance measurement, we have to consider these sources of error and, where possible, correct them. This in turn has an effect on the way we collect the distance measurement results themselves. We therefore begin by examining these sources of error and how to mitigate them.

3. Radio Ranging Errors

Before we begin, we need to draw an important delineation between static sources of error and those that can change during the operation of the ranging measurement system. It is relatively simple to correct static bias, here we want to consider the tougher case of dynamic, time varying, errors. So, to simplify our discussion, we assume that the static sources of measurement error been corrected by calibration. For a detailed treatment of this topic please consult [2].

The remaining dynamic sources of error can be further divided into those internal to the radio and those due to the channel in which the ranging signals propagate.

Assuming a perfect communication channel, we see only errors due to circuit level phenomenon, namley:

- Reference oscillator drift and
- Analog group delay.

Before we look at the influence of the communication channel and propagation on the ranging accuracy, we examine these two sources of error and their correction.

3.1 Reference Oscillator Error

The Requester's timing measurement and the Responder's synchronisation are performed using the local crystal reference oscillator of the SX1280. Thus, any offset in timing between the Requester and Responder crystal oscillators will result in a distance measurement error.

3.2 Correcting Reference Oscillator Error

Because we use the same reference oscillator to derive both the timer for ranging operations and the RF carrier frequency for the 2.4 GHz transmission, we can use the measure of frequency error between transmitter and receiver to indicate reliably and accurately the timing (so distance) offset between Requester and Responder.

The figure below shows the distance measurement correction that should be applied in response to the given frequency (so timing) error.

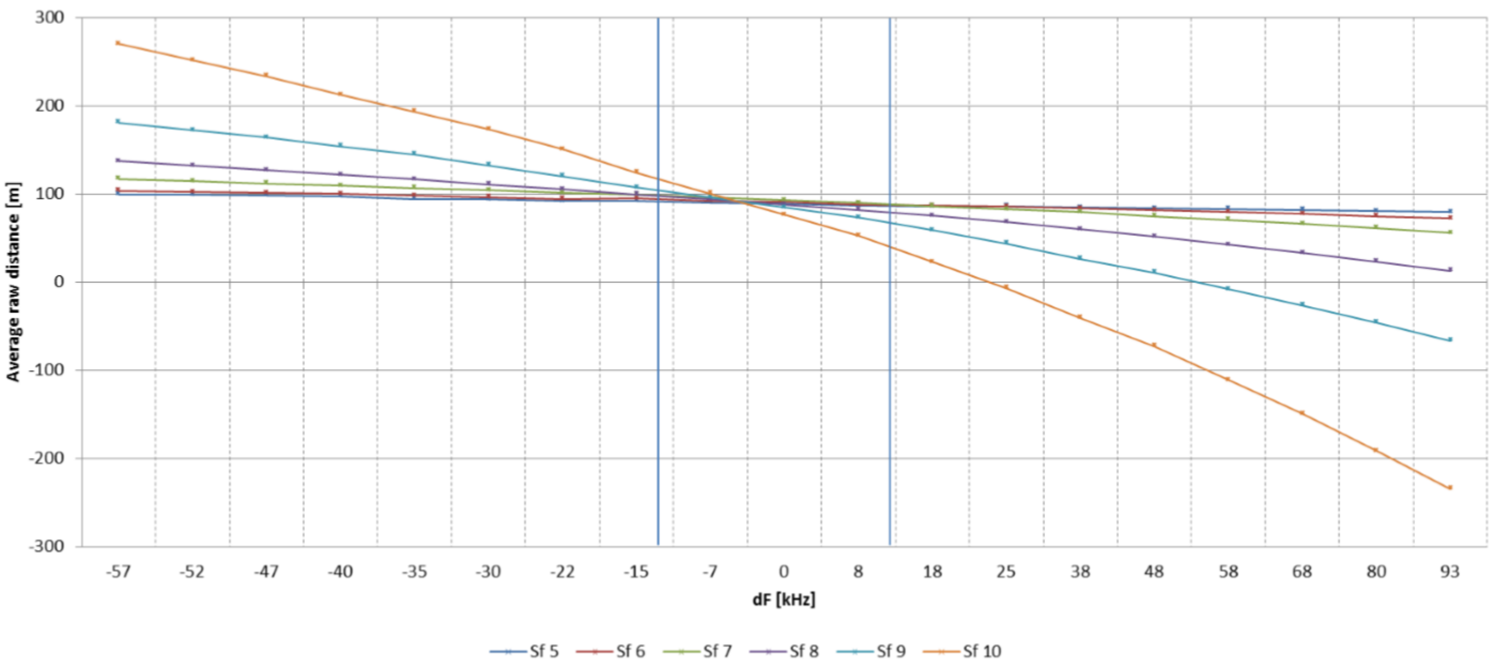
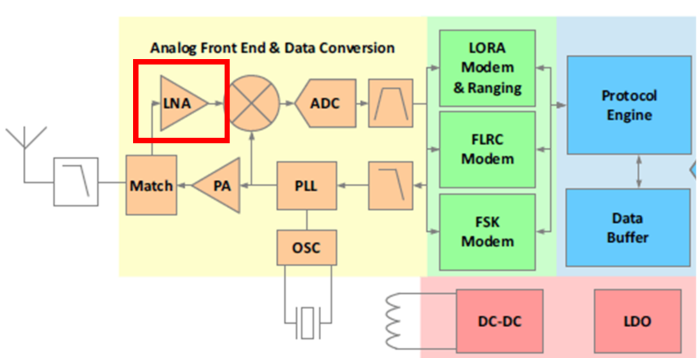


Figure 2. Distance Error versus Frequency Offset Between Requester and Responder

3.3 Analogue Group Delay Error

The SX1280 employs an Automatic Gain Control (AGC) to adapt the LNA gain of the receiver to the received signal strength. This process is standard to almost all radio receivers to allow the reception of both low power signals at the limits of sensitivity and high power signals at short range.



Setting	Gain [dB]
13	Max
12	Max -2
11	Max -4
10	Max -6
9	Max -8
8	Max -12
7	Max -18
6	Max -24
5	Max -30
4	Max -36
3	Max -42
2	Max -48
1	Max -54

Figure 3. The SX1280 LNA Gain (Outlined) and Range of Gain Settings

The impact of this on the operation on RTToF measurement is that the delay through the LNA changes as a function of the amplifier gain.

We have measured this in the figure below, with the measurement setup also shown in the inset image. Each point corresponds to a raw ranging measurement between the Requester and Responder over a 125 m (electrical length, 100 m physical length) cable. Here we can see that as we vary the attenuation between Requester and Responder, at a fixed distance, from high input power (low attenuation) on the left to low input power (high attenuation) on the right - we experience over 8 m of measurement error.

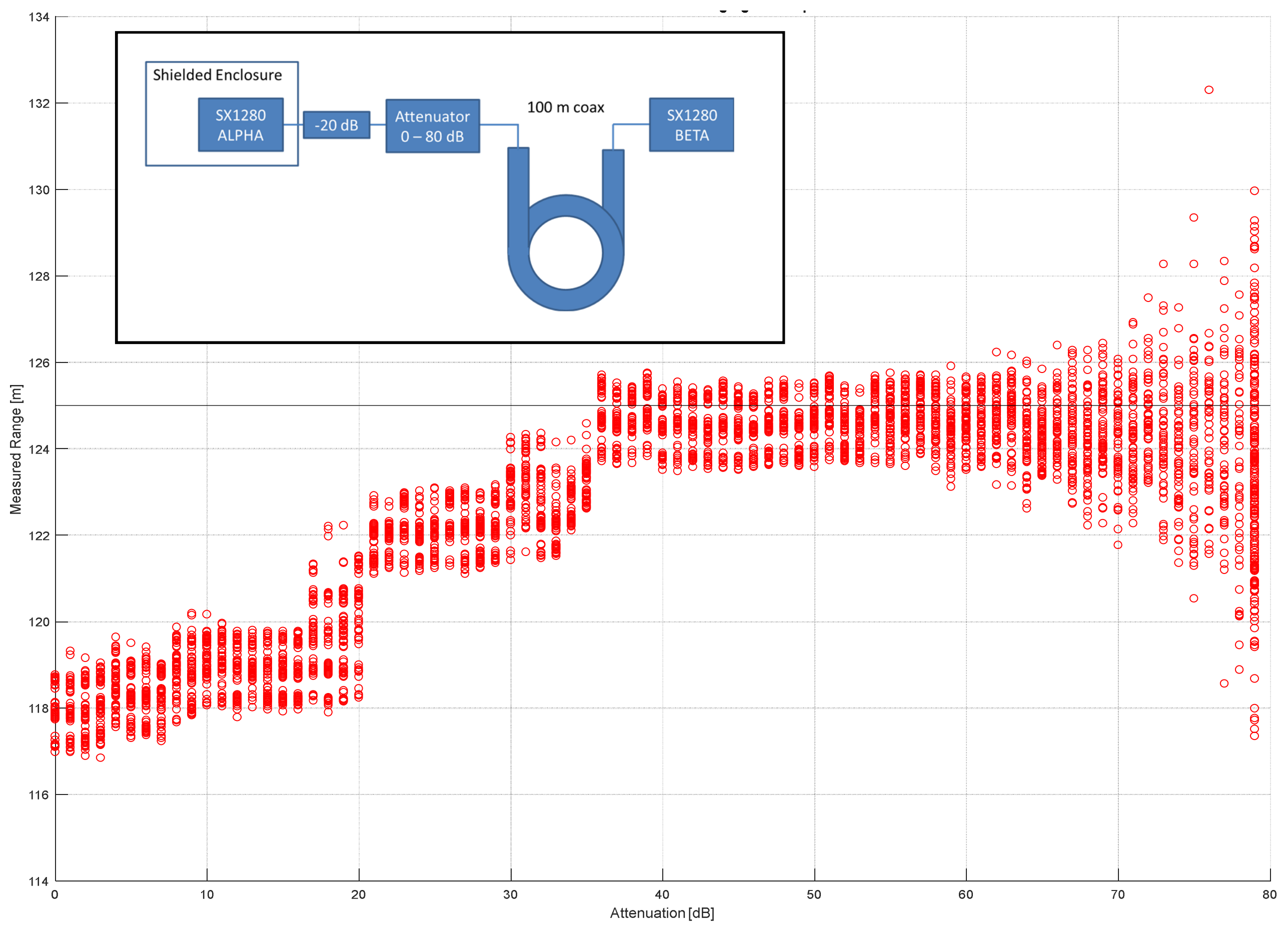


Figure 4. Measured Requester-Responder Distance, at a Fixed Distance, as a Function of Attenuation.

3.4 Correcting Analogue Group Delay Error

Correction of the variable delay due to the LNA gain setting would be possible by sharing information between the ranging Requester and Responder about which gain level was used during each ranging exchange. This approach would impose the following problems:

1) A real time processing operation to monitor gain levels of both the Requester and Responder during each exchange. (i.e. Polling the internal radio registers to measure the gain used for each exchange).

2) The requirement for additional communication overhead to send the LNA gain information from Responder to Requester. (i.e. Send the gain used by the Responder back to the Requester).

In order to avoid this we make use of the RSSI of the ranging exchange measured by the ranging Requester. In doing so we are assuming that the channel remains static for the duration of a single ranging exchange and, secondly, that the signal power seen by the Requester gives an indication of the receiver gain used by both Requester and Responder.

The Ranging RSSI differs slightly from the RSSI measured conventionally, instead of indicating the power in absolute power (dBm) the ranging RSSI indicates the received signal power relative to a signal power threshold in dB.

Correction based upon the ranging RSSI is straightforward. Based upon measurements made in a coaxial cable and variable attenuation, we construct a look-up-table of ranging RSSI versus distance measurement correction. The LUT contains 160 entries per SF BW combination so, for clarity, the LUT data for SF9 1600 kHz is plotted (rather than tabulated) below.

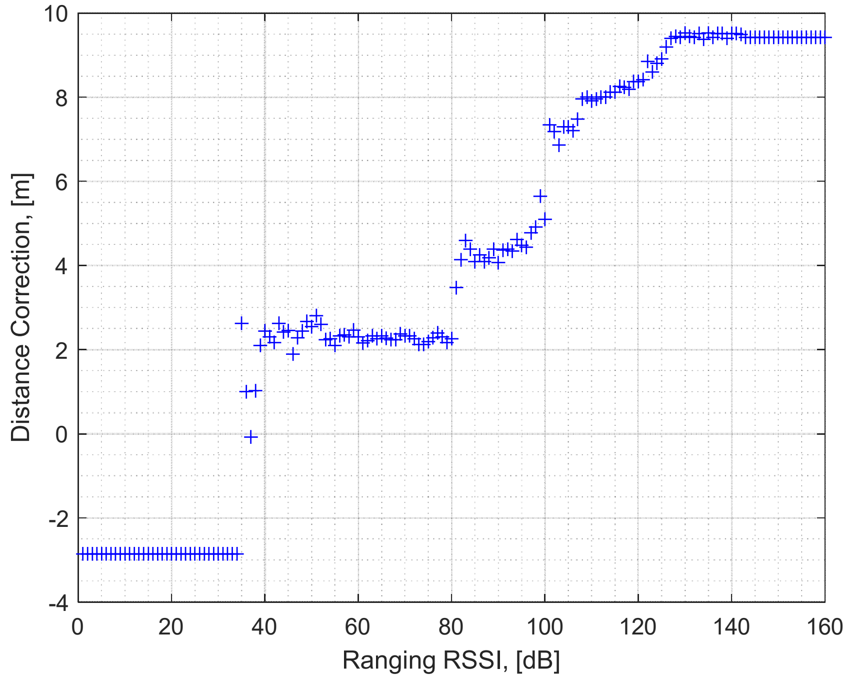


Figure 5. Applied distance correction as a function of Ranging RSSI for SF9 BW = 1600 kHz which is stored as a LUT in firmware.

3.5 Not an RSSI Based Ranging Measurement

Note that it is important to differentiate between this technique and others that employ RSSI as part of the ranging estimation process. Here we are using the RSSI information for a single ranging exchange as a proxy for knowledge of the receiver LNA gains. We are not using the signal strength to determine the range itself – only to compensate the delay of the receiver chains to make our RTToF estimate more accurate.

4. Channel Effects

With the ranging errors inherent to the radio system addressed, we must also consider the influence of the communication channel itself on the ranging result.

4.1 Channel Selectivity & Multipath

At 80 MHz wide, the 2.4 GHz ISM band is much wider than the highest SX1280 LoRa bandwidth of 1.6 MHz. Consequently, we can see significant variation in the ranging performance across the 2.4 GHz ISM band. Below is an extreme case of the point in question. In this example Requester and Responder are 150 m apart with obstructions between them (non-Line-of-Sight, nLoS, channel conditions).

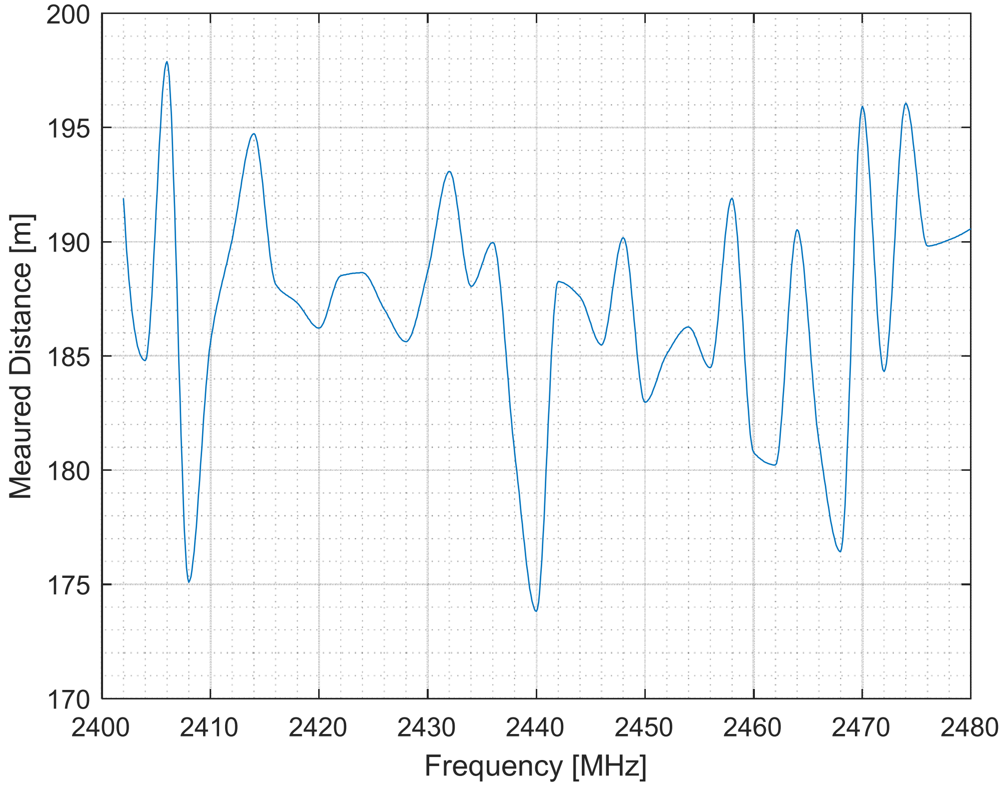


Figure 6. nLoS Ranging Measurement Results at 150 m

This phenomenon arises because the wave propagating between the Requester and Responder happens on different paths, each path being subject to different delays. In the frequency domain, in the case shown above, this delay dispersion translates into frequency selectivity on the scale of the 2.4 GHz ISM band.

In addition to frequency selectivity, the other influence of nLoS propagation between Requester and Responder is the additional bias (extra distance) that the measurement result will exhibit. In the measurement above, we see an additional 30 to 40 m compared with the ground truth distance.

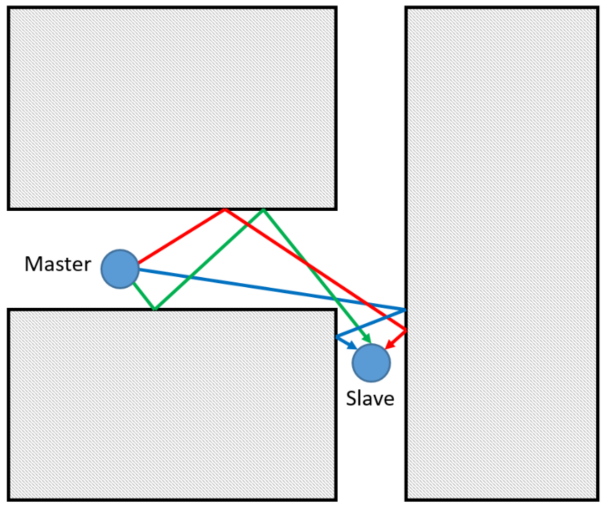


Figure 7. Illustration of Multipath Propagation between a Pair of non-Line of Sight (nLoS) Ranging Radios

The key to understand the origin of both effects is shown in the plan view (above) of the Requester transmitting the ranging request to the Responder. (Noting that the same will apply for the return, ranging response, propagation).

In this scenario, we have nLoS propagation reflected by some surrounding scattering buildings. The red green and blue lines denote the reflected paths. In this nLoS layout we see that there are multiple reflected paths from Requester to Responder, each with differing path lengths. We also see that the reflected nLoS paths are all longer than the physical distance from the Requester to Responder.

This is what causes:

1) The over estimation of the measure distance between radios, as the overall reflected distance is higher.
2) The frequency selectivity of the response of the band due to the delay dispersion of the multiple reflected paths.

4.2 Compensating Frequency Selectivity

Whilst there is no easy way to mitigate the range over estimation effect of multipath, the delay spread, so selectivity effects caused by multipath propagation can be lessened by the use of frequency and antenna diversity. Of the two diversity schemes, frequency diversity is the most important.

4.2.1 Frequency Diversity

As can be seen from the measured 2.4 GHz ISM band response, some frequencies will give less accurate range results than others. The solution to this is frequency diversity, i.e. performing many ranging exchanges across the band to help reduce the influence of measurement on a less accurate channel.

4.2.2 Antenna Diversity

An alternative to changing the frequency at which we perform the measurement, to change the delay spread seen by the radio, is to change antenna. The three types of antenna diversity are spatial, polarisation and pattern diversity. However, what is important to note here is that by changing antenna we aim to change the multipath environment between Requester and Responder so further diminishing the influence of inaccurate channels.

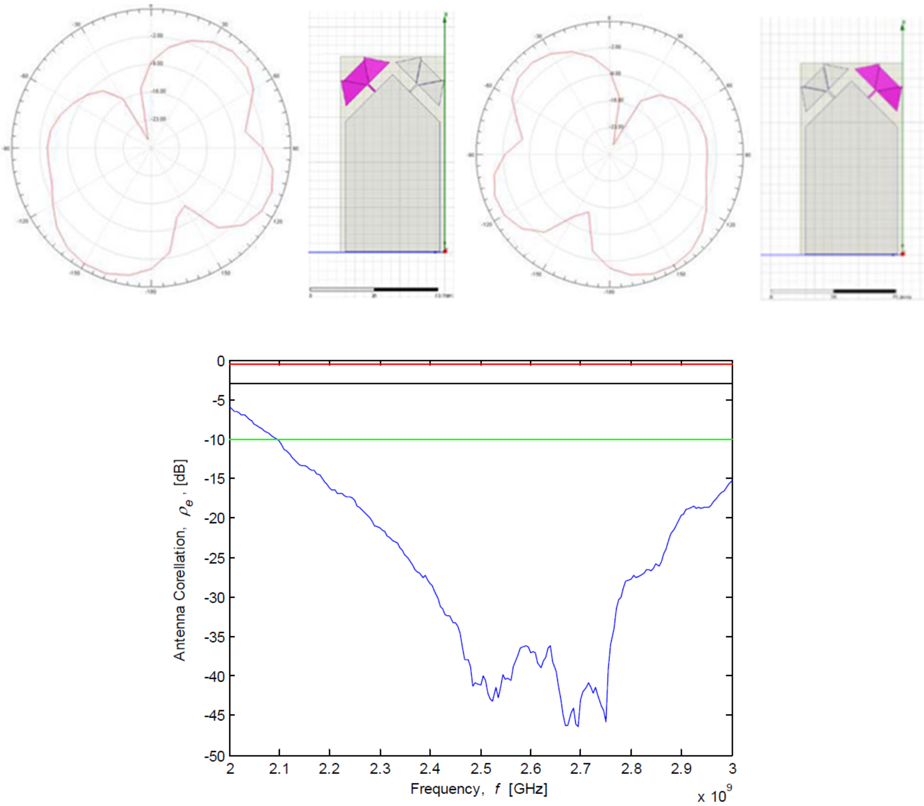


Figure 8. The Antenna Diversity of the SX1280 Evaluation Kit. (Top) the radiation patterns of each antenna and (Bottom) the Measured Antenna Envelope Correlation Coefficient.

Antenna diversity is sometimes beyond the space or cost constraints of an application. Nonetheless, the example of the SX1280 development kit is shown above. Here the top image shows the pattern diversity of the two antennas wherein the antenna radiation patterns are designed to differ from each other.

The bottom figure shows the measurement of the SX1280 development kit antenna envelope correlation factor versus frequency according to the method outlined in [3]. Correlation values below the green line indicate good antenna diversity. In our case this means that switching the antenna should provide a useful source of additional spatially-diverse channels when we are in a nLoS scenario.

4.3 Correcting Measurement Bias by Fingerprinting

Fingerprinting is the technical term used to describe the process of fitting measured ranging or localisation estimates, to a specific environment. In our case, it is used to remove the bias introduced by multipath. Whilst a detailed discussion of this subject is beyond the scope of this Note, we examine the line-of-sight case, noting that these techniques could be modified for ranging measurements in other environments.

Why do we need to correct for the line of sight case? Because even the line-of-sight channel is subject to multipath. The results from a line of sight measurement between a pair of evaluation kits is shown below. Here we see the measured result versus Ground Truth (GT). Even in line of sight conditions we see some range underestimation at short range, some overestimation between 10 to 80 m and then underestimation beyond this. Thus, even in the line of sight case, we perform a polynomial curve fit (red line) with a view to using the polynomial as a correction.

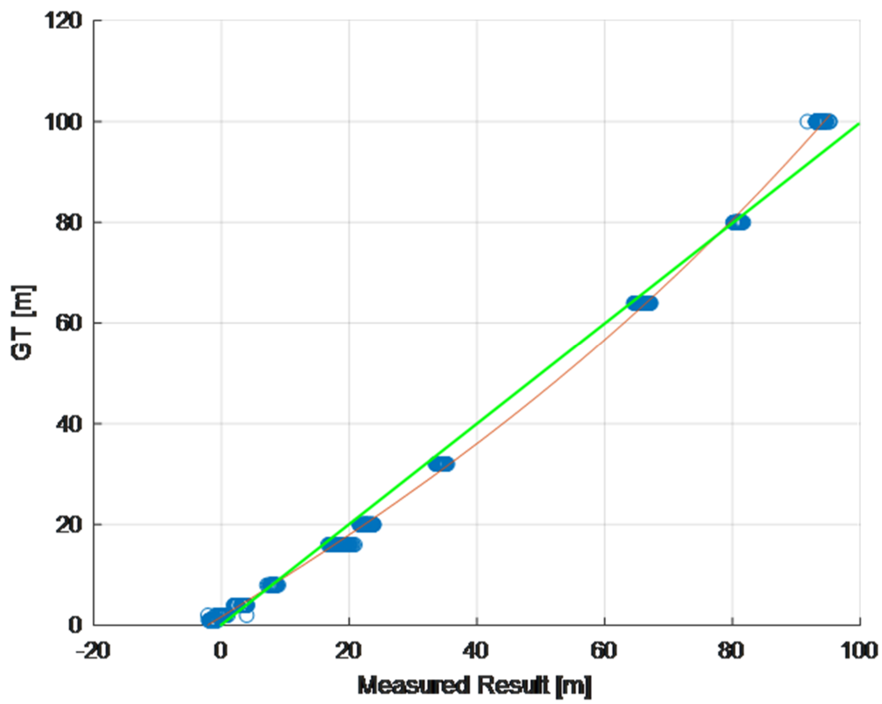


Figure 9. Curve Correction: Measurement Results versus Ground Truth (blue) and Polynomial Fit (red)

Part II Implementation:
A Practical RTToF Measurement System

5. Introduction

In this part, we take the abstract approaches outlined in the previous Section and apply them in practice. This application takes two forms:

i) The ranging protocol used to recover the ranging distance measurements and
ii) The post processing of the set of results recovered and its conversion into a measured distance.

6. The Ranging Protocol

The role of the ranging protocol is twofold: to ensure that the radios are in the right configuration at the right time and to gather adequate ranging results. In addition to this we recover the associated metadata, based upon our understanding from the previous Part, to further improve our measurement accuracy.

The basic protocol we use between the Requester and Responder during a ranging measurement is illustrated below. The ranging exchange begins with a communication phase. In a real application this would allow the Responder radio to remain in a duty cycled sleep mode to reduce power consumption. In our demonstration ranging application the Responder remains in receive mode until it received the first LoRa communication. Once we have established connectivity between the Requester and Responder we then start a series of frequency hopped ranging exchanges, each on a mutually defined channel.

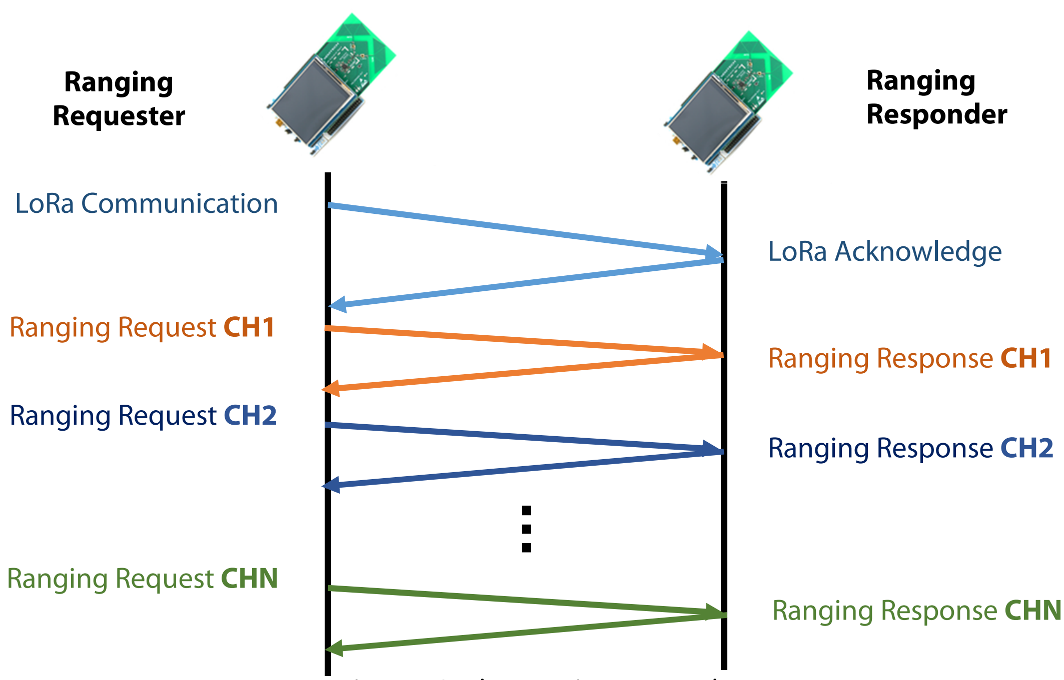


Figure 10. The Ranging Protocol

7. The Channel Plan

The frequency hopped exchanges use the BLE channel plan. In our example, we perform a single exchange on a single frequency and then hop to the next frequency for the following exchange. Due to the frequency selectivity of the ISM band hopping across multiple channels allows us the best chance of avoiding inaccurate channels.

BLE Channel Number	Hopping Sequence	Frequency [MHz]
22	CH1	2450
37	CH2	2402
35	CH3	2476
15	CH4	2436
12	CH5	2430
31	CH6	2468
26	CH7	2458
6	CH8	2416
10	CH9	2424
36	CH10	2478
25	CH11	2456
21	CH12	2448
28	CH13	2462
33	CH14	2472
13	CH15	2432
20	CH16	2446
9	CH17	2422
18	CH18	2442
27	CH19	2460
34	CH20	2474
5	CH21	2414
29	CH22	2464
24	CH23	2454
19	CH24	2444
0	CH25	2404
14	CH26	2434
3	CH27	2410
2	CH28	2408
17	CH29	2440
23	CH30	2452
39	CH31	2480
38	CH32	2426
11	CH33	2428
30	CH34	2466
7	CH35	2418
4	CH36	2412
1	CH37	2406
32	CH38	2470
17	CH39	2438
8	CH40	2420

In the implementation on our EVK the number of hopping channels is variable.

If we define fewer than 40 ranging exchanges per result, say 30, then we only proceed up to the 30th channel in the hopping sequence.

Where we define more than 40 exchanges per result, say 60, then the first 40 exchanges will hop over the full 40 channels, then the first 20 channels in the list will be replayed.

The frequency hopping sequence used in our ranging example. The red highlighted channels denote the BLE advertising channels.

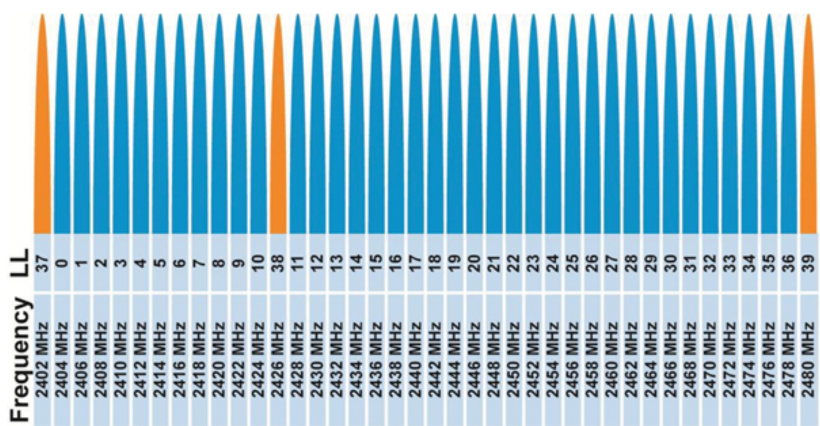


Figure 11. Frequency Hopping Scheme

8. Raw Results

The delineation between communication and ranging exchanges is important. This is because there are different metadata available that can help us improve the ranging accuracy available from each exchange.

8.1 Communication Phase

Given there is only one communication exchange per ranging measurement, we only have one of the following measurements per ranging result.

FEI The Frequency Error Indicator: Indicates the frequency error measured by the Requester upon receiving the communication packet from the Responder.

Communication RSSI The received signal strength of the communication packet.

8.2 Ranging Phase

Each ranging result is based upon multiple ranging exchanges, we have one of the quantities below per hopping channel, so multiple results per ranging measurement sequence.

Ranging Measurement The ranging measurement itself is a 2's compliment hexadecimal number proportional to the round trip time of flight.

Ranging RSSI The signal power seen by the Requester during reception of the ranging exchange.

9. Processing the Ranging Results

With these results collated from both the communication and ranging phases, we can perform the range calculation and correction process. Here we outline the full post processing flow as performed in the evaluation kit.

Here we also examine snippets of source code, to see the relevant code, please consult the SX1280 MBED site [4].

Step 1: Conversion from ranging round trip time of flight to distance in m (multiplication by ranging LSB and divide by 2). The raw result converted into metres is given in the line

```
val = ( double )complement2( valLsb, 24 ) / ( double ) LoRaBandwidth( ) * 36621.09375;
```

Where the value 36621.09375 = (c / (2^12)) / 2 comes from the formula documented in the datasheet [6]. valLsb is the ranging result in 2's compliment hexadecimal.

This conversion is applied to each RTToF result recovered from the SX1280 and stored in an array of RawRngResult[].

Step 2: Apply a correction based upon the FEI measurement from the communication exchange. This assumes that the frequency error is a linear curve with certain gradient. We stock the gradients as a function of Sf and BW as shown below:

```
const double   RNG_FGRAD_0400[] = { -0.148, -0.214, -0.419, -0.853, -1.686, -3.423 };
const double   RNG_FGRAD_0800[] = { -0.041, -0.811, -0.218, -0.429, -0.853, -1.737 };
const double   RNG_FGRAD_1600[] = { 0.103,  -0.041, -0.101, -0.211, -0.424, -0.87  };
```

For example for SF9, 1600 kHz the RngFeiFactor = -0.424. the corrected result is given by subtracting the calculated error as shown below:

```
RawRngResults[i] = RawRngResults[i] - ( RngFeiFactor * RngFei / 1000 );
```

Where RawRngResult[i] is the ith result in the array of raw results and RngFei is the FEI value in Hz read from the first LoRa communication packet.

Step3: The LNA Compensation of SF9 1600 kHz is as shown below. Here the RSSI Ranging is used directly as the look-up index for the distance correction. This correction is applied to each raw result:

```
// Generated LUT and Polynomial Values
#include "rangingCorrection_defines.h"
const double RangingCorrectionSF9BW1600[NUMBER_OF_FACTORS_PER_SFBW] =
{
    -2.8555,
    -2.8555,
    -2.8555,
    -2.8555,
    -2.8555,
    -2.8555,
    -2.8555,
    …
```

Step 4: Take the median value of the remaining filtered results. The code below is used to detemine the median, taking care to also cover the case where the two centremost values must be averaged (if there are an even number of ranging exchanges from which to derive the median).

```
if ((RngResultIndex % 2) == 0)
{
  median = (RawRngResults[RngResultIndex/2] + RawRngResults[(RngResultIndex/2) - 1])/2.0;
}
else
{
  median = RawRngResults[RngResultIndex/2];
}
```

From this point we no longer deal with an array of values, but the single filtered output from the median. Filtered at least in the sense that the median has been found empirically to give much better immunity to outliers that easily distort the average. This is very important in the case where we are seeking immunity from inaccurate channels.

Step 5: Apply a polynomial correction to linearize our results in the line-of-sight case. The polynomial coefficients are stored as shown below:

```
const RangingCorrectionPolynomes_t correctionRangingPolynomeSF9BW1600 = {
  .order = 7,
  .coefficients = {
  6.2781e-10,
  -2.6256e-07,
  3.7632e-05,
  -0.0022408,
  0.05252,
  0.47936,
  0.46183,
  }
};
```

These are derived from real field measurements. For guidance on collecting data in the field please see our ranging measurement app note [5]. Based upon field measurement we use curve fitting to derive the correction polynomial.

Because each SF and BW can have a different polynomial, here the correction is applied by simply looping through and accumulating the corrected polynomial – this allows us to accommodate different order polynomials for each SF and BW combination.

It is also important to note that the polynomial correction is only valid over the range of charachterised distances, here in the range upto 100 m.

```
const RangingCorrectionPolynomes_t *polynome = RangingCorrectionPolynomesPerSfBw[sf_index
][bw_index];
double correctedValue = 0.0;
double correctionCoeff = 0;
for(uint8_t order = 0; order < polynome->order; order++){
  correctionCoeff = polynome->coefficients[order] * pow(median, polynome
    >order - order - 1);
  correctedValue += correctionCoeff;
}
return correctedValue;
```

Complete the corrected value returned by this function is the final output of the corrected range for outdoor ranging measurements in our LoS scenario.

10. Before and After

The plot below shows the difference that the result processing correction makes to the absolute ranging accuracy. These results were measured at a fixed SF and bandwidth of SF9 and 1600 kHz respectively.

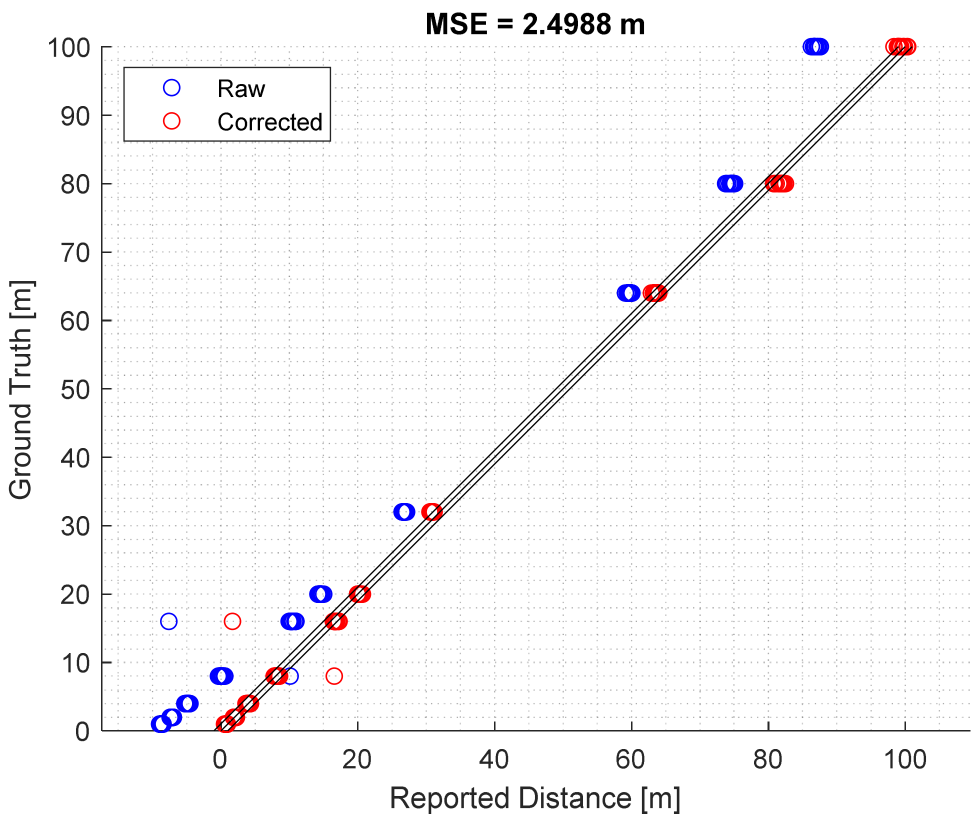


Figure 12. Before and After, the Ranging Accuracy in LoS using the techniques described here

The blue points show the result of simply taking the median ranging result over 40 frequency hopped exchanges in a LoS environment. The black line shows the ideal x=y scale that we should see between measured and Ground Truth distance.

With the full correction process applied as outlined above we obtain the results shown in red. It is important to note that the data set plotted here is not the training data with which the polynomial curve correction was generated but a separate cross validation data set.

In conclusion, it is possible to drastically reduce the bias of ranging results and increase the accuracy by correcting for known and measurable errors in the form of LNA gain and frequency error. Finally, typical environmental influences can be corrected through the use of statistical filtering and polynomial curve fitting to increase absolute accuracy.

11. References

[1] EVK Users Guide: https://www.semtech.com/uploads/documents/sx1280_81_userguide.pdf

[2] T. Cooper "Application Note: An Introduction to Ranging with the SX1280 Transceiver" AN 1200.29 Semtech Corporation, March 2017

[3] https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1199929&tag=1

[4] SX1280 MBED Site: https://os.mbed.com/components/SX1280RF1ZHP/

[5] T. Cooper "Application Note: How to Perform Ranging Tests with the SX1280 Development Kit

[6] SX1280 Datasheet https://www.semtech.com/uploads/documents/DS_SX1280-1_V2.2.pdf

12. Revision History

Version	Comments	Author	Date
1.0	Original version	T. Cooper	July 2019
1.1	Removal of Master-Slave references, Minor typographical correction, update footers	T. Cooper	June 2022

Important Notice

Contact Information

Semtech Corporation
Wireless & Sensing Products
200 Flynn Road, Camarillo, CA 93012
E-mail: sales@semtech.com
Phone: (805) 498-2111, Fax: (805) 498-3804
www.semtech.com