A survey on Time of Arrival Multilateration

Sidong Guo
Eshan Ramesh
*Georgia Institute of Technology VIP

***Abstract*—Time of Arrival(TOA) has attracted great interest in its application in indoor WiFi localization. TOA localization is a two step algorithm that includes timing information extraction with cross-correlation and multi-lateration. The difficulty involved in first part comes from limited sampling rate causing sub-sample TOA offset and multipath propagation skewing the correlation peak. This paper presents a sub-sample based one-way narrow-band TOA approach to mitigate aforementioned effects and multi-lateration techniques.**

***Index Terms*—Cramer-Rao Lower Bound, Estimation, Time of Arrival, Ultra Wide-band, Orthogonal Frequency Division-Multiplexing, Multi-lateration, Least Square, Maximum Likelihood, Multipath, Generalized Cross-Correlation.**

I. INTRODUCTION

The spread of 5G cellular network has incentivized a need for higher resolution indoor and outdoor localization navigation. In modern day multi-access cellular networks, cell-sectoring is employed for reducing intra and inter-cell interferences. Directional antennas with nulls at the other sectors' direction are often used for this purpose. In massive MIMO networks, base stations can employ up to hundreds of antennas on a beamforming array. The direction of the targeted mobile subscriber is therefore of great interest. To that end, the global positioning system (GPS) is widely used for outdoor localization purposes and can attain a high enough accuracy when there is no deep fade. In indoor environment local area networks become necessary due to shadowing and scattering. A natural alternative is to utilize already-deployed WiFi signals for indoor localization. The two most prevalent localization schemes use different indicators: signal strength and timing. Received Signal Strength Indicator (RSSI) methods determines the distance between emitter and receiver by mapping the signal strength difference on an established pathloss model. This method behaves poorly in non-line-of-sight (NLOS) propagation environments with Rayleigh fading and is more susceptible to interference from noise and multipaths. An alternative is therefore proposed whereby the entire RSS topology of the building is pre-measured in a grid-like structure termed RSS fingerprinting. This method requires enormous effort in terms of constructing the RSS map and is also vulnerable to major changes in the topology.

As such, timing-based methods are becoming increasingly popular in two major directions: Time of Arrival (TOA) and Time Difference of Arrival (TDOA). Time of Arrival concerns the absolute Time of Flight (TOF) between the transmitter and receiver, where ideally the internal clocks on both devices are synchronized to produce a cross-correlation measurement that determines the time-delay between the transmitted signal and received signal. Suppose there exists one transmitter and N receivers. Then, N circles of radius $\tau_n c$ can be constructed and the location of the transmitter can be determined by simple triangulation assuming no measurement noise. Similarly, TDOA uses a set of receivers and uses cross-correlation to determine the relative distance difference from each pair of receivers to the transmitter. Instead of circles, a set of hyperbolic functions can be constructed from TDOA measurements that ideally intersect at one point. TDOA has the advantage of not requiring absolute clock information and is therefore widely used in acoustic signal localization where the transmitted signal is some instantaneous speech signal. However, receivers have to be synchronized and multi-lateration for TDOA is more mathematically challenging than TOA, as we will see in later sections.

In this paper, we focus on the method of Time of Arrival localization. Generally, TOA has two major steps: Timing information acquisition, or 1D ranging, and calculation of localization based on the set of timing information obtained. To extract timing information, a number of methods have been proposed that include MAC layer, time domain physical layer and frequency domain physical layer methods [1]. This paper presents a sub-sample based method applicable for single carrier and multiple carrier transmission schemes. After the timing information is acquired, a set of nonlinear equations can be constructed that due to estimation error from previous steps, can't be solved in closed form. Therefore, an error term must be included in all nonlinear equations and be solved using either a linear approach like Least Square (LS) or subspace methods from multiple signal classification (MUSIC) [2] and estimation of signal parameter via rotational invariance techniques (ESPRIT) [3].

A sub-sample based TOA method relies on the assumption that, due to the limited sampling rate of the radios, the cross-correlation will produce a peak that is off by a time margin that is below one sampling period. This is not desirable in indoor localization, where precision to centimeters are required. Assuming a sampling rate of 100 MHz, where the sampling period is 10 ns, a sub-sample offset of 5 ns can produce a distance error of up to 1.5 meters. For multicarrier transmission schemes like OFDM, we suggest that the progressive phase shift between subcarriers can be used to calculate the delay that along with the cross-correlation results, to give a finer timing estimation.

In addition to the error caused by sub-sample offsets, dense indoor scattering shifts the correlation peak with a delay that corresponds to the difference in time between the line-of-sight (LOS) component and the strongest reflected multipath component.

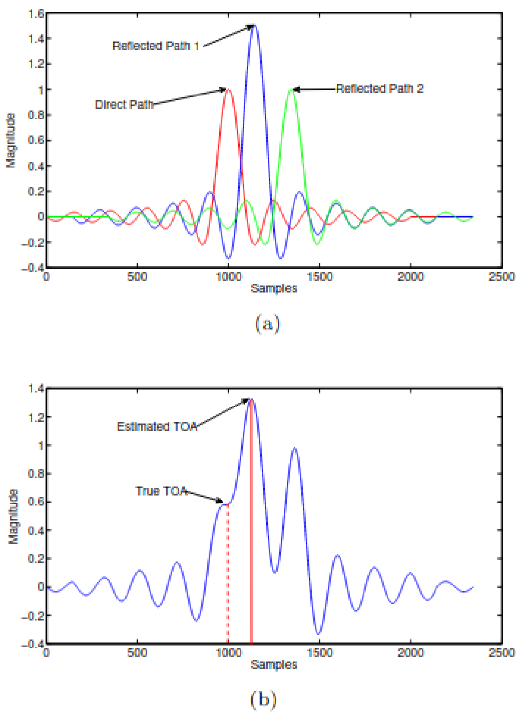

Fig. 1. Illustration of Multipath on Cross-Correlation.

To mitigate the effect of multipath, recent literature has suggested multipath cancellation with Maximum Likelihood Estimation [4]. However, MLE is computationally expensive and therefore the heuristic search technique has been proposed. The method searches for and subtracts each multipath component in descending order of its amplitude by cross-correlation and hopes to eventually restore the original signal [5] [6]. This method can be improved with Generalized Cross-Correlation (GCC) and using a combination of direct search for delay and closed form search for amplitude.

To sum up, this paper presents a one-way, sub-sampled based WiFi TOA method for both OFDM and single-carrier transmission. The rest of the paper is constructed as follows: Section II explains the system model, narrow band signals with flat fading are used that effectively has entire delay spread arriving within one symbol. For that reason, time-invariant channel models are assumed. This section also explains the sources of error. Section III explains the multipath cancellation method. Section IV presents how the progressive phase shift in OFDM subcarrier is used for acquiring sub-sample timing information and Section V lists and compares a number of multilateration techniques used for source localization after timing information is restored.

II. System Model

A. Narrow-band vs Wide-band

Typical indoor delay spread is from 40 ns to 70 ns, in a small lab environment this value tends to be on the lower side. The radios used are USRP 2901, transmitting with a center frequency of 2.4GHz and a bandwidth of 1MHz. The delay spread is therefore much smaller than inverse signal bandwidth and as a result, signal bandwidth is much smaller than coherence bandwidth, a time-invariant channel model can be assumed. All multipath component arrive within one symbol period and causes self-interference and depending on the sampling rate, may or may not be resolvable. Contrary to narrow-band transmission, many researches use ultra wide-band (UWB) so that time pulses span a relative short time, the multipath can easily be resolved if only one symbol is transmitted per delay spread, causing a loss of degrees of freedom. However, if symbols are transmitted continuously subsequent symbol transmission will cause inter-symbol interference (ISI) and frequency diversity techniques such as single-carrier equalization need to employed to counteract this effect. This paper assumes narrow band transmission, therefore the channel can be modeled as a time-invariant FIR filter with N taps (delay-coefficient pairs). The received signal is therefore a linear combination of its delayed versions

$$y[m] = \sum_{l} h_l[m]x[m-l] + w[m] \tag{1}$$

where $h_l[m]$ is the discrete-time equivalent baseband model of wireless channel, the goal is therefore to recover $x[m]$ from received signal $y[m]$

B. Source of error

1) Multipath: The effect of multipath, which in the VIP lab environment, is not so obvious, is such that in NLOS or LOS propagation environment where the LOS component is not the strongest, the Cross-Correlation peak will be shifted by a delay corresponding to the relative delay of that strongest component. Fundamentally, this is because of the distributive nature of linear convolution. Cross-correlation at the receiver between the received signal and the stored transmitted signal (usually a known preamble of a larger frame), is equivalent to a linear convolution between received signal and matched filter of the transmitted signal.

$$f \star \sum_{n=1}^{N} f(t-\tau_n) = \sum_{n=1}^{N} f \star f(t-\tau_n) \tag{2}$$

If the strongest component is the LOS component, then the peak will correspond to the time delay in an ideal situation.

2) Sampling rate: A limited sampling rate will cause two major problems, firstly, too low of a sampling rate will not be able to distinguish multiple peaks with a total delay spread no longer than 40 ns. Secondly, assume the first condition is met, in most cases, the incoming wave will arrive between two sampling points, in which a case of sub-sample offset will occur. Methods have been proposed to reduce this effect such as oversampling at reception. For OFDM, a technique that involves measuring the progressive phase shift of its subcarriers has a lot of potential.

3) Clock drift: Clock drift errors are also called synchronization error. It can arise either from time-synchronization error due to internal clocks having delays between each other or frequency-synchronization error when one oscillator has a slightly different clock rate than other internal oscillators, in which the error will accumulate through time. In VIP lab environment, since propagation distance is relatively small and both radios are connected to the same PC, this error is minimal.

III. Timing information extraction Techniques

There are two major steps to Timing information in time-domain based physical layer one way TOA method, first step is to remove multipath components via means of search and subtract cancellation, and after the final cross-correlation produce a coarse estimation, a sub-sample based method is used to calculate a finer estimation. The combined of which is the true timing information.

A. Multipath Cancellation

(Coming soon...)

B. Sub-sample TOA offset

(Coming soon...)

IV. Multilateration Techniques

Assume a 2-D system model (elevation of antennas usually have very minimal influences on localization accuracy for the sake of convenience I omit the Z-axis) where there are N receivers and 1 transmitter whose location information we are interested in. The timing information (TOF) has been extracted at each receiver based on the method in previous section. A set of non-linear equations can be easily constructed as follows:

$$c\tau_n = \sqrt{(x-x_n)^2 + (y-y_n)^2} + n_n \tag{3}$$

where τ_n is the nth TOA measurement, c is the speed of propagation, x and y are x,y coordinate of the target that we want to estimate, x_n and y_n are x,y coordinate of the nth receiver, which are known values, and n_n is the error term associated with the nth measurement, that is modelled as a zero-mean Gaussian random variable

We now expand the equation by squaring both sides and letting d_n be $c\tau_n$

$$d_n^2 = (x-x_n)^2 + (y-y_n)^2 + n_n^2 + 2n_n\sqrt{(x-x_n)^2 + (y-y_n)^2} \tag{4}$$

By demanding

$$k = x^2 + y^2 \tag{5}$$

and

$$k_n = x_n^2 + y_n^2 \tag{6}$$

equation (4) can be simplified to

$$\begin{gathered} d_n^2 = k + k_n - 2xx_n - 2yy_n + n_n^2 + \\ 2n_n\sqrt{(x-x_n)^2 + (y-y_n)^2} \end{gathered} \tag{7}$$

A. Linear Least Square–Subtract Approach

Here, I first demonstrate a linear least square approach that is also consistent with TDOA LLS. Subtract from equation (7) d_1^2 where

$$\begin{gathered} d_1^2 = k + k_1 - 2xx_1 - 2yy_1 + n_1^2 + \\ 2n_1\sqrt{(x-x_1)^2 + (y-y_1)^2} \end{gathered} \tag{8}$$

By letting

$$r_n = \sqrt{(x-x_n)^2 + (y-y_n)^2} \tag{9}$$

We arrive at the following equation

$$\begin{gathered} d_n^2 - d_1^2 = k_n - k_1 + 2x(x_1 - x_n) \\ +2y(y_1 - y_n) + n_n^2 - n_1^2 + 2n_n r_n - 2n_1 r_1 \end{gathered} \tag{10}$$

Now let

$$H = \begin{pmatrix} x_1 - x_2 & y_1 - y_2 \\ \vdots & \vdots \\ x_1 - x_n & y_1 - y_n \end{pmatrix} \tag{11}$$

and

$$N = \begin{pmatrix} n_1^2 - n_2^2 + 2r_1 n_1 - 2r_2 n_2 \\ \vdots \\ n_1^2 - n_n^2 + 2r_1 n_1 - 2r_n n_n \end{pmatrix} \tag{12}$$

$$x = \begin{pmatrix} d_2^2 - d_1^2 - k_2 + k_1 \\ \vdots \\ d_n^2 - d_1^2 - k_n + k_1 \end{pmatrix} \tag{13}$$

We can obtain a linear matrix equation in the form of

$$2H\theta = x + N \tag{14}$$

where $\theta = [x \quad y]^T$ is the matrix we want to estimate. Note that matrix N is the error matrix that can be approximated to zero due to error terms modeled as zero-mean Gaussian random variable, we therefore arrive at final linear expression

$$2H\theta \approx x \tag{15}$$

By definition, LSE is found by minimizing the metric

$$\begin{aligned} J(\theta) &= \sum_{n=0}^{N-1} (x[n] - s[n])^2 \\ &= (x - 2H\theta)^T (x - 2H\theta) \end{aligned} \tag{16}$$

J is then a quadratic function of θ and by expanding the term

$$J(\theta) = x^T x - 2x^T H\theta - 2\theta^T H^T x + 4\theta^T H^T H\theta \tag{17}$$

Taking the first derivative with respect to θ gives

$$\frac{\partial J(\theta)}{\partial \theta} = -4xH^T + 8H^T H\theta \tag{18}$$

By equating the first derivative to zero we have the estimator

$$\tilde{\theta} = \frac{1}{2}(H^T H)^{-1} H^T x \tag{19}$$

Note that from equation (7) the closed form estimator can be obtained directly by transforming equation (7) into linear matrix form without subtracting d_1^2. Here, we assume that there exist more receivers than dimensions to ensure that $H^T H$ is invertible. Because the number of dimensions is constant, the time complexity of the estimator is dominated by computing H^TH and is $O(n)$.

B. Linear Least Square-Direct Form

It can similarly be shown that the direct TOA form of linear least square estimator $\theta = [x \quad y \quad k]^T$ is the following

$$\tilde{\theta} = (A^TA)^{-1}A^Tb \tag{20}$$

where

$$A = \begin{bmatrix} -2x_1 & -2y_1 & 1 \\ \vdots & \vdots & \vdots \\ -2x_n & -2y_n & 1 \end{bmatrix} \tag{21}$$

and

$$b = \begin{bmatrix} d_1^2 - x_1^2 - y_1^2 \\ \vdots \\ d_n^2 - x_1^2 - y_1^2 \end{bmatrix} \tag{22}$$

Here, in order to make sure the matrix A^TA is invertible, there necessarily exist receivers greater than 3, which is a reasonable assumption for multilateration and that none of the receivers are either approximately or exactly, a scalar multiple position of each other. For direct form, the complexity also scales with multiplication and is $O(n)$.

C. Extension to TDOA

Without loss of conformity, the notations are the same in previous subsections and assuming a 2-D model. A set of non-linear equations can be obtained in the form of

$$\tau_{1n}c = \sqrt{(x-x_1)^2} - \sqrt{(x-x_n)^2} \tag{23}$$

where $\tau_{1n}c$ equals d_{1n}, is the distance difference between receiver n to transmitter and between receiver 1 to receiver, where receiver 1 is used as reference receiver. This is a positive value when receiver 1 is further away. Now with the help of the definition $\sqrt{(x-x_n)^2} = d_n$, equation (22) is transformed into

$$\begin{gathered} d_1^2 - d_n^2 = \sqrt{(x-x_1)^2} - \sqrt{(x-x_n)^2} \\ d_1^2 - d_n^2 = \\ x_1^2 + y_1^2 - x_n^2 - y_n^2 - 2(xx_1 + yy_1) + 2(xx_n + yy_n) \end{gathered} \tag{24}$$

Plugging into the left side $d_n = d_1 - d_{1n}$ and rearrange the right side

$$2d_1d_{1n} - d_{1n}^2 = k_1 - k_n + 2x(x_n - x_1) + 2y(y_n - y_1) \tag{25}$$

Rearranging the equation to have all unknowns on the right sides and by letting

$$H = \begin{pmatrix} x_2 - x_1 & y_2 - y_1 & -d_{12} \\ \vdots & \vdots & \vdots \\ x_n - x_1 & y_n - y_1 & -d_{1n} \end{pmatrix} \tag{26}$$

and

$$x = \begin{pmatrix} -d_{12}^2 - k_1 + k_2 \\ \vdots \\ -d_{1n}^2 - k_1 + k_n \end{pmatrix} \tag{27}$$

We have the linear expression

$$2H\theta = x \tag{28}$$

where θ is $[x \quad y \quad d_1]^T$ It follows that the optimum estimator based on LS is again

$$\tilde{\theta} = \frac{1}{2}(H^TH)^{-1}H^Tx \tag{29}$$

Note however, that H and x matrix have different definition and the estimator is only optimum all distances are comparable based on simulation results. The time complexity of evaluating $\tilde{\theta}$ remains $O(n)$ since the width of x and H are constant.

D. Weighted Linear Least Square

It is not hard to see that the main source of error from traditional linear least square approach comes from the transition from equation from (13) to (14) and the assumption that noise goes to zero. There are two examples in which the variance of noise are taken into account to improve the accuracy of the estimator. Firstly, Minimum Mean Square Error (MMSE) estimator in linear equalization uses information on noise variance and therefore produce better result than Zero-forcing (ZF). Secondly, based on equation (30), the maximum likelihood estimator seeks to maximize the metric $P(d_n;\theta)$, which in other word is to minimize the metric:

$$J(\theta) = (d_n - r(\theta))^TC^{-1}(\theta)(d_n - r(\theta)) \tag{30}$$

There is naturally a weighting factor in MLE that comes in the form of inverse of covariance matrix of noise. Similarly, the estimator that we attempt to minimize takes the form:

$$J(\theta) = (d_n - r(\theta))^TW(d_n - r(\theta)) \tag{31}$$

where W is the weighting factor associated with the error term and the optimum choice is shown to be in fact the inverse of error term covariance [7]. Depending the formulation of linear least square estimator, this term will be different. Here, I make an extension from subsection B and similar approach can be applied to subsection A. However, the nature of the error term N in A means that cross-correlation between two error terms are not statistically independent and therefore making the covariance matrix much more complicated. In B, the matrix form linear equation is:

$$A\theta + q = b \tag{32}$$

where q is the error term defined as:

$$\begin{pmatrix} n_1^2 + 2n_1\sqrt{(x-x_1)^2 + (y-y_1)^2} \\ \vdots \\ n_n^2 + 2n_n\sqrt{(x-x_n)^2 + (y-y_n)^2} \end{pmatrix} \tag{33}$$

Based on [8], the error term q contains the a quadratic term of noise variance, that becomes a combination of cubic or higher terms in autocorrelation, which is statistically insignificant. Therefore, the error term also be written as:

$$q = \begin{pmatrix} 2r_1n_1 \\ \vdots \\ 2r_nn_n \end{pmatrix} \tag{34}$$

The weighting matrix is the inverse of covariance matrix of error term $E[qq^T]$, which is found to be:

$$\begin{aligned} W &= [E[qq^T]]^{-1} \\ &= diag(4r_1^2\delta_1^2, 4r_2^2\delta_2^2, \ldots, 4r_n^2\delta_n^2)^{-1} \end{aligned} \tag{35}$$

It is a diagonal matrix because non-diagonal entries that are cross-correlation of the form:

$$E[2r_in_ir_jn_j] \tag{36}$$

equals to zero because noise n_i and n_j are statistically independent (Assuming based on space-time correlation, receivers are at least half of wavelength apart). Replacing equation (31) and equation (34) into (30) we have the estimator that attempts to minimize the metric:

$$J(\theta) = (b - A\theta)^T W(b - A\theta) \tag{37}$$

which is equivalent to:

$$\tilde{\theta} = \arg\min_{\theta} J(\theta) \tag{38}$$

Expanding the terms:

$$J(\theta) = b^TWb - b^TWA\theta - A^T\theta^TWb + A^T\theta^TWA\theta \tag{39}$$

The partial derivative is:

$$\frac{\partial J(\theta)}{\partial\theta} = -b^TWA - A^TWb + 2A^TWA\theta \tag{40}$$

Rearrange and equating the expression to zero, the estimator takes the form:

$$\tilde{\theta} = (A^TWA)^{-1}A^TWb \tag{41}$$

Now let's take a closer look at the matrix W in equation (40), it can be easily decomposed into matrix multiplications as follows:

$$W = \frac{1}{4}[R^TQR]^{-1} \tag{42}$$

Where matrix R contains the real distance as defined in equation (8), which obviously make it statistically unknown, therefore, often a choice is to replace r_n with d_n, the measured distance, making the matrix effectively:

$$W = \frac{1}{4}[D^TQD]^{-1} \tag{43}$$

In modern day communication, however, the receiver are often configured on the same array and the source is far away, there each d_n is approximately d_1, the weighting matrix can sometimes be expressed as:

$$W = \frac{1}{4}Q^{-1} \tag{44}$$

Because scaling does not affect the estimation result [7]. This approximation does not work in indoor environment where receivers are scattered.

In weighted linear least square, the invertibility of the term A^TWA also depends on whether matrix A attains full rank, because W is the diagonal matrix that doesn't change the invertibility of the product. Therefore, it shared the same assumption with LLS I and II.

Computing the inverse of an $n \times n$ matrix is typically $O(n^3)$, which dominates the $O(n)$ cost of computing the ordinary least squares result. Therefore, the overall time complexity of the estimator is $O(n^3)$

E. Two-step Linear Least Square Method

Two-Step WLLS seeks to improve the accuracy of WLLS and get closer to CRLB by making the following observation: There is actually a statistical dependence between the noise vector q and estimation vector θ, because by definition, q is a function of r_n, and by including the weighting term in form of error vector, the solution of WLLS are assuming independence which gives rise to error. We therefore construct the observation matrix as:

$$\tilde{u} = \begin{pmatrix} \tilde{x^2} \\ \tilde{y^2} \\ \tilde{k} \end{pmatrix} = \begin{pmatrix} x^2 \\ y^2 \\ k \end{pmatrix} + N \tag{45}$$

where N is the error vector and $[x^2 \quad y^2 \quad k]^T$ can be expressed as:

$$\begin{pmatrix} x^2 \\ y^2 \\ k \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x^2 \\ y^2 \end{pmatrix} \tag{46}$$

Now we make the following simplification to the error vector by noting:

$$N = \begin{pmatrix} \tilde{x^2} - x^2 \\ \tilde{y^2} - y^2 \\ \tilde{k} - k \end{pmatrix} = \begin{pmatrix} (\tilde{x} + x)(\tilde{x} - x) \\ (\tilde{y} + y)(\tilde{y} - y) \\ \tilde{k} - k \end{pmatrix} \tag{47}$$

The error vector can be further simplified by making the following observation:

$$\begin{aligned} (\tilde{x} + x)(\tilde{x} - x) &= (2x + e)(\tilde{x} - x) \\ &= 2x\tilde{x} - 2x^2 + e(x + e) - ex \\ &= 2x\tilde{x} - 2x^2 + e^2 \end{aligned} \tag{48}$$

Because the quadratic term of error term e^2 is statistically insignificant, based on this simplification, equation (46) transformed into:

$$N = \begin{pmatrix} 2x\tilde{x} - 2x^2 \\ 2y\tilde{y} - 2y^2 \\ \tilde{k} - k \end{pmatrix} \tag{49}$$

In last subsection it is noted that the optimum choice for weighting factor is the inverse of covariance matrix of the error vector, which in this case is N. Therefore, we are interested in the equivalent covariance matrix of:

$$N = MK = \begin{pmatrix} 2x & 2y & 1 \end{pmatrix} \begin{pmatrix} \tilde{x} - x \\ \tilde{y} - y \\ \tilde{k} - k \end{pmatrix} \tag{50}$$

Using above definitions, we can construct the linear expression:

$$\tilde{u} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x^2 \\ y^2 \end{pmatrix} + \begin{pmatrix} 2x & 2y & 1 \end{pmatrix} \begin{pmatrix} \tilde{x} - x \\ \tilde{y} - y \\ \tilde{k} - k \end{pmatrix} \tag{51}$$

In this case the optimum weighting factor is the error term N, we are therefore interested in the covariance matrix of N. Based on the relation [8], the optimum weighting matrix is given by:

$$\Theta = [diag[M_1, M_2, M_3](A^T W A)^{-1} diag[M_1, M_2, M_3]]^{-1} \tag{52}$$

And the estimator has the expression:

$$\tilde{\theta} = (Z^T \Theta Z)^{-1} Z^T \Theta u \tag{53}$$

where

$$Z = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \tag{54}$$

Evaluating this estimator has time complexity $O(n^3)$ due to the computation of the weighting matrix W.

The method can also be applied to TDOA demonstrated in [9]. After the estimation of $\tilde{u}$ is obtained, estimation of x and y are obtained by:

$$\tilde{\theta} = [\sqrt{\tilde{u}_1} \quad \sqrt{\tilde{u}_2}] \tag{55}$$

In 2SWLLS, in order to guarantee the invertibility of entire calculation chain, two conditions must be met: 1. The initial values used in equation (45) are obtained from LLSI or LLSII, in which case the matrix A has to attain full rank and all restriction on LLS applies. 2. The transmitter can not be modelled at a position, either approximately or exactly, on the two axis-es relative to the coordinate that is being used, because it will cause the M matrix to have a null eigenvalue and the operation in equation (52) to have singular or approximately singular value.

F. Constrained Weighted Linear Least Square

CWLLS is yet another method that tries to improve the accuracy of WLLS by including the constraint in equation (5), meaning the estimator jointly estimate x,y and k while posting additional constraint that trades complexity for accuracy. In matrix form, the constraint is[10]:

$$\tilde{\theta}^T P \tilde{\theta} + B^T \tilde{\theta} = 0 \tag{56}$$

where

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} B = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \tag{57}$$

To find the LSE subject to a constraint, it is best to use the technique of Lagrangian Multiplier. We try to minimize the metric:

$$J(\theta) = (A\theta - b)^T W (A\theta - b) + \lambda^T (\theta^T P \theta + B^T \theta) \tag{58}$$

The value of λ is found by putting the estimation $theta$ back to the constraint. Solving the above equation we have:

$$\tilde{\theta} = (A^T W A + \lambda P)^{-1} (A^T W b - \frac{\lambda}{2} B) \tag{59}$$

Upon substitution into the constraint, we have:

$$\begin{gathered} B^T (A^T W A + \lambda P)^{-1} (A^T W b - \tfrac{\lambda}{2} B) + \\ (A^T W A + \lambda P)^{-1} (b^T W A - \tfrac{\lambda}{2} B^T) \\ P (A^T W A + \lambda P)^{-1} (A^T W b - \tfrac{\lambda}{2} B) = 0 \end{gathered} \tag{60}$$

Based on the following eigenvalue decomposition property:

$$(A^T W A)^{-1} P = U \Lambda U^{-1} \tag{61}$$

The following simplification can be made:

$$(A^T W A + \lambda P)^{-1} = U (I + \lambda \Lambda)^{-1} U^{-1} (A^T W A)^{-1} \tag{62}$$

Now substitute equation 62 into 60 we have the expression in equation (63): Where the following substitutions are used:

$$c^T = B^T U = [c_1, c_2, c_3] \tag{64}$$

$$g = U^{-1} (A^T W A)^{-1} B = [g_1, g_2, g_3]^T \tag{65}$$

$$e^T = b^T W A U = [e_1, e_2, e_3] \tag{66}$$

$$f = U^{-1} (A^T W A)^{-1} A^T W b = [f_1, f_2, f_3]^T \tag{67}$$

The following identity is also used:

$$\Lambda = U^{-1} (A^T W A)^{-1} P U \tag{68}$$

Note that matrix P is of rank 2, therefore the eigenvalue decomposition of the composite matrix $(A^T W A)^{-1} P$ must necessarily produce only two eigenvalues. Therefore the diagonal matrix Λ also has rank 2. Also note that $(I + \lambda \Lambda)$is a diagonal matrix and therefore the inverse is easily calculable by inverting every entry on the diagonal. After the matrix equation is transformed into linear form, it is expressed in equation (68). (Note that this equation assumes the third eigenvalue is zero, which is not necessarily the case, when the zero eigenvalue is first eigenvalue, for instance, the equation needs to be adjusted).

Even with the simplification, solving the set of matrix equation in (68) requires the use of root-finding algorithm. In simulation, the equation is solved with matlab syms toolbox. The equation will produce five roots and the these roots are processed as following: All the complex roots are discarded and all the real roots are plugged into the equation (59) to find the corresponding θ. These θ are then plugged into the cost function in equation (58) to evaluate the results. The θ that produces the minimum cost function is selected as the true result with its corresponding Lagrangian multiplexer.

G. Approximate ML Method

Approximate ML method, or AML, also tries to capitalize on some assumption weakness of WLSS. The transition from equation (42) to (43) avoids the use of true distance by making an approximation, which gives rise to errors. AML seeks to improve this assumption by expanding on the ML expression. The vector probability density function of distance d_n with N element is [11]:

$$\begin{gathered} P(\tau_n; \theta) = \frac{1}{(2\pi^{\frac{N}{2}}) \det^{\frac{1}{2}} Q(\theta)} \\ \exp(-\tfrac{1}{2} (\tau_n - r(\theta)/c)^T Q^{-1}(\theta) (\tau_n - r(\theta)/c)) \end{gathered} \tag{70}$$

$$
\begin{gathered}
B^TU(I+\lambda\Lambda)^{-1}U^{-1}(A^TWA)^{-1}(A^TWb-\tfrac{\lambda}{2}B)+U(I+\lambda\Lambda)^{-1}U^{-1}(A^TWA)^{-1}(b^TWA-\tfrac{\lambda}{2}B^T)\\
PU(I+\lambda\Lambda)^{-1}U^{-1}(A^TWA)^{-1}(A^TWb-\tfrac{\lambda}{2}B)\\
=\\
B^TU(I+\lambda\Lambda)^{-1}U^{-1}(A^TWA)^{-1}A^TWb-B^TU(I+\lambda\Lambda)^{-1}U^{-1}(A^TWA)^{-1}\tfrac{\lambda}{2}B+\\
U(I+\lambda\Lambda)^{-1}U^{-1}(A^TWA)^{-1}b^TWAPU(I+\lambda\Lambda)^{-1}U^{-1}(A^TWA)^{-1}A^TWb-\\
U(I+\lambda\Lambda)^{-1}U^{-1}(A^TWA)^{-1}b^TWA\tfrac{\lambda}{2}B-\tfrac{\lambda}{2}B^TPU(I+\lambda\Lambda)^{-1}U^{-1}(A^TWA)^{-1}A^TWb+\\
U(I+\lambda\Lambda)^{-1}U^{-1}(A^TWA)^{-1}\tfrac{\lambda}{2}B^TPU(I+\lambda\Lambda)^{-1}U^{-1}(A^TWA)^{-1}\tfrac{\lambda}{2}B\\
=\\
c^T(I+\lambda\Lambda)^{-1}f-\tfrac{\lambda}{2}c^T(I+\lambda\Lambda)^{-1}g+e^T(I+\lambda\Lambda)^{-1}\Lambda(I+\lambda\Lambda)^{-1}f-\tfrac{\lambda}{2}e^T(I+\lambda\Lambda)^{-1}\Lambda(I+\lambda\Lambda)^{-1}g\\
-\tfrac{\lambda}{2}e^T(I+\lambda\Lambda)^{-1}\Lambda(I+\lambda\Lambda)^{-1}f+\tfrac{\lambda^2}{4}e^T(I+\lambda\Lambda)^{-1}\Lambda(I+\lambda\Lambda)^{-1}g=0
\end{gathered} \tag{63}
$$

$$
\begin{gathered}
c_3f_3-\tfrac{\lambda}{2}c_3g_3+\sum_{n=1}^{2}\tfrac{c_nf_n}{1+\lambda\kappa_n}-\tfrac{\lambda}{2}\sum_{n=1}^{2}\tfrac{c_ng_n}{1+\lambda\kappa_n}+\sum_{n=1}^{2}\tfrac{e_ng_n\kappa_n}{(1+\lambda\kappa_n)^2}-\tfrac{\lambda}{2}\sum_{n=1}^{2}\tfrac{e_ng_n\kappa_n}{(1+\lambda\kappa_n)^2}-\\
\tfrac{\lambda}{2}\sum_{n=1}^{2}\tfrac{c_nf_n\kappa_n}{(1+\lambda\kappa_n)^2}+\tfrac{\lambda^2}{4}\sum_{n=1}^{2}\tfrac{c_ng_n\kappa_n}{(1+\lambda\kappa_n)^2}=0
\end{gathered} \tag{69}
$$

Now by using the identity:

$$\tau_n = d_n/c \tag{71}$$

where τ is the measured TOA. The probability density function is now

$$
\begin{gathered}
P(d_n;\theta)=\\
\frac{1}{(2\pi^{\frac{N}{2}})\det^{\frac{1}{2}}Q(\theta)}\exp(-\tfrac{1}{2}(d_n-r(\theta))^TQ^{-1}(\theta)(d_n-r(\theta)))
\end{gathered} \tag{72}
$$

We try to minimize the metric:

$$J(\theta)=(d_n-r(\theta))^TQ^{-1}(\theta)(d_n-r(\theta))) \tag{73}$$

where $r(\theta)$ is the matrix of r_n. Now, instead of transforming the metric using approximation on ignoring the error vector, AML seeks to directly differentiate the metric with respect to θ. θ is a vector itself, therefore the differentiating result is a matrix that produces two equations. We begin with expanding the metric as:

$$J(\theta)=d_n^TQ^{-1}d_n-Q^{-1}(2d_n^Tr(\theta))+r(\theta)^TQ^{-1}r(\theta) \tag{74}$$

By differentiating $J(\theta)$ with respect to x, we have:

$$\frac{\partial J(\theta)}{\partial x}=\frac{-d_n^T(x-x_n)}{r(\theta)}+(x-x_n)=0 \tag{75}$$

Based on equation (3), we have the following:

$$\frac{\partial J(\theta)}{\partial x}=\frac{-(n_n)(x-x_n)}{r(\theta)}=0 \tag{76}$$

Expanding the vector:

$$\sum_{n=1}^{N}\frac{(r_n-d_n)(x-x_n)}{r_n}=0 \tag{77}$$

Therefore, similarly:

$$\sum_{n=1}^{N}\frac{(r_n-d_n)(y-y_n)}{r_n}=0 \tag{78}$$

By substituting the below equality into two above equations:

$$r_n-d_n=\frac{r_n^2-d_n^2}{r_n+d_n} \tag{79}$$

We have:

$$\sum_{n=1}^{N}\frac{(r_n^2-d_n^2)(x-x_n)}{r_n(r_n+d_n)}=0 \tag{80}$$

$$\sum_{n=1}^{N}\frac{(r_n^2-d_n^2)(y-y_n)}{r_n(r_n+d_n)}=0 \tag{81}$$

The equivalent matrix equation is

$$2\begin{pmatrix}\sum g_nx_n & \sum g_ny_n\\ \sum h_nx_n & \sum h_ny_n\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}\sum g_n(k+k_n-d_n^2)\\ \sum h_n(k+k_n-d_n^2)\end{pmatrix} \tag{82}$$

where

$$g_n=\frac{x-x_n}{r_n(r_n+d_n)} \tag{83}$$

$$h_n=\frac{y-y_n}{r_n(r_n+d_n)} \tag{84}$$

r_n is true distance and d_n is measured distance. The value of x and y are determined through numerical iteration on these equations as the followings:

1. An initial estimation is produced by another method (LLS for efficiency)
2. g_n and h_n are computed from initial values
3. By using the above values in equation (81), an LS can be estimated in terms of k, the only other variable dependent on x and y
4. Based on the relation in equation (5), choose the value of k that minimizes the cost function and replace k with new value in next loop.
5. Repeat step one, but with the newly estimated x and y
6. After N loops, a set of x and y can be calculated and the combination that produces the smallest cost function is selected as the true value

(Coming soon...)

H. Newton-Raphson Method

All the above methods have attempted to transform a set of non-linear equations into linear equations and using linear least square to search for an answer. Here we present some

$$\nabla^2 r_n(\theta) = \nabla J(\theta) = \begin{pmatrix} \frac{(y-y_1)^2}{((x-x_1)^2+(y-y_1)^2)^{\frac{3}{2}}} & \frac{(x-x_1)(y-y_1)}{((x-x_1)^2+(y-y_1)^2)^{\frac{3}{2}}} \\ \vdots & \vdots \\ \frac{(y-y_n)^2}{((x-x_n)^2+(y-y_n)^2)^{\frac{3}{2}}} & \frac{(x-x_n)(y-y_n)}{((x-x_n)^2+(y-y_n)^2)^{\frac{3}{2}}} \\ \frac{(x-x_1)(y-y_1)}{((x-x_1)^2+(y-y_1)^2)^{\frac{3}{2}}} & \frac{(x-x_1)^2}{((x-x_1)^2+(y-y_1)^2)^{\frac{3}{2}}} \\ \vdots & \vdots \\ \frac{(x-x_n)(y-y_n)}{((x-x_n)^2+(y-y_n)^2)^{\frac{3}{2}}} & \frac{(x-x_n)^2}{((x-x_n)^2+(y-y_n)^2)^{\frac{3}{2}}} \end{pmatrix} \tag{91}$$

non-linear methods that approaches global minimum through means of iterative search. The method of grid search is also an interesting alternative, but for MLE, it is only realizable when the dimension is small. Based on the result from previous subsection, we start with a set of nonlinear equations as determined through MLE, a slight modification of equation (60) is:

$$\frac{\partial J(\theta)}{\partial \theta_j} = 2Q^{-1}\sum_{n=1}^{N-1}[r_n(\theta) - d_n]\frac{\partial r_n(\theta)}{\partial \theta_j} = 0 \tag{85}$$

where we define the N by p Jacobian Matrix as:

$$[\frac{\partial r(\theta)}{\partial \theta}]_{nj} = \frac{\partial r_n(\theta)}{\partial \theta_j} \tag{86}$$

for n=0,1,...,N-1 and j=1,2,...,p. The inverse of error matrix will be cancelled out and we define:

$$g(\theta) = \frac{\partial r_n(\theta)^T}{\partial \theta}(d_n - r_n(\theta)) \tag{87}$$

The equivalent N by 2 matrix for N sensors in TOA localization is

$$[\frac{\partial r_n(\theta)}{\partial \theta_i}] = \begin{pmatrix} \frac{x-x_1}{\sqrt{(x-x_1)^2+(y-y_1)^2}} & \frac{y-y_1}{\sqrt{(x-x_1)^2+(y-y_1)^2}} \\ \vdots & \vdots \\ \frac{x-x_n}{\sqrt{(x-x_n)^2+(y-y_n)^2}} & \frac{y-y_n}{\sqrt{(x-x_n)^2+(y-y_n)^2}} \end{pmatrix} \tag{88}$$

Then in matrix form, the equivalent equation is:

$$\frac{\partial r(\theta)^T}{\partial \theta}(d - r(\theta)) = 0 = g(\theta) = \begin{pmatrix} \sum_{n=1}^{N} \frac{(d_n-r_n)(x-x_n)}{r_n} \\ \sum_{n=1}^{N} \frac{(d_n-r_n)(y-y_n)}{r_n} \end{pmatrix} \tag{89}$$

The Newton-Raphson method searches to iterate:

$$\theta_{k+1} = \theta_k - (\frac{g(\theta)}{\theta})^{-1} g(\theta)|_{\theta=\theta_k} \tag{90}$$

The above iteration involves the Jacobian of $g(\theta)$, which is the scaled Hessian matrix of $J(\theta)$. The Hessian matrix has the expression in equation (90):

$$\frac{\partial [g(\theta)]_i}{\partial \theta_j} = \frac{\partial}{\partial \theta_j}[\sum_{n=1}^{N-1}(d_n - r_n(\theta))\frac{\partial r_n(\theta)}{\partial \theta_i}] \tag{92}$$

Expanding it we have:

$$\frac{\partial [g(\theta)]_i}{\partial \theta_j} = \sum_{n=1}^{N-1}[(d_n - r_n(\theta))\frac{\partial^2 r_n(\theta)}{\partial \theta_i \partial \theta_j} - \frac{\partial r_n(\theta)}{\partial \theta_i}\frac{\partial r_n(\theta)}{\partial \theta_j}] \tag{93}$$

The 2 by 2 scaled Hessian matrix is therefore of the form in equation (94).

Substitute this relation into equation (70), we have the final Newton-Raphson iteration in equation (75). Note that in Newton-Raphson method, the initial value for θ needs to be estimated based on some other algorithm, after which the iteration can be applied to solve for the nonlinear system of equations. When $r(\theta)$ is equal to d_n, this reduces to linear equation and

$$\theta_{k+1} = (H^T H)^{-1} H^T x \tag{95}$$

Here we briefly demonstrate another iterative search method derived from Newton-Raphson called Scoring Method where Fisher Information matrix is used instead of Jacobian of G matrix:

$$\theta_{k+1} = \theta_k - I(\theta)^{-1} g(\theta)|_{\theta=\theta_k} \tag{96}$$

I. Gauss' Method

Gauss' Method attempts to linearize the model against that some nominal value of estimation value. This is different from Newton-Raphson where the derivative of $J(\theta)$ is linearized under the current iteration. It is worth nothing, however, that unlike grid search methods, iterative search methods usually have convergence problem since the metric have multiple local minimum alongside a global minimum.

J. Grid Search Method

Grid search method first finds a set of initial estimation and then center the initial measurement at the center of a square grid to seach for a minimum that minimizes the cost function. We choose a range of 100 meters and a step of 0.1 meter.

K. Subspace Method

Subspace method is a linear algebra based method that attempts to separate noise subspace from signal space. This method is shown to be sub-optimal compared to linear least square approaches in [8].

$$H(\theta) = \begin{pmatrix} \sum_{n=1}^{N} \frac{d_n(y-y_n)^2 - r_n^3}{r_n^3} & \sum_{n=1}^{N} \frac{(d_n-2r_n)(x-x_n)(y-y_n)}{r_n^3} \\ \sum_{n=1}^{N} \frac{(d_n-2r_n)(x-x_n)(y-y_n)}{r_n^3} & \sum_{n=1}^{N} \frac{d_n(x-x_n)^2 - r_n^3}{r_n^3} \end{pmatrix} \tag{94}$$

$$\theta_{k+1} = \theta_k + (\sum_{n=1}^{N-1} \frac{\partial r_n(\theta)}{\partial \theta_i} \frac{\partial r_n(\theta)}{\partial \theta_j} + \sum_{n=0}^{N-1} \frac{\partial^2 r_n(\theta)}{\partial \theta_i \partial \theta_j} (d_n - r_n(\theta)))^{-1} \frac{\partial r_n(\theta)^T}{\partial \theta} (d_n - r_n(\theta)) \tag{97}$$

L. Cramer-Rao Lower Bound

An unbiased estimator's performance can be evaluated by its relation to CRLB, that determines the lowest variance performance of an ideal estimator. In this subsection we derive the CRLB for multilateration problem and in subsequent simulations compare with other estimators' performance.

From equation (2) we have error equation:

$$d_n = r_n + n_n \tag{98}$$

where n_n is again the zero-mean Gaussian random variable with variance δ^2.Then d_n has the Gaussian distribution $d_n \sim \mathcal{N}((r(\theta), \delta^2(\theta))$, the vector extension of the probability density function with N element is then [11] [12].

$$P(d_n;\theta) = \frac{1}{(2\pi^{\frac{N}{2}}) \det^{\frac{1}{2}} C(\theta)} \exp(-\frac{1}{2}(d_n - r(\theta))^T C^{-1}(\theta)(d_n - r(\theta))) \tag{99}$$

where $r(\theta)$ is the N by 1 mean vector with the relation equation (8) and $C^{-1}(\theta)$ is the inverse of the N by N covariance matrix, which is a diagonal matrix because different receiver's signal are independent and therefore only auto-correlation are non-zero. It's diagonal are by definition, the variance of individual receive measurement. Cramer-Rao Lower Bound is defined as the inverse of Fisher Information Matrix, where element of Fisher Information Matrix for this case is:

$$[I(\theta)]_{ij} = -E[\frac{\partial^2 \ln p(d_n;\theta)}{\partial \theta_i \partial \theta_j}] \tag{100}$$

where $\theta = [x \quad y]^T$ is the estimation matrix containing two variable, therefore the Fisher Information matrix is a 2 by 2 matrix. Note that by definition of equation (8), the dependence of d_n on θ is clear. By combining equation (30) and (31), Fisher Information Matrix can be found for general Gaussian Case as:

$$\begin{aligned} [I(\theta)]_{ij} = [\tfrac{\partial r(\theta)}{\partial \theta_i}]^T C^{-1}(\theta) [\tfrac{\partial r(\theta)}{\partial \theta_j}] + \\ \tfrac{1}{2} tr[C^{-1}(\theta) \tfrac{\partial C(\theta)}{\partial \theta_i} C^{-1}(\theta) \tfrac{\partial C(\theta)}{\partial \theta_j}] \end{aligned} \tag{101}$$

And since $C(\theta)$ is independent of θ. Fisher Information can be reduced to:

$$[I(\theta)]_{ij} = [\frac{\partial r(\theta)}{\partial \theta_i}]^T C^{-1}(\theta) [\frac{\partial r(\theta)}{\partial \theta_j}] \tag{102}$$

By definition, vector $r(\theta)$'s partial derivative matrix is given by:

$$[\frac{\partial r(\theta)}{\partial \theta_i}] = \begin{pmatrix} \frac{x-x_1}{\sqrt{(x-x_1)^2+(y-y_1)^2}} & \frac{y-y_1}{\sqrt{(x-x_1)^2+(y-y_1)^2}} \\ \vdots & \vdots \\ \frac{x-x_n}{\sqrt{(x-x_n)^2+(y-y_n)^2}} & \frac{y-y_n}{\sqrt{(x-x_n)^2+(y-y_n)^2}} \end{pmatrix} \tag{103}$$

And because covariance matrix $C(\theta)$ is diagonal matrix, its inverse is the inverse of very diagonal component. We have the following expression:

$$C^{-1} = \begin{pmatrix} \frac{1}{\delta_1^2} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\delta_n^2} \end{pmatrix} \tag{104}$$

Substituting equation (33) and (34) into (32), the Fisher Information matrix is expressed as:

$$I(\theta) = \begin{pmatrix} \sum_{n=1}^{N} \frac{(x-x_n)^2}{\delta_n^2 r_n^2} & \sum_{n=1}^{N} \frac{(x-x_n)(y-y_n)}{\delta_n^2 r_n^2} \\ \sum_{n=1}^{N} \frac{(x-x_n)(y-y_n)}{\delta_n^2 r_n^2} & \sum_{n=1}^{N} \frac{(y-y_n)^2}{\delta_n^2 r_n^2} \end{pmatrix} \tag{105}$$

The lower bound for x and y estimation are therefore denoted by its (1,1) and (2,2) element, the Cramer-Rao Lower Bound can be expressed as:

$$var(\hat{\theta}) \geq \begin{pmatrix} [I^{-1}(\theta)]_{1,1} \\ [I^{-1}(\theta)]_{2,2} \end{pmatrix} \tag{106}$$

V. Simulation Results

The Simulation is done with both MATLAB Phased Array Toolbox through a AWGN channel and numerical method. There are several clarifications:

1.The propagation environment is modelled with 4 anchor points at positions (0,0),(1000,0),(0,1000),(1000,1000). Values are abstractions that can be easily changed. The user is assumed to be in the area enclosed by four APs.

2. The measurement variance is related to the noise variance with the distance square. In other words, assuming a fixed SNR for a transmitter-receiver power (fixed power and noise variance), the larger the distance travelled, larger the measurement variance, which increases as a quadratic function of distance. Fundamentally, this is because Friis Transmission Equation explains that the received power decays with distance square. Therefore, since the distances to each AP are different, in the interest of uniformity, the SNRs in each direction are kept constant. In Multi-user MIMO term, this means the transmitter allocates more power to users that are further away.

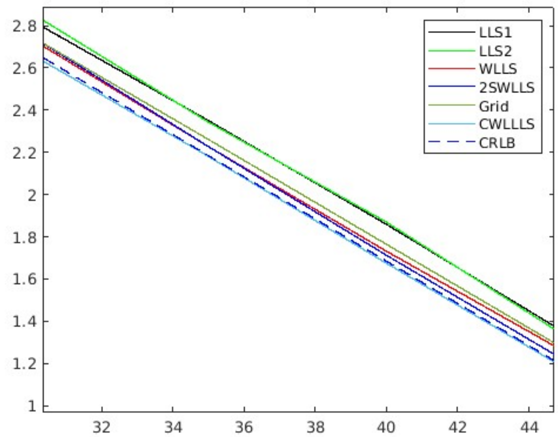

Fig. 2. Simulation of first four linear approaches

3. Each multilateration method is averaged over 1000 independent channel realizations.

4. X axis in unit of SNRdB, where as y axis unit is variance, or mean square error defined as:

$$E[(\tilde{x} - x)^2 + (\tilde{y} - y)^2] \tag{107}$$

The plot for CRLB can be obtained through equation (101).

As a result, 2SWLLS is better than WLLS and significantly better than the rest constantly, two forms of LLS are about similar in performance but trails grid search and WLLS. CWLLS approaches CRLB but has an overall high complexity. Therefore, due to CWLLS's complexity, at high SNR 2SWLLS is more appropriate.

VI. Conclusion

Acknowledgment

References

[1] A. Makki, A. Siddig, M. Saad, and C. Bleakley, "Survey of WiFi positioning using time-based techniques," *Comput. Networks*, 2015. DOI: 10.1016/j.comnet.2015.06.015.

[2] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, Mar. 1986, Conference Name: IEEE Transactions on Antennas and Propagation, ISSN: 1558-2221. DOI: 10.1109/TAP.1986.1143830.

[3] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 984–995, Jul. 1989, Conference Name: IEEE Transactions on Acoustics, Speech, and Signal Processing, ISSN: 0096-3518. DOI: 10.1109/29.32276.

[4] M. Win and R. Scholtz, "Characterization of ultra-wide bandwidth wireless indoor channels: A communication-theoretic view," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 9, pp. 1613–1627, Dec. 2002, Conference Name: IEEE Journal on Selected Areas in Communications, ISSN: 1558-0008. DOI: 10.1109/JSAC.2002.805031.

[5] A. Zayets and E. Steinbach, "Robust WiFi-based indoor localization using multipath component analysis," *2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2017. DOI: 10.1109/IPIN.2017.8115943.

[6] A. Makki, A. Siddig, M. Saad, J. R. Cavallaro, and C. J. Bleakley, "Indoor localization using 802.11 time differences of arrival," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 3, pp. 614–623, Mar. 2016, Conference Name: IEEE Transactions on Instrumentation and Measurement, ISSN: 1557-9662. DOI: 10.1109/TIM.2015.2506239.

[7] J. Chen, R. Hudson, and K. Yao, "Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field," *IEEE Transactions on Signal Processing*, vol. 50, no. 8, pp. 1843–1854, Aug. 2002, Conference Name: IEEE Transactions on Signal Processing, ISSN: 1941-0476. DOI: 10.1109/TSP.2002.800420.

[8] S. Ravindra and S. N. Jagadeesha, "Time of arrival based localization in wireless sensor networks : A linear approach," *Signal & Image Processing : An International Journal*, vol. 4, no. 4, pp. 13–30, Aug. 31, 2013, version: 1, ISSN: 22293922, 0976710X. DOI: 10.5121/sipij.2013.4402. arXiv: 1403.6697. [Online]. Available: http://arxiv.org/abs/1403.6697 (visited on 04/07/2022).

[9] Y. Chan and K. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Transactions on Signal Processing*, vol. 42, no. 8, pp. 1905–1915, Aug. 1994, Conference Name: IEEE Transactions on Signal Processing, ISSN: 1941-0476. DOI: 10.1109/78.301830.

[10] K. Cheung, H. So, W.-K. Ma, and Y. Chan, "Least squares algorithms for time-of-arrival-based mobile location," *IEEE Transactions on Signal Processing*, vol. 52, no. 4, pp. 1121–1130, Apr. 2004, ISSN: 1941-0476. DOI: 10.1109/TSP.2004.823465.

[11] S. M. Kay, *Fundamentals of statistical signal processing: estimation theory*. USA: Prentice-Hall, Inc., 1993, 595 pp., ISBN: 978-0-13-345711-7.

[12] Y.-T. Chan, H. Yau Chin Hang, and P.-c. Ching, "Exact and approximate maximum likelihood localization algorithms," *IEEE Transactions on Vehicular Technology*, vol. 55, no. 1, pp. 10–16, Jan. 2006, Conference Name: IEEE Transactions on Vehicular Technology, ISSN: 1939-9359. DOI: 10.1109/TVT.2005.861162.