

Doubling the Accuracy of Indoor Location: Frequency Diversity

Berthold K.P. Horn^a

Abstract—Determination of indoor location based on fine time measurement (FTM) of the round trip time (RTT) of a signal between an initiator (smartphone) and a responder (Wi-Fi access point) enables a number of applications. However, the accuracy currently attainable — standard deviations of 1–2 meter in distance measurement under favorable circumstances — limits the range of possible application. A first responder, for example, may not be able to unequivocally determine on which floor someone in need of help is in a multi-story building.

The error in location depends on several factors, including the bandwidth of the RF signal, delay of the signal due to the high relative permittivity of construction materials, and the geometry-dependent “noise gain” of location determination. Errors in distance measurements have unusual properties that are exposed here for the first time. Improvements in accuracy depend on understanding all of these error sources.

This paper introduces “frequency diversity,” a method for doubling the accuracy of indoor location determination using weighted averages of measurements with uncorrelated errors obtained in different channels. The properties of this method are verified experimentally with a range of responders. Finally, different ways of using the distance measurements to determine indoor location are discussed and the Bayesian grid update method shown to be more useful than others, given the non-Gaussian nature of the measurement errors.

Index Terms—indoor location, fine time measurement, round trip time, FTM, RTT, IEEE 802.11mc, IEEE 802.11-2016, time diversity, spatial diversity, bandwidth diversity, frequency diversity, Bayesian grid, observation model, transition model

I. OVERVIEW

Determining location accurately indoors, where GPS is not reliable, has many potential applications and has been of interest for some time [1]–[5]. One of the latest entries in this effort is fine time measurement (FTM) of round trip time (RTT) as specified in the 2016 update of the IEEE 802.11 Wi-Fi standard (also referred to as IEEE 802.11mc) [6].

We start by briefly discussing alternative methods for indoor location determination. This is followed by an exploration of the error sources in indoor location determination, particularly those for FTM RTT. Then, different attempts at getting more accurate distance measurements using uncorrelated error contributions are discussed and the frequency diversity method introduced. Experimental results confirm the prediction that frequency diversity can double the accuracy of indoor location, given that there are six non-overlapping 80 MHz channels available in the 5 GHz band [7]. Finally, various methods for turning distance measurements into locations are explored and the Bayesian grid update method shown to be well suited

to the task given the unusual nature of the error in distance measurement.

II. INTRODUCTION

The contributions of the research presented here are as follows: This paper introduces: (1) “frequency diversity” — a method for doubling the accuracy of FTM RTT distance measurements; (2) the “position-dependent error” texture surface — a new way of understanding the nature of the errors in FTM RTT distance measurement; (3) analysis of the unusual properties of the errors in distance measurement in terms of properties of super-resolution algorithms; (4) recognition of the serious impact of signal delay in common building materials resulting from their high relative permittivity — arguably more important than possible multi-path effects;

III. BACKGROUND

A number of different methods for indoor location have been explored, some of which make use of properties of existing radio frequency signals emitted by Wi-Fi access points and Bluetooth beacons (For a quick review see first few chapters of [8]).

A. Signal Strength

Perhaps the simplest approach is to measure the signal strength (RSSI) of a Wi-Fi access point (AP) at a hand-held device such as a smartphone (STA).

Unfortunately, the inverse square law causes the accuracy to drop off inversely with distance and so the measurements are at best only useful close to the AP. Furthermore, signal strength is affected by many factors other than distance. This includes the current power level of the AP and standing waves resulting from interference between signals reflected from material outside the line of sight (LOS) between the transmitter and the receiver. As a result, the relationship between distance and signal strength is not monotonic and not invertible (Fig. 1).

B. Fingerprinting

In light of this, another way of using signal strengths has been explored. So-called “fingerprinting” methods depend on careful mapping signal strengths from *several* sources in the volume of interest. Signal strengths do not vary much with time as long as objects (and people) are not moved. When they *are* moved, the finger-print data may have to be remeasured. Measuring signal strengths of multiple sources at many points in a volume is tedious and does not scale well.

^a Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA 02139, USA, e-mail: bkph@csail.mit.edu

* The corresponding author is Berthold K.P. Horn.

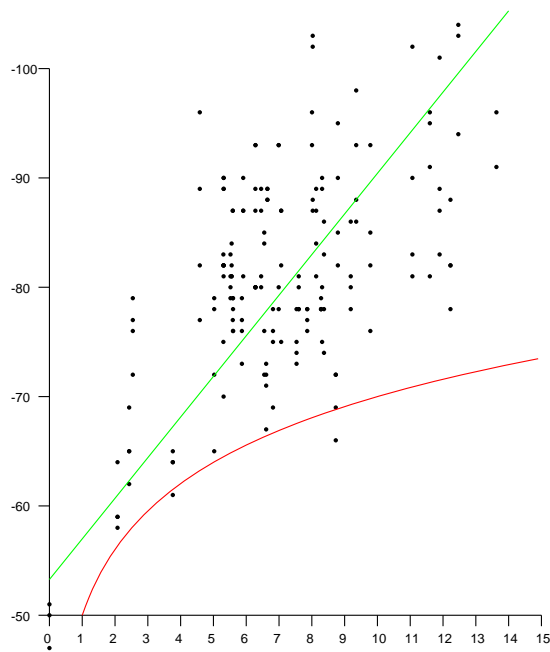


Fig. 1. Scattergram of Signal Strength (RSSI) versus distance in typical three-level wooden building. Horizontal axis: distance between smartphone and APs (in meters). Vertical axis: Signal Strength (in dBm). Red curve: expected inverse square law dependence ($-50 - 20 \log(R)$ dBm). Green line: linear fit ($-53 - 3.7R$ dBm). Because of the large scatter, it should be clear that RSSI is not very useful for estimating distance.

C. Channel State Information

A simple model of the transfer function of the channel from transmitter to receiver is a weighted sum of impulses, each representing a signal that travelled along a different path. If there is a clear line of sight (LOS), the first impulse in that sum is due to the LOS path. So, if the response function can be determined, the first impulse can be isolated and used to determine the time of flight. A network analyzer can be used to measure the frequency response of a communication channel, which is the Fourier transform of the impulse response. It is, however, not practical to deploy network analyzers, in part because they require physical access to both the transmitter and the receiver.

D. Orthogonal Frequency-Division Multiplexing

In the case of orthogonal frequency-division multiplexing (OFDM) signaling — used in all but the earliest IEEE 802.11 physical layer (PHY) standards [6] — the channel is divided into many equi-spaced narrow subchannels. In operation, the response of each subchannel needs to be known and consequently is estimated continuously. This channel state information (CSI) is potentially available (at least since IEEE 802.11n using e.g. Intel 5300). It is a low-resolution approximation to what a network analyzer would measure. Unfortunately, at this point no widely used platform provides access to the CSI.

E. Angle of Arrival

With many antennas, a base station can estimate the direction of arrival of the signal from user equipment (smartphone).

High angular resolution is required since the position error is the product of the distance and the angular resolution. Thus unless distances are very small, base stations with many antennas (and perhaps many radio chains) are needed, since angular resolution varies inversely with the number of antennas. There are also some privacy issues, since here a critical part of the location determination is done by the base stations, not the smartphone.

F. FTM RTT IEEE 802.11-2016

Finally, we come to fine time measurement (FTM) of round trip time (RTT) as specified in IEEE 802.11-2016 (also referred to as 802.11mc) [6]. One might expect this to overcome the limitations of other methods, since time of arrival is based on the *first* signal component, and so should be immune to multi-path problems, such as interference and standing waves.

Access to FTM RTT measurements has been provided on the Android platform since 2018 (Android 9 / Pie), although initially few smartphones and Wi-Fi APs supported the protocol (see also Appendix B).

Experimentally one finds that the distance measurements provided by FTM RTT may have standard deviations of 1–2 meter under favorable circumstances. This is fine for some applications but not others. It is important to understand the underlying causes of the observed errors in distance.

IV. NATURE OF THE ERROR

In FTM RTT, the error — difference between measurement and the actual distance — can be thought of as having several components, which behave very differently. It is important to understand these contributions to the overall error e , since they need to be dealt with in different ways.

$$e = m(c; \dots) + E(\mathbf{r}; c; \dots) + o(c; \dots) \quad (1)$$

Here $m(c; \dots)$ is “measurement noise” (see below) which depends on the channel c (i.e. frequency) and other factors, while $E(\mathbf{r}; c; \dots)$ is the “position-dependent error” (see below) which depends on position \mathbf{r} , the channel c and other factors, while $o(c; \dots)$ is the offset (see below) which depends on the channel c , type of initiator, type of responder etc.

All of the above also depend on the bandwidth, but, except where noted below, we’ll assume use of the highest bandwidth at which FTM RTT is supported by both the initiator and the responder (currently 80 MHz) because that normally leads to the highest accuracy.

Further, where there is a dependence on position as indicated above, there is also dependence on orientation, which we will not continue to refer to explicitly from here on.

A. “Measurement Error”

Remarkably small spreads in results are observed when measurements are repeated without changes in position (or orientation) of initiator and responder, in a fixed environment. In this case, the standard deviation (e.g. 0.1–0.2 meter under favorable circumstances) is considerably smaller than the actual error in distance measurement (which is typically greater

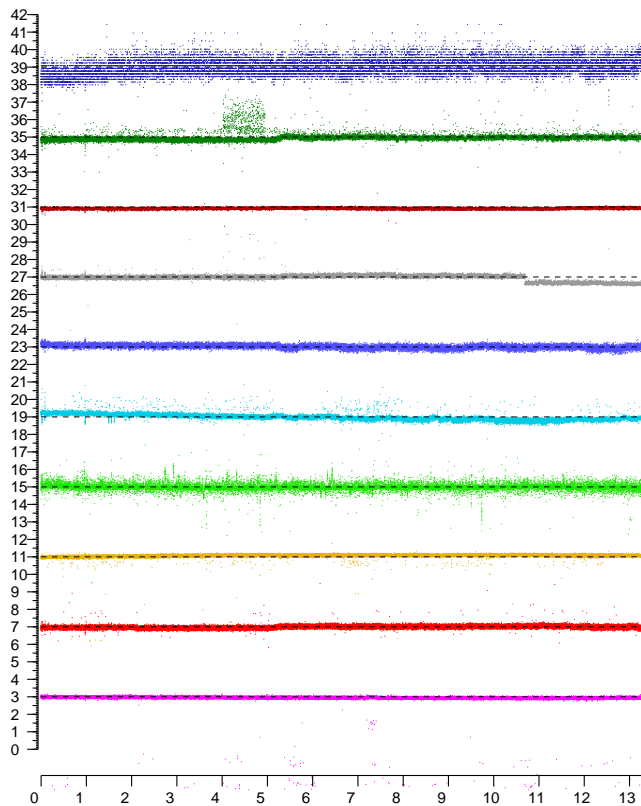


Fig. 2. Sample measurements using ten different responders in fixed positions. (Top plot is for an access point operating in the 2.4 GHz band, the rest in the 5 GHz band). Horizontal axis: time in hours. Vertical axis: distances in meters (individual plots are offset vertically to avoid overlap).

than 1–2 meter). As a consequence, perhaps surprisingly, results are *not* significantly improved by averaging repeated measurements.

While this error component looks a lot like typical measurement error resulting from additive random noise, it should be noted that: (i) its distribution is *not* Gaussian; (ii) there are distant outliers in many cases; and (iii) the distribution is not always even unimodal. Importantly, small changes in position (or orientation) can cause large changes in the distribution. As a result repeated measurement in fixed positions can lead one to grossly underestimate the error in distance. We'll say that repeated measurements obtained in fixed positions exploit "time diversity," and note that time diversity does *not* provide a path to improved accuracy.

B. "Position-Dependent Error"

Perhaps somewhat surprisingly, small movements (millimeters) of the initiator (or the responder) induce large changes (meters) in reported distance measurements. This error component is a function of 3-D position (and orientation). It is difficult to explore and visualize the error dependence fully in 3-D, but much can be learned by simply scanning along lines.

It is clear that the "position-dependent" error in Fig. 3 is much larger than the "measurement noise" in Fig. 2. Careful

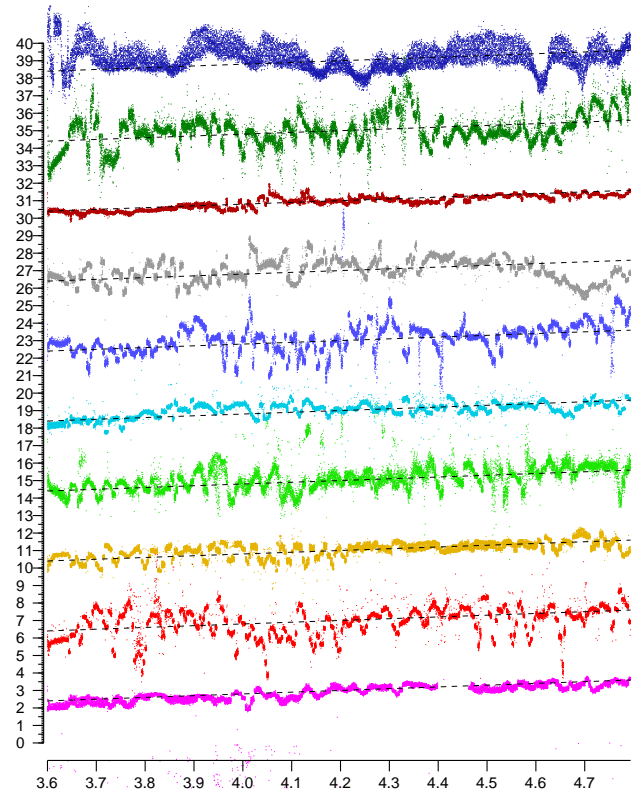


Fig. 3. Sample measurements using ten responders in a range of positions (Top plot is for an access point operating in the 2.4 GHz band, the rest in the 5 GHz band). Horizontal axis: actual position in meter. Vertical axis: reported distances in meters (individual plots are offset vertically to avoid overlap).

measurements along lines surprisingly shows fluctuations in the error surface that have "texture element" size comparable to the wavelength of the radio frequency signal (which ranges from 58 mm for 5210 MHz to 52 mm at 5775 MHz). This is confirmed by inspection of the spatial power spectrum, which has much of its energy at and below the frequency corresponding to about two cycles per wavelength.

Measurements taken at positions separated by more than say a wavelength are fairly uncorrelated. This suggests one way of improving accuracy: average several measurements taken (far enough apart) along points spaced out along a line (or on a regular grid). This indeed leads to a result with considerably higher accuracy than averaging repeated measurements taken in a fixed location. We'll say that repeated measurements obtained on a line (or on a grid of locations) exploit "spatial diversity" and note that spatial diversity *can* improve accuracy significantly.

It is, however, not clear how this observation can be used in practice since it requires either a set of regularly spaced antennas in an array larger than the typical smartphone, or perhaps some mechanism for moving a single antenna into a set of positions in some regular pattern.

For experiments requiring high accuracy, however, such as measurements of the relative permittivities of building materials like concrete, brick and wood, the extra effort in

making measurements in several positions is well justified, since for these types of measurements the raw accuracy of FTM RTT is not adequate.

C. Offset

Over a large range of distances, with a clear line of site, the reported distance varies linearly with the actual distance. The slope of the linear fit is 1 (see e.g. Fig. 4) but there typically is a significant offset, which depends on the type of initiator, the type of responder, the channel in use, bandwidth, and the preamble.

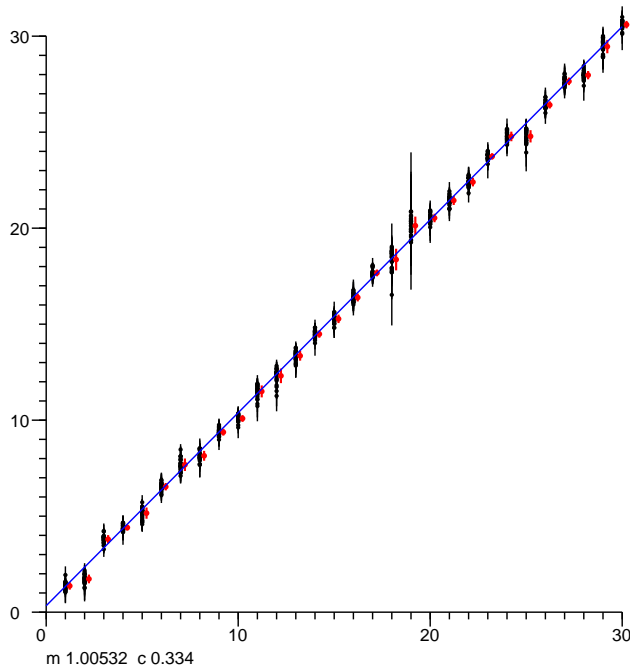


Fig. 4. Linear fit of reported distance to actual distance (outdoors, clear LOS, no obstructions in the first Fresnel zone). The offset in the well calibrated setup tested here happens to be small (less than half a meter), but can be five meters or more in other situations. Horizontal axis: actual position in meter. Vertical axis: FTM RTT reported distances in meters.

Ideally, all initiator/responder combinations would come calibrated to yield zero offset. Presently this is not the case, and different responders will yield different offsets with different initiators (sometimes differing by five or more meter). Even a particular combination of initiator and responder has different offsets when operating in different channels (which can lead to hard-to-track errors when the AP decides to switch channels for some reason!). Presently one must calibrate for the particular combination of initiator and responders to be used in order to eliminate these offsets.

D. Noise Gain

The accuracy of the final location estimate is not the same as the accuracy of the raw measurement of distance between the initiator and the responder. The error in location may be considerably larger than the error in distance measurement, depending on the geometry of the layout of responders and initiator. The ratio of the error in location to the error in the

distance measurement is the “noise gain” — euphemistically referred to as “dilution of precision” (DOP) in GPS terminology. This suggests that there is some benefit to carefully planning the distribution of responders so as to minimize the error in the worst-case position of the initiator (see also Appendix C).

E. Dependence on Bandwidth

The expected accuracy is inversely proportional to the bandwidth of the Wi-Fi signal. Currently the highest bandwidth of initiators and responders that support the IEEE 802.11 FTM RTT protocol is 80 MHz (there are some access points and some Wi-Fi adapters that support 160 MHz but, as of this writing, do not support FTM RTT).

One may consider “bandwidth diversity” as another possible measure to improve accuracy, but the results at 40 MHz and 20 MHz tend to be noticeably worse than those at 80 MHz. As a result, there is only a small gain in accuracy using a best fit weighted average of the three results (aside from that, the offsets are different for different bandwidths and need to be calibrated out).

V. WHERE DOES THE LARGE POSITION-DEPENDENT ERROR COME FROM?

The main component of the error is the position-dependent error. Given the size of the “texture element” of this type of error, it appears to be related to some sort of interference pattern resulting from reflections off objects that are not in the line of sight. This is quite unexpected since the first arriving component of the signal should *not* be affected by any such reflections.

In contrast to this, signal strength (RSSI), being a steady state measurement, *is* subject to large fluctuations (“fast fading”) over relatively small distances due to just such interference. (see upper plot in Fig. 5)

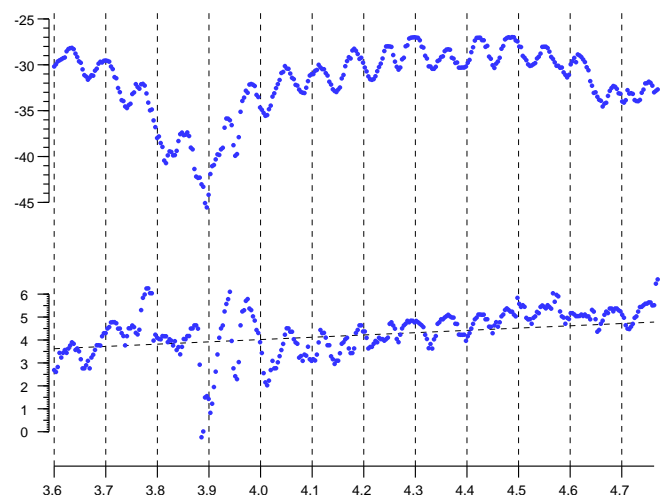


Fig. 5. Upper plot: signal strength (RSSI) in dBm, Lower plot: reported distance in meter. Horizontal axis: actual position in meter. Note undulations with wavelength somewhere between about half the wavelength and the full wavelength of the electromagnetic wave.

Again, at least with a clear line of sight, the first arrival should *not* be affected by later arriving signals reflected from objects in the environment. Thus, it comes as a surprise that FTM RTT distance measurements seem to be affected by some sort of interference patterns or stationary waves (see lower plot in Fig. 5) To understand how this can be, we must know more about how these measurements are made.

A. Super Resolution

With OFDM modulation, demodulation is done by inverse Fourier transform of samples of the signal. For 80 MHz bandwidth, these samples are taken at 80 Msps (actually, both Q (real part) and I (imaginary part) are sampled at that rate, but that does not affect the argument here). That means that samples are taken every 12.5 nsec, which corresponds to 3.75 meter round trip travel of the radio-frequency (RF) wave. So, if first arrival was based merely on which sample exhibits the first sign of a rising waveform, then the (one-way) resolution would be 1.875 m. The measurement actually provided to the user has much finer resolution (RTT, for example, may be given in units of 0.1 nsec, very much smaller than the 12.5 nsec sampling interval). Super-resolution methods are used to “interpolate” between known samples of the signal.

Several super-resolution methods are used, such as MUSIC, ESPRIT, and pencil matrix [9]–[16]. These are based on specific assumptions about the transfer function of the communication channel. In particular, it is assumed that the impulse response of the channel is a weighted sum of shifted impulses, corresponding to different components of a multi-path signal.

While the aim is to provide the user with finer resolution, such methods also have limitations. They are highly non-linear and can exhibit discontinuities and non-monotonicity. Further, information on what actual algorithms are used in the Wi-Fi initiator and in the Wi-Fi access points is not available to the user.

To illustrate the potential problem, consider first an oversimplification. A simple algorithm has arrival time estimated based on when a sample of the signal amplitude exceeds some threshold. However, one cannot use a fixed threshold for deciding when the signal arrives, since the signal can vary over several orders of magnitude (e.g. -100 dBm to -40 dBm — i.e. a ratio of a million to one in power) The threshold to determine whether the “toe” of a signal has arrived must be scaled based on the strength of the signal. But that “signal strength” can only be ascertained later when it has reached a peak. While the “toe” is not affected by multi-path reflections, the amplitude used for normalization *is* subject to the interference pattern. So even though the first arrival is not contaminated by interference, the threshold against which it is compared *is*. This sort of effect can give rise to the wildly fluctuating position-dependent error surface described above (see lower plot in Fig. 5).

VI. FREQUENCY DIVERSITY — SIX CHANNELS

Since the position-dependent error surface has “texture” the order of the wavelength of the radio frequency signal, it

stands to reason that operating at different frequencies would produce different position-dependent errors. There are six non-overlapping 80 MHz channels in the 5 GHz band [7]. This provides for up to six measurements with uncorrelated error contributions, potentially leading to a multiplication of the error by $1/\sqrt{6} \approx 0.408\dots$ (Note that there may be some restrictions on some channels in some parts of the world. The highest channel, for example, is not available in Japan, Israel, Turkey and South Africa [7]).

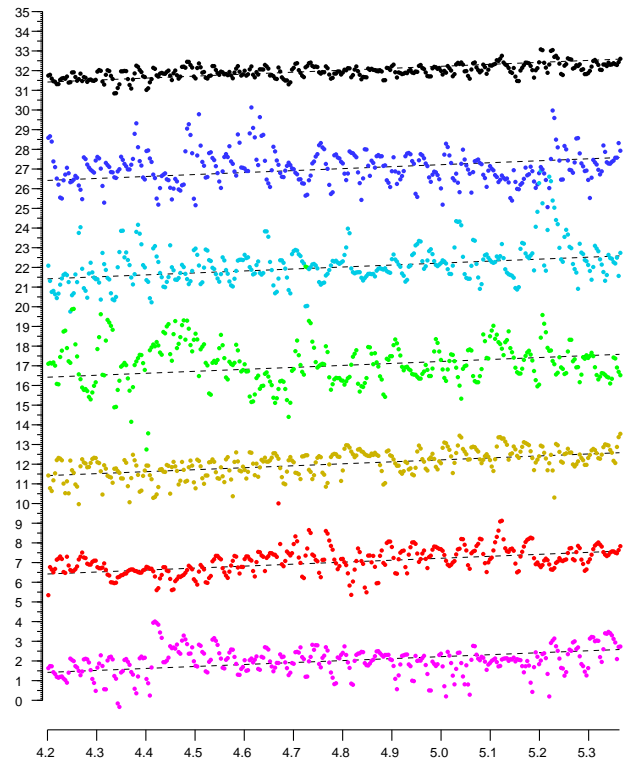


Fig. 6. Sample plots (colored) of reported distances in six 80 MHz wide channels in the 5 GHz band. The top (black) plot is a simple average, which has less than half the error of the individual measurements. Horizontal axis: actual distance in meter. Vertical axis: reported distance in meters (offset to prevent overlap)

Fig. 6 shows plots of distance measurements in six channels as a function of actual position. The channels have center frequencies 5210 MHz (magenta), 5290 MHz (red), 5530 MHz (brown), 5610 MHz (green), 5690 MHz (cyan), and 5775 MHz (blue). The correlation matrix (eq. 2) shows that the position-dependent errors in the different channels are essentially uncorrelated.

$$\begin{bmatrix} 1 & -0.05 & -0.19 & 0.21 & -0.03 & 0.12 \\ -0.05 & 1 & 0.13 & -0.04 & -0.14 & -0.09 \\ -0.19 & 0.13 & 1 & -0.19 & -0.09 & 0.00 \\ 0.21 & -0.04 & -0.19 & 1 & -0.07 & -0.11 \\ -0.03 & -0.14 & -0.09 & -0.07 & 1 & -0.04 \\ 0.12 & -0.09 & 0.00 & -0.11 & -0.04 & 1 \end{bmatrix} \quad (2)$$

The st. dev. of position-dependent errors from the six channels are 0.710 m, 0.578 m, 0.540 m, 1.163 m, 0.944 m, and 0.909 m respectively (average 0.808 m). A simple average of the six distances has st. dev. 0.309 m, which is significantly better than the best channel on its own, and more than twice as accurate as the average.

A weighted sum — rather than a plain average — can do even better. In this particular case, with weights 0.155, 0.234, 0.301, 0.080, 0.130, and 0.100, the st. dev. comes to 0.264 m (which is only about a third of the average st. dev. of the six channels). With six channels, the added refinement of least-squares weighting may not always be worth the effort since the relative quality of the channels depends on the environment and will be different in different situations.

By the way, averaging FTM RTT measurements from six 80 MHz channels does *not* produce the same results as if one were to perform a single FTM RTT measurement in a channel of 480 MHz bandwidth. In the case of a single ultra-wide channel, the error would be multiplied by $1/6$, not $1/\sqrt{6}$.

VII. FREQUENCY DIVERSITY — THREE CHANNELS

It may not always be practical or convenient to use all six 80 MHz channels for FTM RTT distance measurements. In some situations a smaller number may be more easily accessible. Several “tri-band” mesh Wi-Fi APs (e.g. Eero Pro, Netgear Orbi and Linksys Velop). have two radios which make it easy to get measurements for at least two channels in the 5 GHz band (e.g. 5210 MHz in U-NII-1 and 5775 MHz in U-NII-3). Often also, one of the radio chains is shared between the 2.4 GHz and 5 GHz bands and if the device happens to respond to FTM RTT requests in both bands (e.g. Linksys Velop) then this opens up the possibility of taking three measurement with uncorrelated error contributions.

Taking a simple average potentially multiplies the error by $1/\sqrt{3} \approx 0.577 \dots$ (assuming similar distributions for the three channels and with uncorrelated noise). Not as good as with six channels, but still a useful improvement. Actually, this may be a bit optimistic, since the 2.4 GHz channels is not as good as the other two, but suitable weighting of the three contributions can get one close to the ideal.

In Fig. 7, the bottom three plots are for channels with center frequency (i) 5210 MHz (red), (ii) 5775 MHz (green), and (iii) 2442 MHz (blue). The correlation matrix (eq. 3) shows that the position-dependent errors in the different channels are, once again, uncorrelated.

$$\begin{bmatrix} 1 & -0.00 & 0.03 \\ -0.00 & 1 & -0.13 \\ 0.03 & -0.13 & 1 \end{bmatrix} \quad (3)$$

The top plot (black) in Fig. 7 is for a weighted average (weights 0.48, 0.35, and 0.17 respectively). The st. dev. of the position-dependent error in the lower three plots are 0.382 m, 0.480 m, and 0.721 m, for an average st. dev. of 0.528 m. The st. dev. of a simple average is 0.302 m (which is better than any of the individual channel st. dev.). and the st. dev. of the weighted average is 0.270 m (which is almost twice as accurate as the average channel).

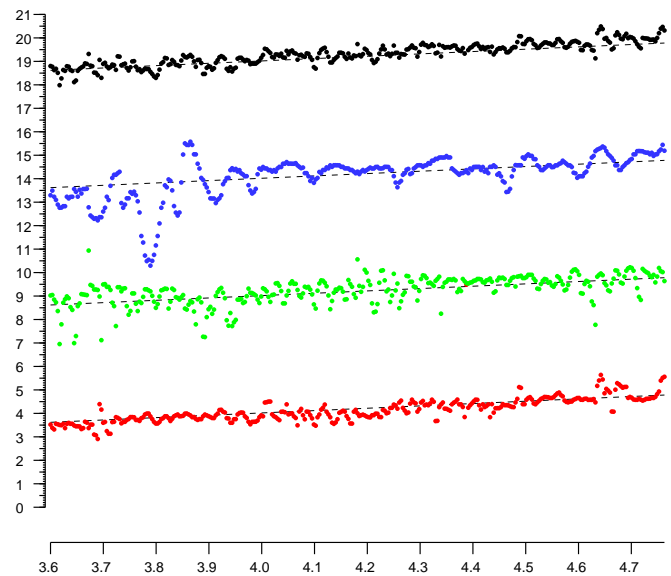


Fig. 7. Sample plots (colored) of reported distances in three channels. The top (black) plot is a weighted average, which has only a bit over half of the average error in the individual plots. Horizontal axis: actual distance in meter. Vertical axis: reported distance in meters (offset to prevent overlap)

Typically different chips are used for the two radio chains. In the case of Eero Pro, for example, the first 5GHz radio (and the 2.4 GHz radio) uses the Qualcomm IPQ4019 chip, while the second 5GHz radio uses the Qualcomm QCA9886 SoC. These have somewhat different measurement qualities and thus weighting their contributions differently (as above) helps improve the overall result.

Finally, if three channels are not available, using two channels can already bring some improvement in accuracy relative to relying on a single channel.

VIII. HIGH RELATIVE PERMITTIVITY OF COMMON BUILDING MATERIALS

Inside buildings, signals often have to travel through walls and floors of concrete, wood, brick, drywall or glass. These materials have high relative permittivity which slows down the signal significantly. Careful measurement of thick layers of various materials show relative permittivities, in the 8–10 range for wood, and 5–7 range for concrete, depending on moisture content and composition (The signal also is attenuated significantly, but this does not directly affect the time-of-arrival) [17]–[20]. Time-of-flight times the speed of light is the equivalent distance travelled in vacuum — which may be considerably larger than the actual distance. A 0.5 meter thick concrete wall can, for example, add 3 or 4 meters to the reported FTM RTT “distance.” This needs to be taken into account somehow in the estimation of position from distance measurements. The effect of thick walls and floors should also be a concern when planning the placements of responders.

Arguably, the effect of high relative permittivities of building materials on distance measurements is more important than that of multi-path. Particularly reminding ourselves again that

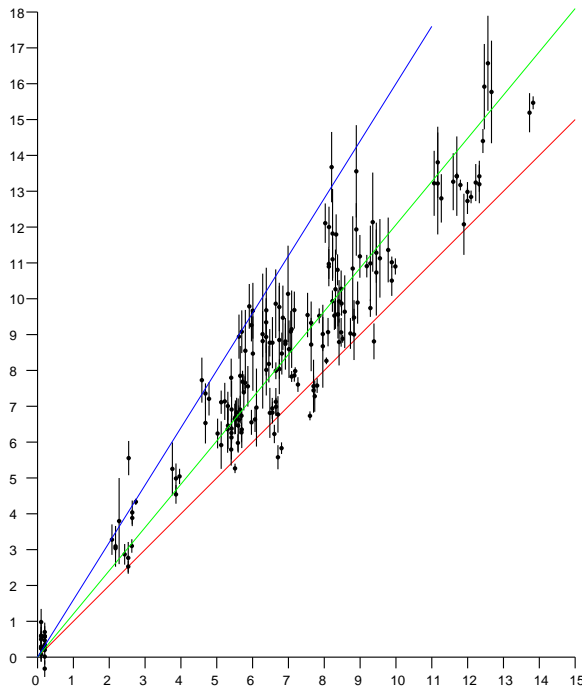


Fig. 8. Scattergram of measured distances versus actual distance in wooden three-story house. Vertical axis: measured distance (meter). Horizontal axis: actual distance (meter). Red line (slope 1) is the ideal relationship; Green line (slope 1.2) is the best linear fit; Blue line (slope 1.6) is an upper extreme. The high permittivity of building materials biases the distances measured by FTM RTT.

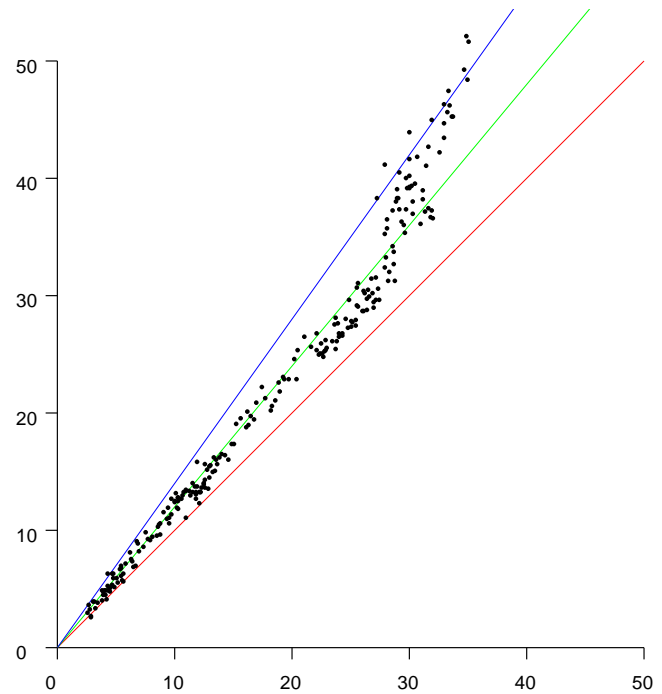


Fig. 9. Scattergram of measured distances versus actual distance in a large open plan office building. Vertical axis: measured distance (meter). Horizontal axis: actual distance (meter). Red line (slope 1) is the ideal relationship. Green line (slope 1.2) is the best linear fit. Blue line (slope 1.4) is an upper extreme. The high permittivity of building materials biases the distances measured by FTM RTT.

the time of first arrival should *not* be affected by reflections that arrive later.

Fig. 8 shows how building materials affect measurements in a three-story wooden house. Fig. 9 show how building materials affect measurements in a large open plan office building. The effect there is less extreme, although over long enough distances just as significant.

IX. RECOVERING LOCATION FROM DISTANCE MEASUREMENTS

Once we have estimated distances from a number of AP responders in known locations we can try and determine where the initiator is.

A. Multi-lateration

If we are dealing with a single level building, we can treat this problem in 2-D. In this case each measurement confines the possible location of the initiator to points on a circle with an AP at the center, — or a circular annulus if we take into account uncertainty in the measurement. Two measurements lead to the intersection of two circles, which typically is two points (These two points lie on a line that is perpendicular to the line connecting the centers of the circles). A third measurement can disambiguate if needed. Three or more measurements are typically inconsistent but can be used in a least squares fashion to reduce the error in location estimation.

This is quite analogous to finding a cellular base station from multiple LTE Timing Advance (TA) measurements — just with much finer resolution [21].

In the more general full 3-D case, each measurement confines the location of the initiator to points on the surface of a sphere with an AP at its center — or a spherical shell if we take into account uncertainty in the measurements. Two measurements restrict the solution to the intersection of two spheres, which typically is a circle (This circle lies in a plane that is perpendicular to the line connecting the centers of the spheres). A third measurement reduces the possibilities to the intersection of a circle and a sphere, which typically occurs in two places. A fourth measurement can disambiguate if needed. Four or more measurements are typically inconsistent but can be used in a least squares fashion to reduce the error in location estimation.

B. Linear multi-lateration?

The equations for the circles — or spheres — are second order, but all with the same higher order terms. Thus it is tempting to subtract them pairwise to obtain linear equations, since sets of linear equations are easy to solve. This is a mistake. While the resulting equations yield the correct solution if the measurements are perfect, the “noise gain” is very high. That is, small errors in distance measurements translate into large errors in position. One way to understand why this happens is that we are throwing away some of the constraint provided by the measurements. For convenience, here we consider the solution to be confined to the planes containing the circles of intersection, *not* to the actual circles, which is a much tighter constraint. (For mathematical details of the argument see [22]).

An aside: this is quite analogous to the infamous “8-point method” in machine vision for solving the relative orientation problem. While it is very appealing because of the linear form of the equations, minimization of errors in those equations does not minimize the sum of errors in image positions [23], [24]. As a result, this method cannot be recommended (other than perhaps in the hope of finding plausible starting values for methods that do the right thing).

C. Least Squares minimization and brute force grid search

For a given hypothesized location for the initiator, the distance from each AP can be computed and compared with the measured distance. One can then find the location that minimizes the sum of squares of the differences between computed and measured distances. Gradient-descent may not work reliably to find the global minimum of this error sum, since the shape of the error surface can be complex. We can, however, divide the space into pixels (2-D) or voxels (3-D) and simply compute the error for each cell. This is, after all, not computationally expensive, since, given the limited accuracy of FTM RTT measurements, the cells need not be very small (e.g. perhaps 0.5 m on a side). So even a typical building with side lengths of tens of meters would be represented by just a few thousand cells.

D. Kalman filtering

Kalman filtering [25], [26] provides a way to update an estimate of the position along with an estimate of the covariance matrix of uncertainty in the estimated position over time a measurement is made. It is based on assumptions of Gaussian noise independent of the measurement, Gaussian transition probabilities and linearity.

Unfortunately the measurement error is not Gaussian nor is it independent of the measurement itself. Further, when near one of the responders, the area of likely positions is shaped more like a kidney (i.e. part of a circular arc) — or even bimodal — rather than something that can be well approximated by a linearly stretched out Gaussian distribution (Fig. 13). As a consequence, Kalman filtering does not provide the best way to use the available information.

E. Particle filter

If a probability distribution is not easily modeled in some parameterized way (such as a multi-dimensional Gaussian), then other means may be used to represent it. One such method is that of particle filters which uses weighted samples to represent a distribution [27]. The distribution is in effect approximated by the sum of weighted impulses. At each step, the position of the particles is updated based on a transition model. The weights of the particles are adjusted based on the measurements. Particles with low weight are then discarded, while new particles are sampled to keep the overall number of particles at a desired value.

F. Bayesian grid update

Another way of dealing with a probability distribution that can't be easily parameterized is to represent it with values on a regular grid. Sequential Bayesian updates can be applied to such a grid of probabilities [28]. This method starts with a prior distribution (perhaps uniform). A transition model is invoked at each step which modifies the distribution based on likely movement of the initiator (e.g. a random walk). If a floor plan is available, impenetrable walls can be taken into account in the transition model if desired. This is followed by Bayesian update based on distance measurements, which uses an observation model which estimates the probability of seeing a measurement given the actual geometric distance between a voxel and the responder.

If a single location is required as output, rather than a distribution, one can, for example, use the mode (maximum likelihood) or the centroid (expected value) of the distribution.

As with other forms of “filtering,” there can be a lag in the response when the initiator moves more rapidly than the transition model expects. Also, a bad solution may get “trapped” behind walls, when a floor plan is used to prevent “tunneling” through walls in the transition model.

G. Observation Model

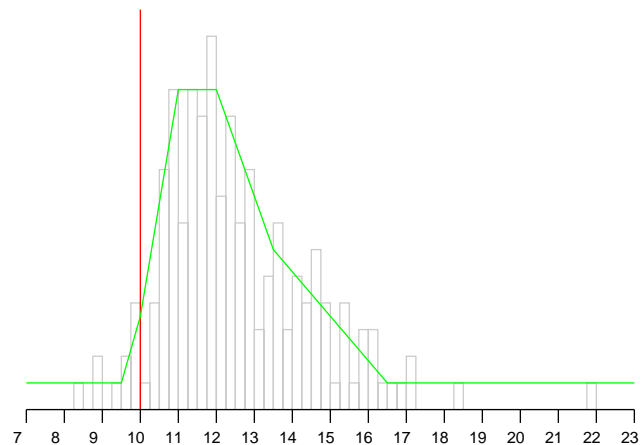


Fig. 10. A slice through an observation model. Horizontal axis: measured distance (when actual distance is 10 m) Grey histogram: measurements from typical three-level residence. Green curve: observation model — probability of measuring the specified distance (piece-wise linear fit to grey histogram).

Fig. 10 shows a section of an observation model. It shows the probability of various measured “distances” on the horizontal axis in meter) given that the actual distance between initiator and responder is 10 meter (i.e. the vertical red line). The actual distance is a lower bound on the measurement. It can be considerably larger since the signal may pass through building materials with large relative permittivity. In the figure, the observation model (green curve) is a piece-wise linear fit to experimental data from a three-level residence (grey histogram).

The observation model is used to update the probability at each grid cell. For each cell on the grid, the distance from the AP is known and so the appropriate slice of the observation

model can be accessed. The observed FTM RTT distance is then used to look up the probability that this observation would occur, given the known actual distance for this grid point. This value is then used to multiply the current value in that cell. Optionally, the resulting grid of values can then be normalized so it once again adds up to one.

H. Transition Model

We use a simple transition model of a random walk of a step size based on comfortable walking speed of 1.4 m/sec [29]). In a simple implementation this just “pushes” probabilities into neighboring cells (except for cells on the edge of the grid). If more information is available from inertial measurement (IMU) and magnetic compass, then this can be used to refine the transition model. But the simple model appears to be adequate for location determination. A floor plan can be used to limit “forbidden transitions” such as walking through a wall. This can further improve the tracking of a location solution as the user progresses through the environment.

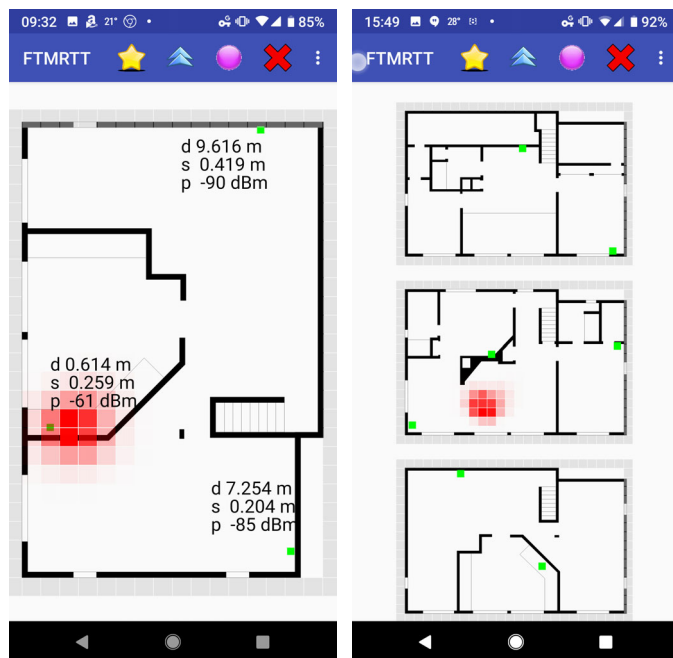


Fig. 11. Sample “heat maps” of Bayesian grids. Left: 2-D case (single level) with 3 responders (green dots). Text shows current FTM RTT distance, st. dev. and signal strength. Right: 3-D case (three levels) with 7 responders. Voxels in each floor were collapsed into a single layer for display purposes.

Fig. 11 shows probability distributions on grids with cells 0.5 meter on a side. The green dots mark the positions of the responders (In this case, the floor plan was not utilized to limit the transition model). For an MP4 movie showing the Bayesian grid evolve as someone moves on one level, see [30]. For an MP4 movie showing the Bayesian grid evolve as someone moves through a three-story building, see [31].

X. NOISE GAIN (A.K.A. DILUTION OF PRECISION — DOP)

The geometric arrangement of responders determines the “dilution of precision” (DOP, or “noise gain”), that one can expect in various parts of the volume of interest.

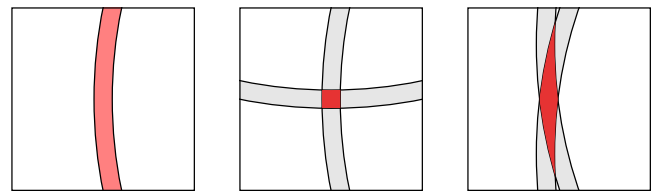


Fig. 12. Dilution of Precision. Left: constraint from single distance measurement; Middle: favorable combination of constraints; Right: unfavorable combination of constraints. The area of the overlap grows as $1/\sin(\theta)$, where θ is the angle between the directions to the APs.

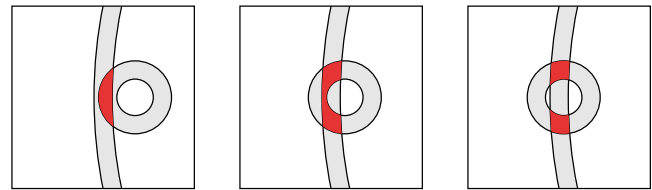


Fig. 13. Dilution of Precision when close to responder. Left: Intersection is more or less an oblong oval; Middle: Intersection is sort of kidney shaped; Right: Intersection is bimodal. Such distributions cannot reasonably be approximated by multi-variate Gaussians.

On the left in Fig. 12, is shown the annulus within which the initiator position is constrained when a single, noisy distance measurement is available. In the middle is the situation when two measurements are available from responders that are more or less at right angles in directions as seen from the initiator. Plausible solutions in this favorable case are confined to a small area. On the right is the less fortunate situation where the directions to the responders are similar, and not much new information is provided by the second measurement. Correspondingly, the likely position of the initiator is not as well confined.

When close to one of the responders, the geometry becomes more intricate, and, counter-intuitively, the solution may be less well determined. This is illustrated in Fig. 13.

It is generally not a good idea to have the responders close together, since then the distance measurements will be correlated and redundant. The effect of errors typically not isotropic, but is stronger in some directions than others (as, for example, in the case of GPS, where the vertical DOP is considerably larger than the horizontal DOP, as a result of the fact that the “visible” satellites are not distributed evenly over a sphere of possible directions). In some cases curves of constant error may be quite elongated, meaning that while the position may be well defined in some directions, it is not in others. Finding the “best” layout of responders in a given 3-D volume is an open research problem.

For additional detail see Appendix D.

XI. CONCLUSIONS

The accuracy of FTM RTT distance determination can be doubled using frequency diversity. The error in FTM RTT distance has peculiar properties (for a start, it is non-Gaussian) that derive from the super-resolution algorithms used. Common building materials can introduce large errors

in FTM RTT distance estimates because of their high relative permittivity. Bayesian grid estimation is well suited to the task of recovering location from distance measurements given the unusual nature of the errors. The “noise gain” in location determination can be kept low by carefully planning the geometric arrangement of access points.

ACKNOWLEDGMENT

We wish to express our gratitude for the helpful suggestions of Irad Stavi at Compulab Ltd., Mark Schwartz of Kevadiya, Inc., and Roy Want and Wei Wang of Google, Inc.

REFERENCES

- [1] (2017, Feb.) Wi-Fi CERTIFIED location brings Wi-Fi indoor positioning capabilities. Wi-Fi Alliance. [Online]. Available: <https://wi-fi.org/news-events/newsroom/wi-fi-certified-location-brings-wi-fi-indoor-positioning-capabilities>
- [2] F. V. Diggelen, R. Want, and W. Wang. (2018, Jul.) How to achieve 1-meter accuracy in Android. [Online]. Available: <https://www.gpsworld.com/how-to-achieve-1-meter-accuracy-in-android/>
- [3] J. Cobb. (2018, Sep.) Testing Wi-Fi RTT on Android P for indoor positioning. [Online]. Available: <https://www.crowdconnected.com/blog/testing-wifi-rtt-on-android-p-for-indoor-positioning/>
- [4] M. Ibrahim, H. Liu, M. Jawahar, V. Nguyen, M. Gruteser, R. Howard, B. Yu, and F. Bai, “Verification: Accuracy evaluation of WiFi fine time measurements on an open platform,” in *MobiCom-18*, New Delhi, India, Oct. 2018.
- [5] (2019, Sep.) Indoor positioning with WiFi RTT and Google WiFi. [Online]. Available: <http://www.darryncampbell.co.uk/2019/09/27/indoor-positioning-with-wifi-rtt-and-google-wifi/>
- [6] *IEEE 802.11 - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE Std., 2016. [Online]. Available: https://standards.ieee.org/standard/802_11-2016.html
- [7] List of WLAN channels. [Online]. Available: https://en.wikipedia.org/wiki/List_of_WLAN_channels
- [8] J. Xiong, “Pushing the limits of indoor localization in today’s Wi-Fi networks,” Ph.D. dissertation, University College of London, Sep. 2015.
- [9] X. Li and K. Pahlavan, “Super-resolution toa estimation with diversity for indoor geolocation,” *IEEE Transactions on Wireless Communications*, vol. 3, no. 1, pp. 224–234, Jun. 2004.
- [10] A. A. Ali and A. S. Omar, “Time of arrival estimation for WLAN indoor positioning systems using matrix pencil super resolution algorithms,” in *Proc. 2nd Workshop on Positioning, Navigation and Communications (WPNC)*, 2005.
- [11] F. X. Ge, D. Shen, Y. Peng, and V. O. K. Li, “Super-resolution time delay estimation in multipath environments,” *IEEE Transactions on Circuits and Systems I*, vol. 54, no. 9, pp. 1977–1986, Sep. 2007.
- [12] D. Humphrey and M. Hedley, “Prior models for indoor super-resolution time of arrival estimation,” in *Proc. IEEE 69th Conference: Vehicular Technology Conference (VTC)*, May 2009.
- [13] C. Morhart and E. M. Biebl, “High resolution time of arrival estimation for a cooperative sensor system,” *Advances in Radio Science*, vol. 8, pp. 61–66, 2010.
- [14] T. J. S. Khanzada, A. R. Ali, and S. A. Napoleon, “Use of super resolution algorithms for indoor positioning keeping novel designed wlan signal structure,” *ACM IWDE*, 2010.
- [15] M. Noto, F. Shang, S. Kidara, and T. Kirimiti, “Super-resolution time of arrival estimation using random resampling in compressed sensing,” *IEICE Trans. Commun.*, vol. E101-B, no. 6, Jun. 2018.
- [16] J. A. Bhatti, B. M. Ledvina, R. W. Brumley, W. Zhang, N. Wong, and E. Chiu, “Time of arrival estimation,” U.S. Patent 10 361 887, Jul., 2019.
- [17] R. Wilson. (2002, Aug.) Propagation losses through common building materials - 2.4 ghz vs 5 ghz. [Online]. Available: <http://www.am1.us/Papers/E10589%20Propagation%20Losses%202%20and%205GHz.pdf>
- [18] B. Davis, C. Grosvenor, R. Johnk, D. Novotny, J. Baker-Jarvis, and M. Janezic, “Complex permittivity of planar building materials measured with an ultra-wideband free-field antenna measurement system,” *Journal of Research of the National Institute of Standards and Technology*, vol. 112, no. 1, Jan./Feb. 2007.
- [19] B. Kapilevich, M. Einat, A. Yahalom, M. Kanter, B. Litvak, and A. Gover, “Millimeter waves sensing behind walls - feasibility study with FEL,” in *Proceedings of FEL 2007*, Novosibirsk, Russia, 2007.
- [20] R. Rudd, K. Craid, M. Ganley, and R. Hartless. (2014, Sep.) Building materials and propagation final report - fcom. [Online]. Available: https://www.ofcom.org.uk/_data/assets/pdf_file/0034/55879/ihp_final_report.pdf
- [21] B. K. P. Horn. (2017) Timing advance and Android smartphones. [Online]. Available: https://people.csail.mit.edu/bkph/cellular_repeater_TA
- [22] —. (2018) Reduction to linear equations and the pseudo inverse. [Online]. Available: https://people.csail.mit.edu/bkph/ftmrtt_location#Linear_Equations
- [23] —. (1999) What is wrong with so-called ‘linear’ photogrammetric methods? [Online]. Available: https://people.csail.mit.edu/bkph/articles/Orientation_2D_Illustration.pdf
- [24] —. (1999) Projective geometry considered harmful. [Online]. Available: <https://people.csail.mit.edu/bkph/articles/Harmful.pdf>
- [25] Kalman filter. [Online]. Available: https://en.wikipedia.org/wiki/Kalman_filter
- [26] Thorvald N. Thiele. [Online]. Available: https://en.wikipedia.org/wiki/Thorvald_N._Thiele
- [27] Particle filter. [Online]. Available: https://en.wikipedia.org/wiki/Particle_filter
- [28] Bayesian inference. [Online]. Available: https://en.wikipedia.org/wiki/Bayesian_inference
- [29] Preferred walking speed. [Online]. Available: https://en.wikipedia.org/wiki/Preferred_walking_speed
- [30] B. K. P. Horn. (2019) FTM RTT locations while moving around on one level. MP4. [Online]. Available: https://people.csail.mit.edu/bkph/images/FTM_RTT_2D-2019-01-03-170903.mp4
- [31] —. (2019) FTM RTT locations while moving through a three-story home. MP4. [Online]. Available: https://people.csail.mit.edu/bkph/movies/FTM_RTT_Bayesian-2018-12-20-164419.mp4
- [32] Java reflection. [Online]. Available: <https://docs.oracle.com/javase/tutorial/reflect/>

APPENDIX A

CURRENT STATE OF SUPPORT FOR FTM RTT

The ability to perform FTM RTT measurements is announced by an access point in the beacon it emits at regular intervals (typically every 100 msec). Presently only Compulab’s “Wi-Fi Indoor Location Device” (WILD) (with a modified Intel AC-8260 Wi-Fi adapter) and Google Wifi (with Qualcomm IPQ4019) do this. Importantly, however, quite a number access points do *respond* to FTM RTT requests even though they do not announce this capability. This includes several of the recent “mesh” APs (e.g. Eero Pro, Netgear Orbi, Linksys Velop) — as well as some older APs such as ASUS RT-ACRH13.

APs that support FTM RTT, but do not advertise this capability, may, in some cases, not support it properly yet, be subject to large offsets and measurement errors, frequent outliers, or crashes when asked to respond “too often.”

Many Wi-Fi adapters cannot be used as access points because of regulatory restrictions on their channels. Channels may be marked “passive scan only” or “no IR” (i.e. cannot “initiate radiation”). Generic Intel 8260, Intel 8265, Intel 9260 Wi-Fi cards do “support” FTM RTT, but are not allowed to act as access points (due to “no IR” restriction on channels in the 5GHz band) and so are not useful as FTM RTT responders.

Wi-Fi access points tend to be replaced less often than say smartphones and laptops. APs tend to be replaced only when some major new feature is touted (such as higher data rates and more channels in 802.11AC). Unfortunately FTM RTT was introduced at a time when no such major advantage was simultaneously being proffered. As a result relatively few installed APs support FTM RTT at this point. This will

change as soon as powerful new features are introduced, as in 802.11AD for example.

APPENDIX B ANDROID API AND JAVA REFLECTION

Access points that support the IEEE 802.11mc FTM RTT protocol, but do not advertise this capability, are awkward to use because Android API `WifiManager.getScanResults()` marks them as not supporting 802.11mc in the `ScanResult`, and so the `WifiRttManager.startRanging()` call on the corresponding `RangingRequest` fails — without even trying.

One work-around is to use Java reflection [32], to set the `FLAG_80211mc_RESPONDER` bit in the `flag` field in the `ScanResult` (the “setter” methods `setFlag()` and `clearFlag()` are blacklisted and so can’t be used by third-party applications).

A more flexible approach is to use the hidden `addResponder` method in the `Builder` inner class of the `RangingRequest` class. For this one needs to build an instance of the `ResponderConfig` class “by hand.” A `ResponderConfig` instance contains the MAC address (BSSID), responder type (AP), 80211mc support flag, channel width, frequency, center frequency, and preamble type. One advantage of this approach is that one can build a `RangingRequest` without needing the results of a Wi-Fi scan (which takes time, and is heavily throttled in Android 9) — The information about the APs may come “out of channel” — from a file say (which may also contain information about the physical location of the APs — see also Appendix C). Needless to say, this requires more Java reflection magic.

In this regard, it may be helpful to know that the specified center frequency field in the `ResponderConfig` is *ignored* and replaced by a stored value from the most recent Wi-Fi scan. There are a number of implications, aside from the obvious one that one cannot control the center frequency of the AP in this fashion. One is that an AP can’t be used for ranging if it hasn’t been “seen” recently in a Wi-Fi scan. Another is that an AP can’t be used right after it switches channels — at least not until the next Wi-Fi scan picks up the new channel information. By the way, it is important to know which channels APs use, since the offset in the FTM RTT reported distance is typically different in different channels.

As an aside, Windows 10 does not currently support 802.11 FTM RTT (while it does support Wi-Fi scans using `Wlan-Scan()`).

APPENDIX C HOW TO GET THE LOCATIONS OF THE RESPONDERS

In recovering the location of the initiator (smartphone), one needs to know the locations of the responders (APs). This information can be provided “out of channel” in a file that lists all of the APs in a building — along with their properties.

It may be more convenient (and the method scales better) if the APs themselves broadcast this information. The IEEE 802.11-2016 standard provides for that. Location Configuration Information (LCI) can provide latitude, longitude, altitude and their uncertainties. Location Civic Report (LCR

or CIVIC) can provide a “civic” address in a standardized key-value format. Corresponding “getter” methods `getLci()` and `getLcr()` of `RangeResult` are blacklisted in Android and so not available to third-party applications. However, the `getUnverifiedResponderLocation()` method is available to obtain a `ResponderLocation` from a `RangeResult` and this has the available location information.

Presently, only CompuLab’s “Wi-Fi Indoor Location Device” (WILD) provides for specification of LCR and CIVIC information about the access point (using entries `-lci=...` and `-civic=...` in the `hostapd.config` file). Sadly, at this point, no other Wi-Fi access point allows specification of the location of the access point.

APPENDIX D PLACEMENT OF RESPONDERS

In the 2-D examples in Fig. 14 the green dots are the locations of responders (APs), while the red dots are potential positions for the initiator (smartphone, STA). The constant error curves show how position may be poorly localized in some direction yet well constrained in a direction at right angles. In placing the responders, the aim is to make the constant error curves small and round in most of the work space. Symmetrical layouts for the responders seem to work well, as shown on the left in Fig. 14, while somewhat surprising results may be achieved with asymmetrical layouts, as shown on the right in Fig. 14. Note also that location can be recovered reasonably well even outside the convex hull of the responders — up to a point.

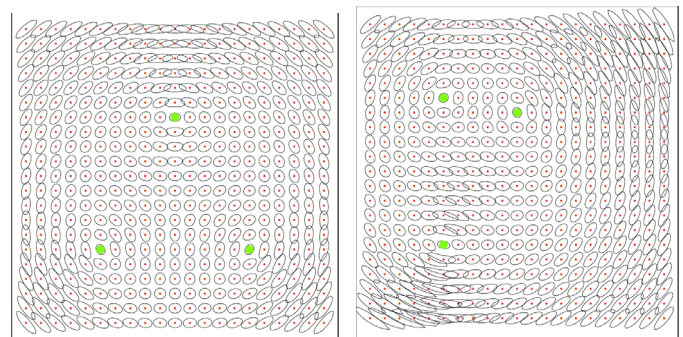


Fig. 14. Quality of location determination near three responders (2-D case). Left: Symmetric arrangement. Right: Asymmetric arrangement.

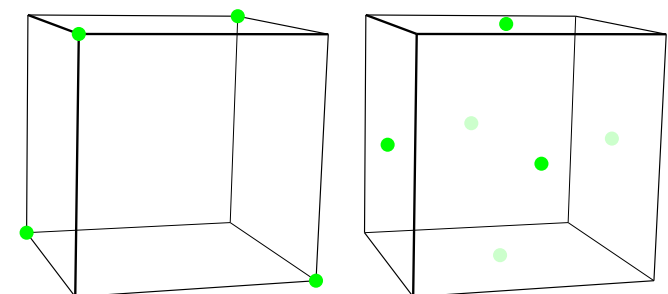


Fig. 15. Left: Placement of 4 responders (3-D case) Right: Placement of 6 responders (3-D case)

For 3-D, cubic volume of interest (or a rectangular brick shaped volume with not-too-different side lengths), placing four responders at the vertices of a tetrahedron embedded in the cube has appealing properties (these points are at the four “even” vertices of the cube (left side of Fig. 15). With six responders, the vertices of an octahedron have good properties (these six points are at the face centers of the cube (see right side of Fig. 15). Both of these configurations avoid placing any subset of (more than three) responders in a plane.

Adding a responder somewhere in the middle of the volume also improves overall location accuracy determination quality (See e.g. right side of Fig. 11).

Placing responders at regular intervals along a line (a corridor say), while providing simplicity of installation, may not be a good idea if location accuracy is of importance.