Indoor Localization Accuracy of Major Smartphone Location Apps

Nader Moayeri, Chang Li, and Lu Shi
National Institute of Standards and Technology
Gaithersburg, MD, USA
Email: {nader.moayeri, chang.li, lu.shi}@nist.gov

***Abstract*—NIST carried out comprehensive testing in five large buildings to compare the accuracies of Fused Location Provider (FLP) Application Programming Interface (API) for Android phones and Core Location running on iPhones. The testing was done under various mobility modes using the performance metrics provided in the international standard ISO/IEC 18305. This paper presents the results of these evaluations for horizontal, vertical, and 3D errors. Even though in some cases and respects Core Location is better than FLP, it can be said that overall FLP has better accuracy than Core Location, at least in the buildings we used for testing. While Core Location consistently provides location estimates at a fast rate, FLP provides location estimates at a slower rate and sometimes it does not provide elevation information. The paper also compares the accuracy of FLP with that of the best Android app that won the PerfLoc Prize Competition organized by NIST for development of Android indoor localization apps.**

***Index Terms*—indoor localization, smartphone apps, location apps, accuracy, Android, Fused Location Provider (FLP), iOS, Core Location, E911, PerfLoc Prize Competition, ISO/IEC 18305**

I. Introduction

The Global Navigation Satellite System (GNSS) is the enabler for the widely used outdoor navigation capabilities of the smartphone, such as vehicular navigation or looking for a store or establishment while on foot in unfamiliar surroundings in a city. Less is known about smartphone's indoor localization capability. It is not as accurate as its outdoor counterpart, but it provides some idea of where one is in a large building. This may prove useful, for example, when trying to reach a store in a shopping mall or a work of art in a museum. While outdoor localization accuracy benefits from the availability of road and street maps, lack of access to floor plans in most buildings limits indoor localization accuracy. In addition, lack of line-of-sight (LOS) signal propagation paths to GNSS satellites and presence of severe multipath signal propagation inside buildings make the indoor localization problem much harder. The goal of this paper is to assess the indoor localization capability of the smartphone.

The vast majority of smartphones in use today are either Android phones or iPhones, with almost 85% and 15% market shares [1], respectively. While Google has developed the Fused Location Provider (FLP) Application Programming Interface (API) [2] for indoor/outdoor localization in Android phones, iPhones use Apple's Core Location framework [3]. Google Maps uses FLP and Apple Maps uses Core Location. Understandably, Google and Apple do not reveal how exactly FLP and Core Location work, but it is known that they both collect and use anonymous crowdsourced data, including cellular signals, GNSS signals, Wi-Fi signals, and air pressure measured by the smartphone, when the location service on the phone is enabled. It is also known that Core Location uses magnetometer measurements and Bluetooth signals when the latter is available. Apple has an indoor localization solution that relies on Wi-Fi fingerprinting at the site prior to use that produces much more accurate location estimates indoors than the "baseline" Core Location evaluated in this paper. We do not evaluate Apple's fingerprinting solution in this paper. We wish to determine if either FLP or (baseline) Core Location is clearly better than the other. We are not aware of any paper in the technical literature that has addressed this question, the answer to which should be of interest to many.

The U.S. National Institute of Standards and Technology (NIST) is well positioned to test FLP and Core Location in a comprehensive manner. We possess a unique test bed, the largest of its kind in the world, for such performance evaluations. We led the development of the international standard ISO/IEC 18305, Test and evaluation of localization and tracking systems [4], which we use in this study. Last but not the least, NIST organized the PerfLoc Prize Competition [5], [6] for development of Android smartphone indoor localization apps. This paper not only evaluates FLP and Core Location performance, but it also compares FLP with the app that won the top prize in PerfLoc.

According to the U.S. Federal Communications Commission (FCC) [7], more than 10,000 people, who would otherwise be saved, die every year when calling 911 from a cellphone because emergency dispatchers cannot get a quick and accurate location on them. An emerging application of FLP and Core Location is to provide emergency responders with more accurate location for 911 calls placed from cell phones inside buildings than with legacy 911 through the use of cellular telephony signals. Google is already using FLP in 14 countries around the world to provide "supplemental location" information to emergency call centers [8]. Recently, Google launched Android Emergency Location Service (ELS) in the U.S. with T-Mobile and RapidSOS [8]. The initial test results show that ELS data "dramatically shrunk the estimated radius of a call's location, from 522 feet down to 121 feet

U.S. Government work not protected by U.S. copyright

and arrived faster than carrier data" [9]. Apple has launched its own initiative so that "iPhone users in the United States who call 911 will be able to automatically and securely share their location data with first responders" with the introduction of iOS 12 [10], [11]. The results presented in this paper provide timely information, based on comprehensive testing, on how FLP and Core Location help with E911 indoor location accuracy. The latest rules from the FCC require 50-meter horizontal location accuracy or providing dispatchable location for 80% of E911 calls by April 2021 [12]. Both FLP and Core Location meet or exceed this benchmark *today*. The caveat is that the building has to have Wi-Fi for these gains to materialize.

The rest of the paper is organized as follows. Section II describes the methodology we used for testing Android FLP and iOS Core Location. Section III presents comprehensive test results for FLP and Core Location. Comparisons between FLP and the best PerfLoc app are presented in Section IV. Finally, concluding remarks are provided in Section V.

II. Testing Methodology

The methodology we used to evaluate FLP and Core Location had much in common with what we used in the PerfLoc Prize Competition. The performance evaluation procedure in PerfLoc was a two-phase process. In Phase I, we used an over-the-web procedure to evaluate the PerfLoc "algorithms" using 30 test data sets, each collected with four different brands/models of Android phones in four NIST buildings. In Phase II, we invited the teams that had developed the best algorithms to NIST and evaluated their "apps" using live testing in a fifth building at NIST they knew nothing about. We made training data sets collected in the first four buildings available to PerfLoc participants in Phase I and the participants could use the entire data in a set to estimate location at a designated time instance. That is, they could use "future" smartphone data to estimate location at the "present time". In Phase II, the finalists invited for live testing at NIST were required to provide location estimates in "real-time"; they could not use any lookahead. In addition, they did not have any training data sets for the fifth building [5], [6]. We used live testing only to evaluate the performance of FLP and Core Location, but we tested in all five buildings that had been used to evaluate the performance of PerfLoc algorithms and apps.

A. Buildings and Test Points

The five buildings we used for testing FLP and Core Location were selected based on guidance from ISO/IEC 18305. These buildings are instrumented with almost 1300 test points, henceforth called dots, laid on the floors and professionally surveyed. Therefore, the ground truth 3D coordinates (latitude, longitude, and elevation) of all dots are known to NIST. Table I presents descriptions of the buildings, their sizes, the number of dots deployed in each, and the number of Wi-Fi access points (APs) in each.

Building 1 is subterranean with two floors under grade level, i.e. it has a sub-basement in addition to the basement indicated in Table I. It had no Wi-Fi APs when we collected PerfLoc data [5] in February 2016. As shown in Table I, it now has four APs, but there is no Wi-Fi signal available in roughly 85% of its floor area. No cellular signal can be received in that building, and naturally no GNSS signal is available either. Neither FLP nor Core Location worked in that building. Therefore, limited testing was done in that building after we observed that the location estimate provided by FLP or Core Location hardly ever changed once we entered the building. We recognize that this is a corner case representing a very challenging building from a radio frequency (RF) coverage viewpoint, but this situation can arise in a multi-level underground parking structure without Wi-Fi coverage.

TABLE II: T&E Scenarios and Numbers of Dots in Each Building/Scenario Combination

Scenario #	Description	# of Dots Used in Each Building				
		B1	B2	B3	B4	B5
1	Walking with 3-sec. Stops at Dots	123	157	109	41	96
2	Walking w/o any Stops	126	146	147	52	139
3	Cart	—	41	127	48	58
9	Using Elevators	—	—	—	—	67
10	Walking in/out of the Building	—	144	126	88	97
12	Sidestepping	—	123	153	50	34
13	Walking Backward	—	115	144	49	32
14	Crawling	—	145	151	—	10
16	Cart	—	87	—	—	—

B. Test and Evaluation (T&E) Scenarios

A T&E scenario for a given building is a pre-determined path with known starting and ending points in the building that the test subject follows while carrying or wearing the localization device of the system under test [4]. In this study, we tested two localization devices simultaneously, namely, an Android phone and an iPhone. Several T&E scenarios were used in each building corresponding to different mobility modes. The test subject would prompt each of the two phones to generate a location estimate at each dot visited on the path for a given scenario. In general, our testing procedures closely follow the guidelines provided by the ISO/IEC 18305 standard. Table II shows various scenarios used and the number of dots on the path for each scenario in each building. Most of these scenarios involved moving around in a building for at least 20 minutes.

In Buildings 1-4, we used the same T&E scenarios as those used in PerfLoc Phase I evaluations, except that we stopped testing in Building 1 after two scenarios, because FLP and Core Location were hardly updating their location estimates. In Building 5, we used the same T&E scenarios as those used in PerfLoc Phase II evaluations. We used the same paths in the five buildings for this study as those used in PerfLoc. To be exact, each path followed the same sequence of dots as the corresponding scenario in PerfLoc, but naturally how we got from one dot on a path to the next one was not exactly the same to centimeter-level accuracy as the corresponding path

TABLE I: Buildings Used in Our Tests

Building #	Building Type	Gross Floor Area (m^2)	Basement	# of Dots	# of Wi-Fi APs
1	Multi-Story Laboratory & HVAC Equipment	9,308	Y	217	4
2	Multi-Story Office and Laboratory	10,709	N	299	46
3	Single-Story Warehouse, Shop, and Office	13,342	N	257	26
4	Single-Story Machine Shop	4,598	N	100	8
5	Multi-Story Office	31,870	Y	403	131

in PerfLoc. More importantly, the RF landscape can never be replicated, because the performance evaluations for PerfLoc (data collection for Phase I and live testing in Phase II) and FLP/Core Location were done at different times. Subject to these caveats, we can claim that we used the same procedures for evaluating the performance of FLP and Core Location as those used for PerfLoc. Therefore, it would be reasonable to compare the performance of FLP with that of the best PerfLoc app in Building 5, which we do in Section IV.

Any scenario involving a mobility mode other than walking (sidestepping, walking backward, crawling, and the cart) in Buildings 1-4 had a good bit of walking and then one or more segments of the other mobility mode specified. The cart scenarios are exceptions, because they used the cart from the start to the end. In Building 5, we opted for obtaining performance results for other mobility modes without mixing them with walking, and that is how the live tests in Phase II of PerfLoc were done.

The starting points for most building/scenario combinations were outside the building, where we spent 2-3 minutes to make sure we had a good GPS fix before we entered the building. In two cases only, the starting point was inside the building.

We evaluated FLP and Core Location simultaneously. Figure 1 shows how (a) we attached the phones to a "plastic stick" and the test subject walked while holding the stick in his/her hand, (b) the test subject crawled on the floor with the stick in his/her hand, and (c) we placed the phones on the bed of a plastic cart while the test subject pushed the cart around in the building. ISO/IEC 18305 includes crawling, sidestepping, and walking backward scenarios because fire fighters engage in those activities while moving around in a building on fire. The cart scenario would arise for asset tracking in a warehouse, in a hospital, or on the factory floor, when assets are transported on a cart or other type of wheeled transport vehicle. We took into account the average elevation of the phones with respect to the floor in walking, crawling, and cart scenarios so that our computation of localization error would be accurate. These elevations were 1.53 m, 0.13 m, and 0.77 m, respectively, for the Android phone. They were 1.74 m, 0.18 m, and 0.77 m, respectively, for the iPhone.

C. Phones Used and the Evaluation Apps

In this study, we used a Google Pixel XL phone running on Android OS 9 / Google Play Services 14.3.67 and an iPhone 8 running on iOS 11.4.1. One may wonder whether FLP and Core Location indoor performance would be affected by these specific choices of phones. We suspect the answer to this question is negative, because FLP and Core Location primarily rely on Wi-Fi signals, and the number of smartphone Wi-Fi chipset manufacturers is very limited.

We used an Evaluation App for each phone to assess the performance of FLP or Core Location. Specifically, the test subject would click on a button on each Evaluation App when the two phones were above a designated dot on the path for a given T&E scenario as the test subject followed the path for the scenario. The Evaluation App would then call FLP or Core Location to obtain a location estimate. We developed the Evaluation App for the Android phone ourselves. The Evaluation App for Core Location was provided by Carnegie Mellon University. Core Location was always able to respond with a location estimate, as it generates such estimates at a fixed 60 Hz rate. This was not the case with FLP, however. At best, it can generate location estimates at a 1 Hz rate. Updated location estimates from FLP were not available at all dots on the path for a T&E scenario. Therefore, we used the last available location estimate for FLP, whenever an update was not available.

D. Additional Testing Beyond PerfLoc Scenarios

Each dot in each building was visited several times in the course of our test and evaluations. We noticed significant variation in location estimates obtained at the same dot, regardless of whether the location estimates were obtained from FLP or Core Location. Therefore, we spent another ten days of data collection, in addition to the initial ten days for the PerfLoc T&E scenarios described above, to collect a reasonably large number of estimates with each phone at each dot in each building. Location estimates from Core Location were always available, but this was not true for FLP. In FLP's case, sometimes no updated location estimates were available. There were also cases where FLP would produce a latitude and a longitude, but no elevation.

The analysis of the additional data we collected is simply beyond the scope of this paper due to space limitations.

III. Test Results

We already mentioned that neither FLP nor Core Location worked in Building 1. When we tried to go through a walking scenario in that building, both FLP and Core Location provided location estimates at the starting point of the scenario outside the building. Core Location provided frequent location updates all the time we walked inside the building, but all of them were the same as the last location estimate it provided outside the building. FLP did not provide any location updates

(a) Walking (b) Crawling (c) Cart

Fig. 1: Various Mobility Modes

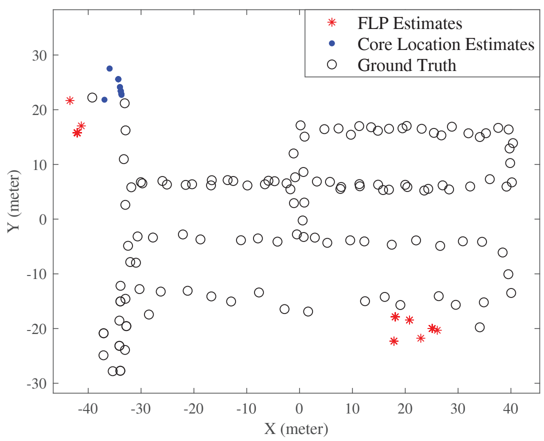

Fig. 2: Neither FLP nor Core Location provided accurate location estimates in Building 1.

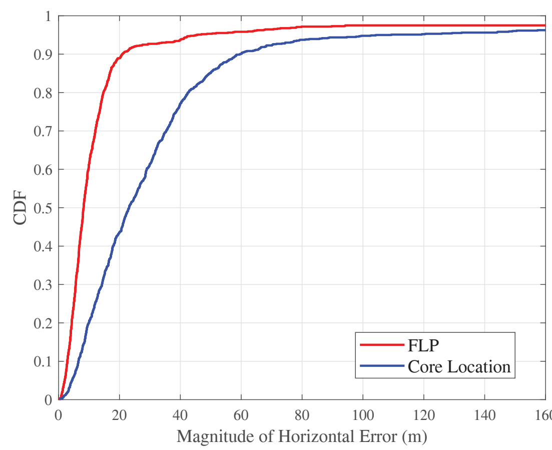

Fig. 3: CDF of Horizontal Error Magnitude over All T&E Scenarios in Building 2

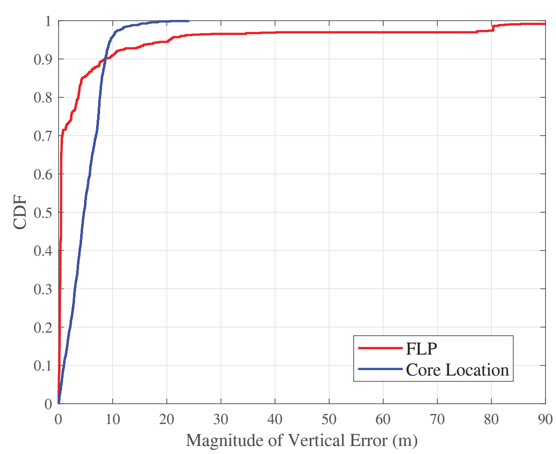

Fig. 4: CDF of Vertical Error Magnitude over All T&E Scenarios in Building 2

inside the building, except for a few in a small area of the building. This is depicted in Figure 2, where the path taken on the two floors of the building has been collapsed onto a single horizontal plane. Therefore, no results for Building 1 are presented in this section. Rather, we present performance results for Buildings 2-5 only.

Table III shows the performance metrics we used in this study that are defined in ISO/IEC 18305 and can be found online in Appendix A of [13]. The only exception is elevation availability, which is defined as the ratio of number of dots at which elevation information was available to the total number of dots visited in a scenario.

Tables IV and V present, respectively, the results obtained in Buildings 2 and 5 for FLP and Core Location. In these tables as well as others later in the paper, "walking" refers to T&E Scenarios 1, 2, 9, and 10 as well as the walking segments in T&E Scenarios 12-14. "Others" refers to other mobility modes, namely, T&E Scenario 3 and 16 as well as non-walking segments in T&E Scenarios 12-14. Looking at the overall results in the tables, we conclude that FLP has an edge over Core Location in these buildings. This is true in the sense

TABLE III: Performance Metrics Used in this Paper

Symbol	Performance Metric
δ_z	Elevation Availability
$\mu_{\|\varepsilon_h\|}$	Mean of the Magnitude of Horizontal Error
$\mu_{\|\varepsilon_z\|}$	Mean of the Magnitude of Vertical Error
$\mu_{\|\varepsilon\|}$	Mean of the Magnitude of 3D Error
$\sigma^2_{\|\varepsilon_h\|}$	Variance of the Magnitude of Horizontal Error
$\sigma^2_{\|\varepsilon_z\|}$	Variance of the Magnitude of Vertical Error
$\sigma^2_{\|\varepsilon\|}$	Variance of the Magnitude of 3D Error
CE95	Circular Error 95%
VE95	Vertical Error 95%
SE95	Spherical Error 95%
CEP	Circular Error Probable
VEP	Vertical Error Probable
SEP	Spherical Error Probable

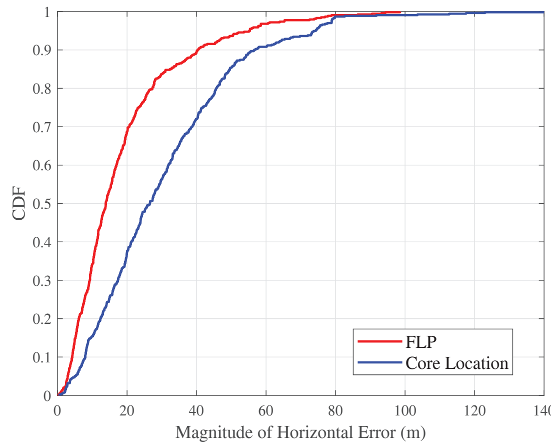

Fig. 5: CDF of Horizontal Error Magnitude over All T&E Scenarios in Building 5

TABLE IV: FLP vs Core Location Performance in Building 2

	Google FLP			Apple Core Location		
	Walking	Others	Overall	Walking	Others	Overall
# of Dots	774	184	958	774	184	958
δ_z	91.0%	99.5%	92.6%	100.0%	100.0%	100.0%
$\mu_{\|\varepsilon_h\|}$	14.51	19.80	15.53	34.35	28.17	33.16
$\mu_{\|\varepsilon_z\|}$	4.93	4.88	4.92	4.89	5.46	5.00
$\mu_{\|\varepsilon\|}$	16.57	20.82	17.38	35.27	29.44	34.15
$\sigma^2_{\|\varepsilon_h\|}$	862.74	1239.08	938.16	1437.10	1018.88	1361.57
$\sigma^2_{\|\varepsilon_z\|}$	280.21	227.62	269.86	8.53	15.37	9.88
$\sigma^2_{\|\varepsilon\|}$	1103.31	1449.13	1171.09	1405.38	990.21	1329.81
CE95	42.22	57.26	45.56	125.92	67.76	109.33
VE95	20.89	20.22	20.66	9.21	11.87	9.63
SE95	56.42	60.22	60.22	126.15	68.19	109.57
CEP	7.94	10.19	8.22	23.86	21.84	23.15
VEP	0.45	0.51	0.45	4.63	6.02	4.69
SEP	8.39	10.32	8.65	25.11	22.82	24.54

TABLE V: FLP vs Core Location Performance in Building 5

	Google FLP			Apple Core Location		
	Walking	Others	Overall	Walking	Others	Overall
# of Dots	399	134	533	399	134	533
δ_z	65.9%	47.0%	61.2%	100%	100%	100%
$\mu_{\|\varepsilon_h\|}$	15.71	27.76	18.74	28.23	39.96	31.18
$\mu_{\|\varepsilon_z\|}$	6.96	3.28	6.04	8.22	3.66	7.08
$\mu_{\|\varepsilon\|}$	18.84	28.24	21.21	30.54	40.38	33.01
$\sigma^2_{\|\varepsilon_h\|}$	137.32	572.79	273.28	879.98	704.50	860.38
$\sigma^2_{\|\varepsilon_z\|}$	73.65	6.43	59.27	58.60	2.71	48.45
$\sigma^2_{\|\varepsilon\|}$	151.01	562.98	270.35	870.41	686.88	841.14
CE95	39.99	77.79	55.34	56.42	79.81	74.30
VE95	24.47	7.19	22.91	20.39	6.60	19.68
SE95	42.87	78.12	55.48	58.82	79.84	74.31
CEP	12.88	17.71	13.87	23.91	36.46	26.53
VEP	2.29	2.96	2.69	5.48	3.57	4.58
SEP	16.72	17.71	17.12	25.55	37.05	27.84

of horizontal, vertical, and 3D error magnitudes, regardless of whether we look at the means of these error magnitudes or the 50- or 95-percentile points of the respective cumulative distribution functions (CDFs). The only exception is vertical error 95%, where Core Location has an advantage over FLP in Building 2 and a slight advantage in Building 5.

Figures 3 and 4 depict, respectively, the CDFs of horizontal and vertical error magnitude for FLP and Core Location over all T&E scenarios in Building 2. Figures 5 and 6 show corresponding information for Building 5. Figures 3 and 5 show once again that FLP has lower horizontal error than Core Location in these buildings. Figures 4 and 6 show crossing CDFs for FLP and Core Location. This implies that the probability distribution for the magnitude of vertical error for FLP has a heavier tail than the corresponding probability distribution for Core Location. As for vertical error, neither FLP nor Core Location has an advantage over the other one in these buildings.

In general, the performance results for Buildings 2 and 5 exhibited the same patterns. These are both multi-story buildings with a fairly high density of Wi-Fi APs (one in roughly every 250 m^2 of area).

Tables VI and VII present, respectively, the results obtained in Buildings 3 and 4 for FLP and Core Location. Looking at the overall results in the tables, we conclude that FLP and Core Location have comparable horizontal and 3D error magnitudes in these buildings. As for vertical error magnitude, Core Location is better than FLP. These statements are true in the sense of the means of these error magnitudes and the 50- or 95-percentile points of the respective CDFs.

Figures 7 and 8 depict, respectively, the CDFs of horizontal and vertical error magnitude for FLP and Core Location over all T&E scenarios in Building 3. Figures 9 and 10 show corresponding information for Building 4. Figures 7 and 9 show once again that FLP and Core Location have comparable horizontal error in these buildings. As for vertical

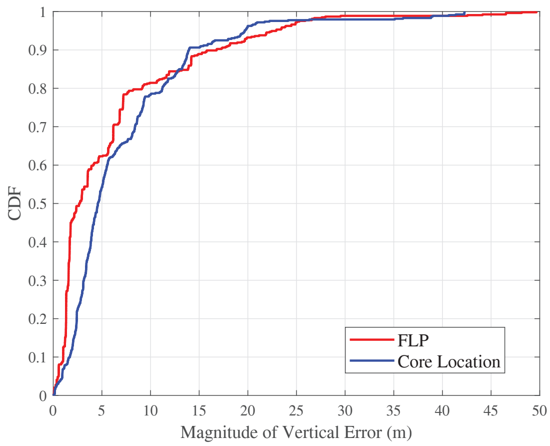

Fig. 6: CDF of Vertical Error Magnitude over All T&E Scenarios in Building 5

TABLE VI: FLP vs Core Location Performance in Building 3

	Google FLP			Apple Core Location		
	Walking	Others	Overall	Walking	Others	Overall
# of Dots	774	183	957	774	183	957
δ_z	86.4%	85.2%	86.2%	100%	100%	100%
$\mu_{\|\varepsilon_h\|}$	27.27	18.14	25.52	28.10	25.26	27.55
$\mu_{\|\varepsilon_z\|}$	7.04	0.95	5.88	2.05	2.08	2.06
$\mu_{\|\varepsilon\|}$	29.64	18.19	27.45	28.32	25.38	27.76
$\sigma^2_{\|\varepsilon_h\|}$	1204.09	156.70	1016.35	2065.29	187.46	1706.88
$\sigma^2_{\|\varepsilon_z\|}$	121.57	0.44	104.13	2.79	0.61	2.37
$\sigma^2_{\|\varepsilon\|}$	1240.33	156.10	1052.91	2059.41	186.09	1701.95
CE95	54.80	38.89	54.23	52.12	47.66	52.12
VE95	27.71	1.50	27.30	5.21	3.62	4.97
SE95	58.86	38.91	57.60	52.13	47.80	52.13
CEP	19.65	15.80	18.88	22.66	22.58	22.64
VEP	1.50	0.72	1.49	1.37	2.02	1.68
SEP	22.09	15.82	20.68	22.85	22.68	22.81

error, Figures 8 and 10 show once again that Core Location has better performance than FLP in these buildings. We notice that the probability distribution of magnitude of vertical error for FLP has a much heavier tail than the corresponding probability distribution for Core Location.

In general, the performance results for Buildings 3 and 4 exhibited the same patterns. These are both single-story buildings with a rather low density of Wi-Fi APs (one in roughly every 500 m^2 of area).

Tables IV-VII present data on two other issues that are worth discussing. One is a comparison of walking with other mobility modes. We focus on horizontal error, because all scenarios involving other mobility modes were carried out on the same floor in each of the four buildings. (Buildings 3 and 4 are single-story buildings anyway.) Therefore, there were no floor changes, which made these scenarios rather easy for elevation estimation. The mean of horizontal error magnitude and the 50- or 95-percentile points on its CDF are mostly lower for walking than for other mobility modes in Buildings 2 and 5. In Buildings 3 and 4, it is the other way around. Therefore, it

TABLE VII: FLP vs Core Location Performance in Building 4

	Google FLP			Apple Core Location		
	Walking	Others	Overall	Walking	Others	Overall
# of Dots	253	75	328	253	75	328
δ_z	87.7%	100.0%	90.5%	100.0%	100.0%	100.0%
$\mu_{\|\varepsilon_h\|}$	16.39	17.40	16.62	17.46	15.40	16.99
$\mu_{\|\varepsilon_z\|}$	7.95	10.34	8.49	2.94	3.76	3.13
$\mu_{\|\varepsilon\|}$	20.00	22.01	20.46	17.88	16.10	17.47
$\sigma^2_{\|\varepsilon_h\|}$	84.94	90.54	86.13	117.77	45.91	101.90
$\sigma^2_{\|\varepsilon_z\|}$	79.95	61.32	76.50	1.99	3.22	2.38
$\sigma^2_{\|\varepsilon\|}$	96.44	75.92	92.22	113.53	41.03	97.33
CE95	32.17	37.86	32.74	40.60	28.91	38.35
VE95	24.57	23.58	23.58	5.45	5.49	5.48
SE95	38.07	38.75	38.29	40.64	28.91	38.40
CEP	15.42	15.31	15.35	14.99	14.71	14.88
VEP	2.56	10.56	5.57	2.48	4.71	3.35
SEP	19.41	21.51	20.29	15.24	15.28	15.24

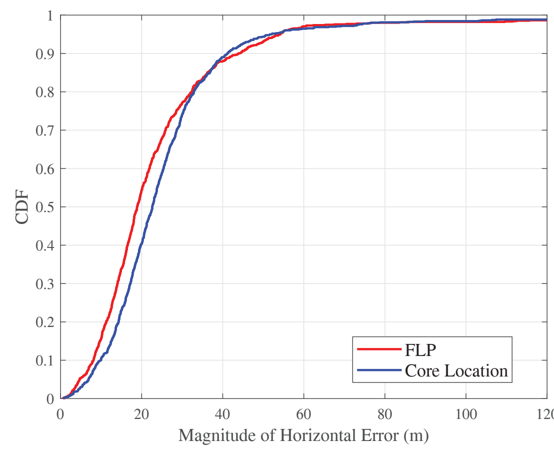

Fig. 7: CDF of Horizontal Error Magnitude over All T&E Scenarios in Building 3

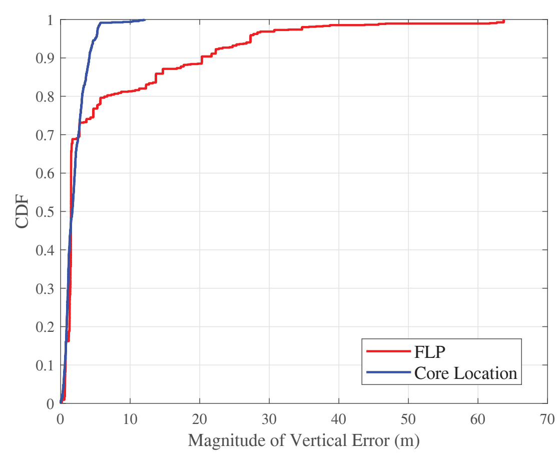

Fig. 8: CDF of Vertical Error Magnitude over All T&E Scenarios in Building 3

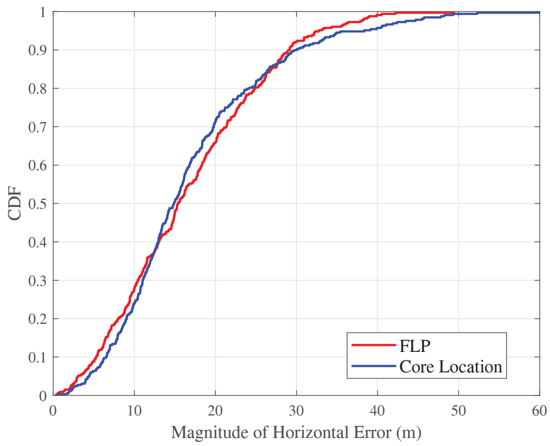

Fig. 9: CDF of Horizontal Error Magnitude over All T&E Scenarios in Building 4

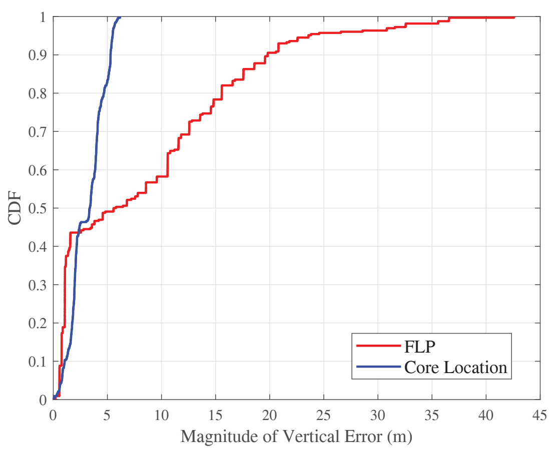

Fig. 10: CDF of Vertical Error Magnitude over All T&E Scenarios in Building 4

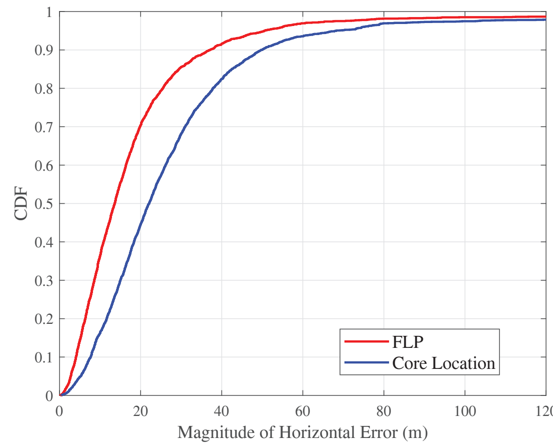

Fig. 11: CDF of Horizontal Error Magnitude over All T&E Scenarios and All Buildings

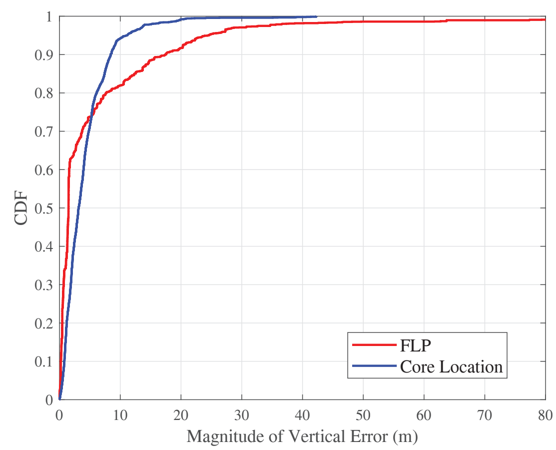

Fig. 12: CDF of Vertical Error Magnitude over All T&E Scenarios and All Buildings

is plausible to conclude that the mobility mode does not have much of a bearing on FLP or Core Location performance.

The second issue is that FLP sometimes does not generate an elevation estimate. The percentage of dots at which FLP produced elevation information in each scenario is shown in Tables IV-VII. Specifically, the overall elevation availability was only 61.2% in Building 5. Note that Core Location produced elevation information all the time.

Figures 11 and 12 depict, respectively, the CDFs of horizontal and vertical error magnitudes for FLP and Core Location over all T&E scenarios and all buildings. They show that FLP clearly provides better horizontal accuracy than Core Location. However, the CDFs for vertical error magnitude cross, and hence there are no winners as far as vertical error is concerned. Figure 11 shows that FLP and Core Location provide, respectively, 25.5 m and 37.8 m horizontal accuracy, hence meeting FCC's 2021 requirement for horizontal accuracy for E911 calls *today*.

IV. Comparison of FLP with the Best PerfLoc App

Table VIII provides a comparison of FLP and the best PerfLoc app in Building 5. Over all T&E scenarios, the app that won the PerfLoc Prize Competition is better than FLP by 1.08 m in the sense of mean horizontal error magnitude. When it comes to the 95-percentile point on the CDF of horizontal error magnitude, FLP is better than the PerfLoc app by 17.5 m, which is quite significant. As for vertical error magnitude, the best PerfLoc app is significantly better than FLP. For the walking scenarios (namely, 1, 2, 9, and 10) the PerfLoc app outperforms FLP in every respect. For other mobility modes, sometimes FLP is better than the best PerfLoc app, and sometimes it is the other way around. It is already known that the best PerfLoc app, which uses IMU sensors in addition to various RF signals (Wi-Fi, GPS, and cellular), does not do a good job of handling other mobility

TABLE VIII: Comparison of FLP and PerfLoc Performance in Building 5

	Walking		Cart		Sidestepping		Walking Backward		Crawling		Overall	
	FLP	PerfLoc	FLP	PerfLoc	FLP	PerfLoc	FLP	PerfLoc	FLP	PerfLoc	FLP	PerfLoc
$\mu_{\|\varepsilon_h\|}$	15.71	10.76	25.16	49.67	33.36	27.79	32.75	38.07	7.75	10.31	18.74	17.66
$\mu_{\|\varepsilon_z\|}$	6.96	2.09	2.15	0.72	4.58	0.57	4.94	0.20	0.09	0.34	6.04	1.71
$\mu_{\|\varepsilon\|}$	18.84	11.43	25.42	49.68	34.18	27.81	33.43	38.07	7.75	10.32	21.21	18.16
CE95	39.99	24.86	60.56	78.92	94.36	79.17	73.87	83.10	14.26	18.56	55.34	72.84
VE95	24.47	5.22	3.56	1.16	6.18	0.95	7.21	0.47	0.11	0.57	22.91	5.20
SE95	42.87	24.86	60.63	78.92	94.56	79.17	74.22	83.10	14.26	18.56	55.48	72.84

modes [6]. Therefore, as far as NIST is concerned, much room for improvement is still left in developing smartphone indoor localization apps.

The results presented in Table VIII should be viewed with one important caveat. The precise locations and MAC addresses of all Wi-Fi APs in Building 5 were available to the best PerfLoc app. FLP does not have access to such information, but it collects and uses crowdsourced smartphone data from anyone that visits any building and has location service enabled to estimate the locations of Wi-Fi APs and cell towers. Overall, it is fair to say that the best PerfLoc app had to do better than FLP, because it had access to the Wi-Fi AP locations.

V. Conclusions

In this paper, we presented comprehensive test results obtained in five large buildings for Android FLP and iOS Core Location based on the testing procedures of the international standard ISO/IEC 18305. Based on our overall results, FLP clearly achieves better horizontal accuracy than Core Location, but no clear picture emerges when one looks at vertical accuracy. Putting these together, overall FLP yields better accuracy than Core Location. It is observed that Core Location typically provides more stable elevation information with lower error variance than FLP.

We also presented performance results from one building showing that the Android app that won the PerfLoc Prize Competition has a slight edge over FLP. This comparison, however, is not entirely fair. While the best PerfLoc app had access to the precise 3D coordinates of Wi-Fi APs in the building, FLP has access to the vast crowdsourced data that Google collects from Android phones to provide more accurate location estimates, among other things.

We plan to analyze another treasure trove of data we collected by obtaining a large number of location estimates from FLP and Core Location at every single dot in four of the five NIST buildings we used for testing. The purpose of the additional work is to study the statistical variations of location estimates at a given location.

Acknowledgments

We would like to thank Professor Anthony Rowe of Carnegie Mellon University and his team, particularly Niranjini Rajagopal, for modifying an iOS app they had developed so that it would collect the iPhone data we needed and serve as the Evaluation App for Core Location in this study.

Disclaimer

Certain commercial entities, equipment, or materials may be identified in this document in order to describe an experimental procedure or concept adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose.

References

[1] International Data Corporation, “Smartphone OS Market Share.” [Online]. Available: https://www.idc.com/promo/smartphone-market-share/os

[2] Google, “Fused Location Provider API.” [Online]. Available: https://developers.google.com/location-context/fused-location-provider/

[3] Apple, “Core Location.” [Online]. Available: https://developer.apple.com/documentation/corelocation

[4] “ISO/IEC 18305:2016, Information technology – Real time locating systems, Test and evaluation of localization and tracking systems.” [Online]. Available: https://www.iso.org/standard/62090.html

[5] N. Moayeri, M. Ergin, F. Lemic, V. Handziski, and A. Wolisz, “PerfLoc (Part 1): An extensive data repository for development of smartphone indoor localization apps,” in *Proc. IEEE Int. Symp. Pers. Indoor Mobile Radio Comm. (PIMRC)*, Valencia, Spain, Sep. 2016, pp. 1–7.

[6] N. Moayeri, C. Li, and L. Shi, “PerfLoc (Part 2): Performance evaluation of the smartphone indoor localization apps,” in *Proc. Int. Conf. Indoor Positioning Indoor Navigation (IPIN)*, Nantes, France, Sep. 2018, pp. 1–8.

[7] Federal Communications Commission, “Third Further Notice of Proposed Rulemaking.” [Online]. Available: https://www.documentcloud.org/documents/2195636-fcc-third-nprm-february-2014.html#document/p15

[8] FireceWireless, “Google launches Emergency Location Service in U.S. with T-Mobile.” [Online]. Available: https://www.theverge.com/2018/2/16/17021690/google-conducts-test-help-911-accurately-locate-callers

[9] The Verge, “Google conducts test to help 911 accurately locate callers.” [Online]. Available: https://www.fiercewireless.com/wireless/google-launches-emergency-location-service-u-s-t-mobile

[10] Apple, “Apple’s iOS 12 securely and automatically shares emergency location with 911.” [Online]. Available: https://www.apple.com/newsroom/2018/06/apple-ios-12-securely-and-automatically-shares-emergency-location-with-911/

[11] CNN, “These iOS 12 features may improve your life.” [Online]. Available: https://money.cnn.com/2018/09/17/technology/ios-12-features/index.html

[12] Federal Communications Commission, “Indoor Location Accuracy Benchmarks.” [Online]. Available: https://www.fcc.gov/public-safety-and-homeland-security/policy-and-licensing-division/911-services/general/location-accuracy-indoor-benchmarks

[13] N. Moayeri, L. Shi, and C. Li, “Test & evaluation of the safety app from 911Cellular, LLC at NIST,” *NIST Technical Note 1962*, Apr. 2017. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/TechnicalNotes/NIST.TN.1962.pdf