Computer Communications 132 (2018) 96–110

Contents lists available at ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

Pre-crowdsourcing: Predicting wireless propagation with phone-based channel quality measurements

Rita Enami *, Yan Shi, Dinesh Rajan, Joseph Camp

Southern Methodist University, United States

ARTICLE INFO

Keywords:
Crowdsourcing
Cellular network measurements
LTE
Path loss evaluation
Radio propagation modeling

ABSTRACT

Conducting in-field performance analysis for wireless carrier coverage and capacity is extremely costly in terms of equipment, manpower, and time. At the same time, there is a growing number of opportunities for crowdsourcing of network and sensor information via smart applications, firmware, and cellular standards. These facilities offer carriers feedback about user-perceived wireless channel quality. Crowdsourcing provides the ability to rapidly collect feedback with dense levels of penetration using client smartphones. However, mobile phones often fail to capture the fidelity and high sampling rates of more-advanced equipment (*e.g.*, a channel scanner) used when drive testing for analysis of propagation characteristics. In this work, we quantify the impact of various effects induced by mobile phones when interpreting signal quality such as averaging over multiple samples, imprecise quantization, and non-uniform and/or less frequent channel sampling. To do so, we conduct extensive in-field experiments across heterogeneous devices and environments to empirically characterize the path loss via phone measurements on LTE networks. We find that mobile phones are comparable to advanced equipment in inferred radio propagation, within 0.1 of the calculated path loss exponent by the channel scanner across downtown, single-family, and multi-family residential areas. Lastly, we develop an intuition for the importance crowdsourcing-based propagation prediction by evaluating the effect on coverage estimation when deploying operational networks. For example, we find that even a prediction error of 0.4 for the path loss exponent would cause a 40% redundancy in the covered area or coverage holes for 25% of the targeted area based on whether the error was above or below the actual value, respectively.

Acronym List

API Application Program Interface
ASU Arbitrary Strength Units
BER Bit Error Rate
DRX Discontinuous Reception
GSM Global System for Mobile
KPI Key Performance Indicator
KS Kolmogorov–Smirnov Test
LiDAR Light Detection and Ranging
LTE Long Term Evolution
MDT Minimization of Drive Test
QoS Quality of Service
RMSE Root Mean Square Error
RSRP Reference Signal Received Power
RSRQ Reference Signal Received Quality
SNR Signal-to-Noise Ratio
TETRA Terrestrial Trunked Radio
UE User Equipment
UMTS Universal Mobile Telecommunications System

1. Introduction

Cellular network providers collect and analyze radio signal measurements continuously to improve network performance and optimize network configuration. Available methods to obtain the signal measurements consist of drive testing, network-side-only tools, dedicated testbeds, and crowdsourcing [1]. The former three methods are extremely resource intensive. For example, one common approach for capturing radio signal measurements is to outfit a backpack with six mobile phones running various applications and network protocols alongside an expensive mobile channel scanner (see Fig. 1) for network engineers to gather data on foot. Vehicles are often used for a greater number of and potentially higher-powered and more costly devices, allowing higher levels of mobility in a targeted region. In congested areas with various technologies (*e.g.*, LTE, GSM, UMTS, and TETRA) the problem becomes worse: to get an acceptable quality of service, data collection should be repeated multiple times per roll out of each technology to appropriately configure the network [2]. Further complicating matters, physical changes to the environment such as construction of new

* Corresponding author.
E-mail addresses: renami@smu.edu (R. Enami), shiy@smu.edu (Y. Shi), rajand@smu.edu (D. Rajan), camp@smu.edu (J. Camp).

https://doi.org/10.1016/j.comcom.2018.10.003

Available online 16 October 2018

Nomenclature

γ_T	Path loss exponent from TSMW measurements
Q	Qualipoc application
T	Channel scanner, TSMW
W	WiEye application
$\#B$	Number of buildings
$\#T$	Number of trees
α	A constant from the transmitter power, antenna heights and gains
Δ_γ	Path loss exponent range
γ	Path loss exponent
γ_Q	Path loss exponent from Qualipoc measurements
γ_W	Path loss exponent from WiEye measurements
$\gamma_W{}'$	Path loss exponent from matched measurement of WiEye
$\gamma_Q{}'$	Path loss exponent of matched measurement from Qualipoc
$\overline{B_h}$	Average height of the buildings
$\overline{R_{GE}}$	Ground elevation of the receiver
$\overline{T_h}$	Average height of the trees
σ	Standard deviation of a Gaussian distribution
B_{h_σ}	Standard deviation of buildings' height
BTS_{GE}	Ground elevation of the transmitter
F_u^γ	Percentage of a useful service area
R_{GE_σ}	Standard deviation of the ground elevation
T_{h_σ}	Standard deviation of trees' height
x	Received signal level at distance d
x_0	Receiver threshold
$\triangle$Q&T	Path loss exponent offset between channel scanner & Qualipoc measurements
$\triangle$W&T	Path loss exponent offset between WiEye & Qualipoc measurements

Fig. 1. Typical Rohde & Schwarz backpack for walk/drive testing (left) and TSMW channel scanner (right) [3].

buildings or highways can decrease the effectiveness of the obtained data.

Crowdsourcing is an economical alternative to these resource-intensive methods that has the additional benefit of considering the in-situ performance at the end-user device. Consequently, many carriers are rolling out smart applications, firmware, and standardization efforts to crowdsource perceived channel state by user equipment (UE). Furthermore, LTE release 11 in 3GPP TS 37.320 [4] has developed a Minimization of Drive Test (MDT) specification to monitor the network Key Performance Indicators (KPIs) via crowdsourcing. In [5], the use case scenarios for MDT are determined as follows: coverage optimization, mobility optimization, capacity optimization, parametrization for common channels, and Quality of Service (QoS) verification. The coverage optimization topic contains some other use cases, such as coverage mapping, detection of excessive interference, and overshoot coverage detection.

While there is less control of the factors leading to a recorded channel quality, there are many advantages to crowdsourcing this information in terms of lessening the need for costly equipment, reduced in-field man hours, rapid scalability of data sets, and penetration into restricted physical locations. These advantages have sparked a number of works where crowdsourcing has been utilized to identify network topology [6], perform real-time network adaptation [7], characterize Internet traffic [8], detect network events [9], fingerprint and georeference physical locations [10], assess the quality of user experience [11], and study network neutrality [12]. The bandwidth, latency, and throughput have previously been used as crowdsourced KPIs [13,14] to evaluate wide-area wireless network performance [15] and in-context performance [16].

However, mobile phones possess a number of shortcomings when compared to a channel scanner in reporting channel quality, such as: *(i)* averaging over multiple samples, which can flatten channel fluctuations [13] with manufacturer-specific methodologies to estimate the received signal power [17], *(ii)* coarse quantization, which can impose a unit step for minuscule changes, *(iii)* sampling at non-uniform intervals when crowdsourcing information as opposed to long, consecutive testing periods recorded when drive testing, and *(iv)* clipping that results from less sensitive receivers with fringe network connectivity. The accuracy of the received signal reporting by mobile phones as compared to a channel scanner was evaluated in [18], but the effect of averaging, the impact on path loss calculation, and the resulting coverage estimation impact was not considered. Hence, while a crowdsourcing framework for characterizing wireless environments would have tremendous impact on drive testing costs, we believe that a first step in doing so requires understanding the viability of mobile phones to replace more advanced measurement equipment as channel modeling probes.

In this work, we study the impact of various effects induced by user equipment when sampling signal quality. These shortcomings include averaging over multiple samples, imprecise quantization, and non-uniform and/or less frequent channel sampling. We specifically investigate the accuracy of characterizing large-scale fading using crowdsourced data in the presence of the aforementioned phone measurement shortcomings. To do so, we perform extensive in-field experimentation to quantify the impact of each of these four effects when evaluating the viability of mobile phones to characterize the path loss exponent, a metric commonly used by carriers for deployment planning, frequency allocation, and network adaptation. Our results indicate that the inferred propagation parameters by smartphone measurements in GSM and LTE networks is comparable to those obtained by the advanced equipment that are frequently used by drive testers (*e.g.*, channel scanners). Finally, we analyze the impact of the path loss prediction error on a carrier's misinterpretation of coverage area and predicted network throughput. In wireless networks, the percentage of coverage area is determined by a region over which the signal level exceeds the sensitivity level with a specified level of probability. This value is the likelihood of coverage at the cell boundary and a function of the received signal level. Therefore, an accurate network design will avoid possible gaps in the network (overestimating propagation) or interference in adjacent cells (underestimating propagation), which both affect the network throughput [19]. In particular, our work makes the following four contributions.

First, we set forth a framework to evaluate the impact of strictly using mobile phones (as opposed to a channel scanner) in propagation

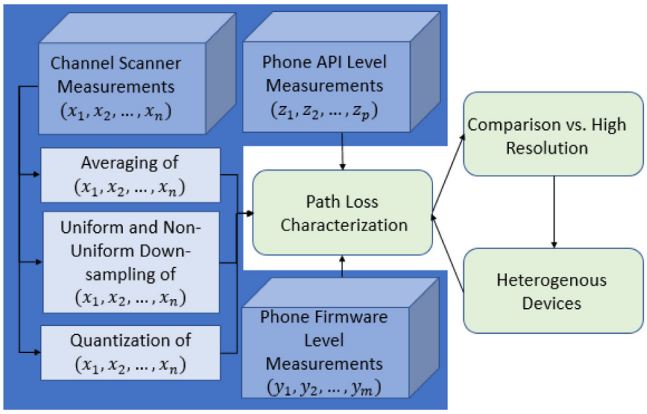


Fig. 2. Pre-processing and post-processing of collected data by channel scanner and mobile phones.

prediction. As depicted in Fig. 2, we consider how the averaging, uniform and non-uniform downsampling over time and space, and quantization of mobile phone channel quality samples at both the firmware and API levels affect the path loss characterization. At the API level, we have designed a freely-available Android application called WiEye, which can help users globally analyze spectrum in an economical manner. Additionally, WiEye functions as a crowdsourcing tool, which has captured over 250 million signal quality measurements from over 60 thousand users worldwide (protected by an IRB). At the firmware level, we capture signal quality directly from the hardware via a Rohde & Schwarz tool called Qualipoc.

Second, we compare the perceived channel quality across the channel scanner, multiple mobile phone models, and various levels of the software stack. To do so, we perform extensive local experiments across downtown, single-family residential, and multi-family residential regions and directly compare the received channel quality as reported by the channel scanner to mobile phone firmware-level and API-level data, where each mobile phone measurement considered has a corresponding channel scanner measurement for comparison. We initially observe that even over different sectors from the same base station in the same region type there can be a 0.4 difference in inferred path loss exponent and identify some of the geographical features that are responsible for this variation. More generally, as compared with the path loss exponent calculated for each region based on the channel scanner, we find that the firmware level measurements had an average path loss exponent estimation error of 0.06, 0.06, and 0.1 and API-level measurements had an error of 0.12, 0.13, and 0.11 for the single-family, multi-family, and downtown regions, respectively. This result considers the same number of samples from each device for direct comparison. Each prediction error occurred in the positive direction, meaning the value of the predicted path loss exponent from the mobile phone was greater than that predicted by the channel scanner data, an observation that can be used for future MDT calibration. We also examine the range over which each user-side device and software is able to receive cellular base station transmissions (*i.e.*, their sensitivities) to understand where clipping of crowdsourced data might occur.

Third, we quantify the impact on inferring propagation characteristics from the various calculations and imperfections that mobile phones induce on received channel quality before reporting it to the user. To do so, we consider numerous data sets from the channel scanner in the aforementioned environmental contexts and impose these imperfections to understand their role by evaluating against the root mean-squared error of path loss prediction from the original channel scanner data set in that region. Our results show that the path loss parameters obtained by mobile phone samples are sufficiently comparable to the advanced drive testing equipment, paving the way for crowdsourcing as a viable solution for in-field performance analysis.

Fourth, considering the fact that any error in path loss estimation will ultimately affect the coverage area estimation and Bit Error Rate (BER), we build intuition about the prediction errors reported throughout the previous sections of the paper as they relate to network planners and operators by quantifying the impact on coverage estimation and user BERs. Since we observe path loss exponents ranging from 2 to 4 from our crowdsourcing platform, we consider a situation in which the actual path loss exponent is 3, but errors in prediction range from −1 to +1. In doing so, we allow a continuum of analysis about how much the network holes (overestimating propagation) or redundancy (underestimating propagation) might exist from the original targeted area. In particular, a modest propagation overestimation error of −0.4 (13% error) from an actual path loss exponent of 3 results in a quarter (25%) of the targeted area having coverage holes in regions that were assumed to be covered. Conversely, the same modest propagation underestimation error of +0.4 from actual path loss exponent of 3 would result in a 40% overlap in the targeted coverage region. While the percentage of error is very small (-/+13%), the impact on coverage estimation is large. In terms of user BER, such a 0.4 prediction error frequently raises the BER by an order of magnitude for many situations (*e.g.*, predicting 2.1 but an actual path loss exponent of 2.5, a relative error of only 16%, for an SNR of 15). In other words, at locations where there was assumed to be moderate to high SNR, the prediction errors can have a dramatic effect on user performance. For example, some services like video streaming require a specific throughput. Small variations in throughput will increase the latency of live streams, especially at the cell boundaries.

The remainder of the paper is organized as follows. We discuss related work in Section 2. In Section 3, we experimentally quantify the channel quality reporting differences of mobile phones versus a channel scanner. In Section 4, we analyze the role of mobile phone imperfections in terms of path loss prediction. Section 5 relates path loss prediction error to coverage estimation for operational networks. Lastly, we conclude in Section 6.

2. Related work

The Minimization of Drive Tests (MDT) initiative in the 3GPP standard has been created to exploit the ability of smartphones to collect radio measurements in a wide range of geographical areas to enhance coverage, mobility, capacity, and path loss prediction [20]. Also, a few measurement studies have used API-level measurements to estimate different KPIs of cellular networks [13,14,21,22]. They each measured KPIs in terms of throughput, received signal power, and delay and involved regular users to provide measurements (*i.e.*, crowdsourcing) across large geographical regions in some cases. In contrast, we focus on characterizing the wireless channel using diverse end-user devices at different levels of the software stack. Predicting the cellular network coverage by using the crowdsourced data has been studied in a few studies. For example, network coverage maps using crowdsourced data is studied in [23]. However, the authors provided the observed received signal level without a discussion of the differences across end-user devices. In addition, another work used a similar idea of using crowdsourced data along with interpolation techniques to predict the coverage area [24]. Although, the impact of location inaccuracy and data distribution of the interpolation techniques was investigated, the impact of the imperfections of end-user devices was not explored. In fact, [25] argued that [24] suffers from a lack of control and repeatability of capturing data and piggy-backing mobile broadband measurements onto public transport infrastructures.

Furthermore, others proposed the Bayesian Prediction method to improve the coverage estimation obtained by drive testing and MDT measurements, but the results were strictly based on advanced devices as opposed to mobile phone measurements [26]. The provided X-map accuracy from simulated data in [27] has been evaluated in terms of position inaccuracy, UE inaccuracy, and number of measurements. However, to analyze crowdsourced data, using in-field experimentation

is important to distinguish between the performance of more advanced equipment versus a mobile phone in channels similar to those experienced by user devices. Furthermore, three major application scenarios for spatial big data obtained by performing MDT in a wireless network are depicted in [28]. Also, it has been shown that massive amounts of data needs a high-performance processing platform and solutions to obtain meaningful conclusions. Hence, [28,29] have focused on providing a platform to deal with big data regarding different applications. To estimate the channel quality, we are using RSRP as our metric from the LTE standard. It was previously observed by [17] that the reported value by a mobile phone in terms of RSRP is influenced by averaging but did not consider the compounding effects. Similarly, [18] depicts that the received signal power by commercial phones is comparable to an advanced tool. While this is close in nature, we also consider many of the spatial and temporal downsampling effects that would cause imprecise estimation of the path loss estimation for a given environment and develop a carrier-focused intuition of the network and user impact of these errors.

3. In-field calibration of received signal power from mobile phones

The purpose of this study is to compare the ability of mobile phone measurements, captured either at the API level or the firmware level, to an advanced measurement tool in characterizing wireless channels in terms of path loss. Before doing so, in this section, we compare and calibrate the raw measurements provided by diverse mobile phones at different levels of the software stack with data provided by a channel scanner.

API-level phone data. At the API level, we modify our Android application WiEye, which we designed to crowdsource measurements, to log signal quality measurements at the highest sampling rate that the operating system will allow (1 Hz). Since WiEye can be installed on any Android-based phone, we can compare API-level measurements across a wide array of devices. In our study, we use four different mobile phones: (*i.*) Samsung S5, (*ii.*) Nexus 5, (*iii.*) Google Pixel, and (*iv.*) Samsung S8. While the former two phones are not the latest models, they provide a comparison across multiple generations, and the Samsung S5 is the phone that allows a firmware-based tool that we will now discuss.

Firmware-level phone data. At the firmware level, we have purchased a software tool called Qualipoc from Rohde & Schwarz, which allows signal strength measurements to be reported directly from the chipset. Qualipoc can receive the channel quality information from many diverse technologies, such as LTE, GSM, and WCDMA. The sampling rate of the Qualipoc is approximately 3 Hz. Unlike the channel scanner, the mobile phones continuously search for the best visible base station by measuring the signal power received from multiple base stations, affecting both the API-level and firmware-level measurements.

Channel scanner data. To replicate the measurement process typically performed by drive testers, we use a commonly-used Rohde & Schwarz TSMW Channel Scanner for obtaining detailed signal quality measurements. The TSMW can passively and continuously monitor numerous technologies in 30 MHz - 6 GHz frequency range, with a sampling rate of 500 Hz. The scanner is controlled by Romes software (version 4.89), which is installed on a laptop connected via an Ethernet cable to the TSMW.

In-field measurement setup and calibration. We conduct a measurement campaign across three diverse regions of Dallas, Texas with respect to terrain type: single-family residential, multi-family residential, and downtown. All five device types are connected to the same network operator for direct comparison and perform measurements in parallel on a co-located roof of a car. In each region, we observe cellular transmissions and record data from 11 total base stations.

We first quantify the signal quality sensitivities of each device for measurements taken at the same time and location. To do so, we applied a post-processing procedure on the entire collected data set. Since the sampling rate of the channel scanner is higher than that of Qualipoc (firmware) or WiEye (API), we extract the samples from the channel scanner data set, which are the closest in time to that of WiEye and Qualipoc. The matching process consists of two steps: (*i.*) grouping measurements based on the transmitting base station, and (*ii.*) downsampling channel scanner data to have the same number of samples as the Qualipoc and WiEye's data set, where each mobile phone sample has a corresponding channel scanner measurement in time. If the channel scanner did not report a measurement within one second of the mobile phone measurement, we do not consider that data point in our comparison.

Table 1
Field-tested range of reported signal quality (dBm) from channel scanner (TSMW), Qualipoc, and WiEye.

Device	Min	Max	Range
Channel scanner	−134	−56	77
Qualipoc phone	−129	−55	74
WiEye phone	−128	−57	71

Table 2
Average signal quality offsets (dBm) reported from Qualipoc and WiEye with matched channel scanner measurement.

Device Location	Qualipoc dBm Diff. (Mean)	WiEye dBm Diff. (Mean)
Downtown	−1.5 (−75.6)	−4.4 (−78.5)
Single-family	−1.3 (−82.5)	−3.8 (−85.0)
Multi-family	−1.9 (−78.4)	−4.1 (−80.3)

Table 1 shows the minimum, maximum, and range of the received signal power for all of these measurements across all cell towers in each region. As it is seen from the results, the widest range (77) and greatest sensitivity (−134 dBm) is captured by the channel scanner with the least range (71) and sensitivity (−128 dBm) captured by WiEye. The reduced range experienced by the mobile phone will cause some clipping on the extreme ends of the connectivity ranges, especially with poor signal quality.

Next, we again consider this downsampled data set which matches the time stamps across devices to consider the difference in reported signal quality per signal quality sample across devices. Table 2 shows the difference of WiEye compared to the matched channel scanner measurement and Qualipoc compared to the matched channel scanner measurement across the three region types. This measurement shows the dBm offsets that mobile phones could induce on a crowdsourced data set as compared to more advanced equipment. We also report the mean reported signal strength per region for completeness.

We observe that the difference in reported received signal level is on average 1.57 dBm higher on the channel scanner versus Qualipoc across the three regions with a range of 1.3 to 1.9. In contrast, the difference in reported received signal level is on average 4.43 dBm higher on the channel scanner versus WiEye across the three regions with a range of 4.1 to 4.8. These dBm offsets could affect the path loss characterization as a higher reported channel quality could lower the path loss exponent versus a lower reported channel quality which could raise the path loss exponent. In the following section, we will consider the role of these dBm offsets as well as multiple other mobile phone imperfections.

4. Leveraging mobile phone based measurements on path loss prediction

One of the most common metrics which drive testers use to evaluate a given region is path loss. Since we ultimately want to use mobile phone measurements in a crowdsourcing manner to obtain this metric, we need to understand the role of mobile phone imperfections on evaluating the path loss of a given environment. In particular, reported signal quality from mobile phones will have the following effects: averaging, uniform and non-uniform downsampling, and different resolutions caused by quantization. In this section, we first provide some background on path loss modeling and then experimentally evaluate the role of these mobile phone imperfections on path loss estimation.

4.1. Modeling large-scale fading: Path loss

Large-scale fading refers to the average attenuation in a given environment for transmission through and around obstacles in an environment for a given distance [30]. Path loss prediction models are classified into three different categories: empirical, deterministic, and semi-deterministic. Empirical models such as [31] and [32] are based on measurements and use statistical properties. However, the accuracy of these models is not as high as deterministic models to estimate the channel characteristics. These models are still widely-used because of their low computational complexity and simplicity. Deterministic models or geometrical models consider the losses due to diffraction, and detailed knowledge of the terrain is needed to calculate the signal strength [33,34]. These models are accurate. However, their computational complexity is high, and they need detailed information about the region of interest. Semi-deterministic models applied in [35] and [36] are based on empirical models and deterministic aspects.

In our study, we use empirical methods since it is the type of modeling that would be most appropriate to leverage crowdsourcing. The large-scale fading is a function of distance (d) between the transmitter and the receiver, and γ is the path loss exponent, where the path loss exponent varies due to the environmental type from 2 in free space to 6 in indoor environments. Some typical values are 2.7–3.5 in typical urban scenarios and 3–5 in heavily shadowed urban environments [30]. In this work, we focus on the inferred path loss exponent from mobile phone measurements and use a linear regression model to calculate the path loss exponent.

4.2. In-field analysis of inferred path loss across region and device types

As discussed in Section 3, our experimental analysis spans three region types (single-family residential, multi-family residential, and downtown) with multiple mobile phone types at the API-level (WiEye), with mobile phones at the firmware level (Qualipoc), and with a channel scanner (TSMW). All of these devices report which base station sector is transmitting the received signal. We performed the measurements while the car speed was maintained at approximately 20 mph. To avoid stopping at the traffic lights, we observed traffic light patterns, and we drove each route many times to record data regarding our requirements. In the future, we could consider predicting the future received signal strength concerning the UE speed and direction regarding the base station. To do so, we can record the compass data from the phone along with the signal strength, location information, and time stamp.

Since prior works have shown per region performance [30] and per sector performance can differ [15], we first analyze the variation of the path loss exponent from each region and each sector in three regions from the channel scanner to show some examples of the γ diversity. We consider the following three types of areas:

(1) A downtown region containing tall buildings and trees, which are non-uniformly distributed over the region.

(2) A single-family area that is covered by a high density of foliage and mostly two-story buildings.

(3) A multi-family area that has a mixture of vegetation and buildings of two stories or more in height.

Since path loss is not only a function of distance but is also affected by obstacles between the transmitter and receiver and environment type, we consider the geographical features in different areas to explain the variation of the observed path loss exponents (even within the same region type). For a more thorough investigation on the relationship between the geographical features and the role they play on propagation effects please refer to our recent work [37,38]. Here, we make the following observations:

(i) Path loss slope varies in different region types: To study the path loss exponent's variation in each region, we inferred all available γs corresponding to different sectors in each region. We eliminated the sectors with a low number of measurements as defined by the results in Fig. 10. We performed linear regression on each sector's signal strength measurements independently to find the path loss exponent for that sector. Then, in each region, we select the sectors containing the minimum and maximum γ.

Table 3
Minimum and maximum observed path-loss exponent per region and corresponding geographical features.

Region		γ	$\overline{B_h}$	B_{h_σ}	#B	$\overline{T_h}$	T_{h_σ}	#T	$\overline{R_{GE}}$	R_{GE_σ}	BTS_{GE}	Δ_γ
Single-family	Min	3.2	8.34	2.7	209	10	2.9	1846	176	2	180	.5
	Max	3.7	8.6	2.6	470	11	3.5	2300	186	7	180	
Multi-family	Min	2.9	10.4	15.3	21	8.7	3.1	230	184	1.9	186	1
	Max	3.9	11.8	11.5	98	12	14	500	184	3	176	
Downtown	Min	2.8	35	35	36	9.7	7.8	197	135	2	136	1
	Max	3.8	37	32	41	14	10	233	136	2.8	127	

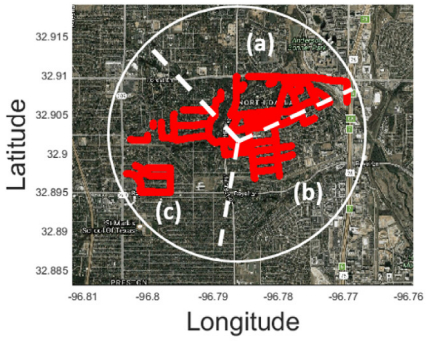
A received signal is a combination of transmitted signals, composed of reflected or scattered transmissions that are obscured by buildings or trees. Thus, the propagation environment is profoundly influenced by the path loss and affects the network performance. The impact of the buildings would be more visible in an urban environment where a diversity in building height surrounds the UE. In this work, we considered three different area types to measure. Table 3 shows the minimum and maximum path loss exponent obtained using channel scanner measurements for a particular sector in each region. As we can see, there are differences between the path loss exponent readings from different regions. To enhance our understanding about these differences, we provide detailed information of the buildings and foliage corresponding to each region. To provide 3-dimensional geographical features of a region, we used a database of Light Detection and Ranging (LiDAR) information from that region. This dataset contains detailed information of buildings and trees as we discussed extensively in our recent work [37].

The height of surrounding objects plays an important role on the signal attenuation, because the receiver height is typically lower than the clutter height. Therefore, we provide the average and standard deviation of object heights in that area. Here, B_h and T_h depict the average height of the buildings and trees and the standard deviation of these object heights for and γ_{max} and γ_{min} are depicted as B_{h_σ} and T_{h_σ}, respectively. Furthermore, we consider the number of objects (scatterers and reflectors) and ground elevation information of an area. The ground elevation information of the receiver and the transmitter would ultimately influence the difference between the clutter height and transmitter height.

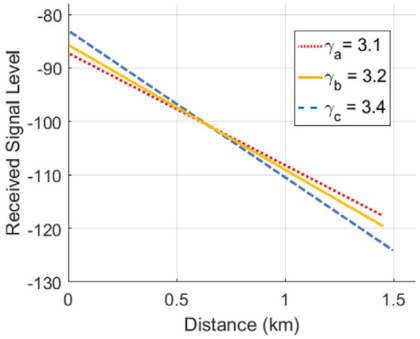
We observe that downtown and multi-family regions report the highest variation range of the path loss slope. In the single-family area, there is not as big of a difference in the average and standard deviation of the object heights for γ_{max} and γ_{min}. However, the number of the trees ($\#T$) and buildings ($\#B$) located in the sector corresponding to γ_{max} is much higher than the others. Furthermore, the ground elevation of the area is about 7 m higher than the ground elevation of the base station in the γ_{max} case. The range of the observed path loss exponent in each environment is depicted by Δ_γ, and the results show that the variation in the path loss exponent of multi-family and downtown regions are higher than the single-family area.

ii. γ varies in different sectors of a particular base station (even in the same region type). We consider a particular base station consisting of three sectors in each region and the corresponding geographical characteristics of each sector. Fig. 3a depicts the spatial distribution of signal strength measurements obtained from a channel scanner for a base station located in a single-family residential region. The measurement locations across the three sectors are represented by red dots. In Fig. 3b, we see that the path loss exponent of sector (a) to sector (c) ranges from 3.1 to 3.4, even from the same base station.

To generalize this behavior over multiple region types, Table 4 depicts the path loss exponent of three sectors of a particular base



(a) Measurements from a base station in the single family region.



(b) Regression line fit to the measurements of each sector.

Fig. 3. Signal quality data from three sectors around a base station (left) related path loss exponents of each (right).

Table 4
Field-tested path-loss exponent per cell from channel scanner (TSMW) and corresponding geographical features.

Region	Sector	γ	$\overline{B_h}$	B_{h_σ}	#B	$\overline{T_h}$	T_{h_σ}	#T	$\overline{R_{GE}}$	R_{GE_σ}
Single-family	$Sector_3$	3.6	8.6	2.7	330	10.7	3.2	2760	186	6
	$Sector_2$	3.5	8.3	2.8	270	10	2.9	2572	186	2
	$Sector_1$	3.2	8.34	2.7	190	10	2.9	1846	182	2
Multi-family	$Sector_1$	3.7	10.3	2.7	130	8.8	3	1117	182	3.5
	$Sector_2$	3.6	11	5	276	10.5	4.5	1400	180	2.5
	$Sector_3$	3.4	9.7	3	230	9.7	4.6	1960	184	2.5
Downtown	$Sector_1$	3.5	60	53	38	16	10	305	134	2.6
	$Sector_2$	3.2	43	30	56	12.5	9	300	133	3.6
	$Sector_3$	2.8	35	35	36	9.7	7.8	147	135	2

station in three different areas (downtown, single-family, and multi-family residential areas). The γ in the downtown area shows a higher variation (0.7) than two other regions with the multi-family and single-family areas having a value of 0.3 and 0.4, respectively. Furthermore, we can see that the small variations in average height of the objects in each region account for small changes in estimated γ because the taller objects around the receiver or in between the base station and the UE are more likely to scatter or reflect the signal. Also, in single- and multi-family areas, the number of objects located in a sector and near to the receiver has an impact on the path loss slope [39].

The γ in the downtown area shows a higher variation (0.7) than two other regions with the multi-family and single-family areas having a value of 0.3 and 0.4, respectively. We can see that the small variations in average height of the objects in each region accounts for small changes in estimated γ. Also, in single- and multi-family areas, the number of buildings and trees located in a sector has an impact on the path loss slope.

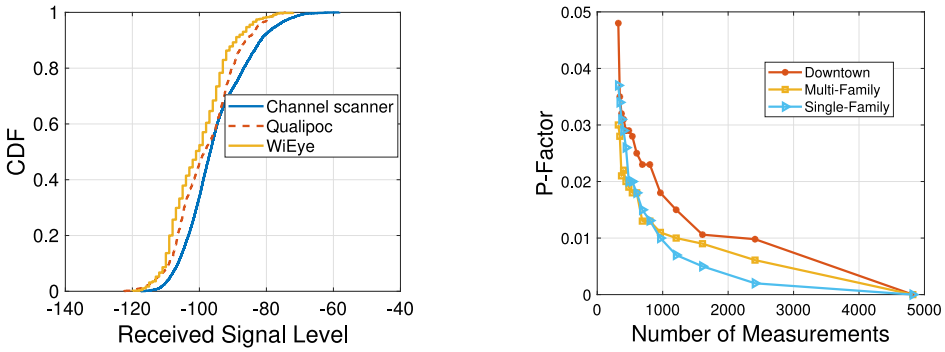
iii. RSRP samples form diverse statistical distributions based on device and region types. To depict the difference in received signal power between the channel scanner, firmware, and API level, we plot the distribution of the Reference Signal Received Power (RSRP) values obtained by each tool for a specific base station sector in Fig. 4a. We observe that the difference between the CDF's median of the channel scanner (−77 dBm), Qualipoc (−78.8 dBm), and WiEye (−79.5 dBm) are about 1.8 dB and 2.5 dB, respectively. This difference is similar to that discussed in Table 1, especially for the firmware measurements but shows that the API level samples are subject to other effects such as averaging of samples, which will be explored in greater depth in Section 4.3.1.

To evaluate the viability of a measurement sample size for each region, we use the Kolmogorov–Smirnov test (KS test), which attempts to determine if the samples come from the same distribution. There are two metrics with the test, h and p, which are the results of the hypothesis test at the default 5% significance level. In particular, h determines if the test is passed or failed, and p is the estimated significance for the specific test evaluated. An h flag will be reported as 0 (false) if the null hypothesis that the two distributions have a common distribution and cannot be rejected at the chosen significance level and concurrently does not have enough evidence to support the similarity.

In our application, if the number of measurements is too small to be representative of the region, the p value will be above a threshold of 0.05. Conversely, if $p \leq 0.05$, it signifies that there are a sufficient number of measurements to be confident in the path loss exponent prediction. In Fig. 4b, we show the p value of the KS test based on the measurement number per region type. In our case, when $p \leq 0.05$, then h is always 1, which means the test passes for these values of p. When the threshold is crossed, the number is approximately 600 samples for each region. Furthermore, the figure shows that decreasing the number of measurements in a single-family area has a lower impact on the KS value as compared to multi-family and downtown areas. This effect can be credited to the relative homogeneity of the geographical features in the single-family area as opposed to the more heterogeneous multi-family or downtown regions. We will extend this investigation on measurement number size in Section 4.3.2, where we focus on the role of downsampling in both time and space from a very large number of measurements taken by the channel scanner.

iv. Matching the mobile phone samples in time to the channel scanner samples provides precise path loss exponent prediction. We now focus on a single mobile phone (Samsung Galaxy S5) to directly compare the path loss exponent inferred from the received signal quality reported at the API and firmware levels to that reported by the channel scanner in the same environment. We consider the most densely-measured sector from each region type in our comparison and calculate three different path loss exponents. First, we consider the path loss exponent γ_X as calculated from all measurements in the chosen sector for device X, where X is T for TSMW, Q for Qualipoc, or W for WiEye. Second, we downsample the TSMW measurements according to the matching process mentioned in Section 3, where the TSMW measurement with the closest time stamp to the mobile phone measurement is chosen for Qualipoc and then for WiEye. This second calculated path loss exponent is represented by $\gamma_{Q'}$ and $\gamma_{W'}$, respectively, and allows the path loss exponent to be considered for the same number of measurements as Qualipoc and WiEye but with the signal strength readings from the TSMW. This approach inherently controls for the number of samples, which we later evaluate extensively.

These two γ values are shown in Table 5. By comparison across these path loss exponents, γ_T with $\gamma_{Q'}$ and $\gamma_{W'}$, we observe that even when the same device is used (TSMW) to capture the signal strength measurements, downsampling the number to match the mobile phones raises the estimate of the path loss exponent in every environment. This effect could be explained by the inclusion of lower quality measurements (*i.e.*, considering the measurements that were clipped from the mobile phone



(a) CDF of RSRP of channel scanner versus Qualipoc and WiEye.

(b) The impact of decreasing the number of samples on the KS value.

Fig. 4. Statistical distribution properties of RSRP based on device (a) and region type (b).

Table 5
Path loss exponents inferred from mobile phone signal quality reported at the firmware (Qualipoc) and API level (WiEye) from total measurements versus those matched closest in time to that of the channel scanner (TSMW).

Region	TSMW		Qualipoc				WiEye			
	Samples	γ_T	△Q&T (dB)	Samples	γ_Q	γ_Q'	△W&T (dB)	Samples	γ_W	γ_W'
Single-family	2 063	3.1	1.4	620	3.23	3.17	3.8	293	3.31	3.19
Multi-family	1 961	3.41	0.9	970	3.48	3.42	3.1	350	3.51	3.38
Downtown	11 634	3.85	1.2	512	3.97	3.87	3.5	225	4.00	3.89

measurements), which in turn lowers slightly increases the value of the path loss exponent. Despite the consistently positive path loss exponent error in prediction, the matched γ values of γ_Q' and γ_W' are within 0.06, 0.06, and 0.1 for the firmware level measurements and 0.12, 0.13, and 0.11 for the API-level measurements for the single-family, multi-family, and downtown regions, respectively. Therefore, matching the mobile phone measurements in time to the channel scanner measurements allowed highly-effective path loss exponent prediction, especially at the firmware level.

4.3. Mobile phone and crowdsourcing impact on path loss estimation

In this section we study the impact of different shortcomings with mobile phone measurements (averaging, temporal downsampling, and quantization) and imperfections that arise with crowdsourcing wireless signal strengths (non-uniform downsampling in both time and space) as opposed to drive testing in a known physical pattern with a known periodic sampling frequency in a particular region under test. In this Sections 4.3 and 4.4, we use signal strength samples from the channel scanner exclusively in our analysis and emulate each mobile phone imperfection in isolation to evaluate the impact of that effect.

4.3.1. Averaging of the received signal power

Network interfaces often use some form of hysteresis to suppress sudden fluctuations in channel state that might lead to overcompensation in adaptive protocols. Many times this hysteresis is performed by averaging multiple received signal qualities before reporting it to the higher layers (*e.g.*, within the firmware) and/or the user (*e.g.*, within the operating system in support of API calls). Each device uses its own policy (often proprietary) to take a specific number of samples over a certain period of time. In particular, a mobile phone in an LTE network is required to measure the Reference Signal Received Power (RSRP) and Reference Signal Received Quality (RSRQ) level of a serving cell at least every Discontinuous Reception (DRX) cycle to see if the cell selection criteria is satisfied [40]. To do so, a filter is applied on the RSRP and RSRQ of the serving cell to continually track the quality of the received signal. Within the set of measurements used for the filtering, two measurements shall be spaced by no longer half of DRX cycle [41]. On the other hand, a mobile phone receives multiple resource elements and measures the average power of resource elements. However, the number of resource elements in the considered measurement frequency and period over which measurements are taken to determine RSRP by the mobile phone depends on the manufacturer.

As a result, even if two devices are in the same environment in close proximity and experience virtually the same channel quality fluctuations, differences in averaging window sizes could be interpreted as diverse fading behaviors. More importantly, when crowdsourcing signal strengths, we are forced to accept the averaging behavior of a broad range of devices. Hence, there is a question as to the degree to which an MDT update should be filtered. Since we are focused on large-scale path loss in this paper, we assume that applying a filtering mechanism on the measurements to average out the effect of the fast fading avoids misinterpretation of extreme instantaneous behavior. In other words, the eNodeB does not want to misinterpret the channel condition due to uncharacteristic spurs in the measurement, which could lead to erroneous actions such as excessive handover.

Hence, we seek to empirically quantify the degree to which a range of averaging windows (*i.e.*, the number of samples used in the average reported) affects the calculation of the path loss exponent parameter. We depict the variation of the γ parameter in Fig. 5 when we vary the averaging window from 0.25 to 6 s on the collected measurements by the channel scanner, which corresponds to a window size of 0 to 200 samples. We averaged the RMSE corresponded to each window size over multiple base stations in each region. As we see, by increasing the filter size, the maximum error in three regions is about 0.1. In other words, we show that decreasing the window size does not improve the results dramatically.

Also, we represent the average and standard deviation of the estimated errors in γ estimation caused by averaging for each region via a box plot in Fig. 6. The x-axis shows a period of 6 s with a step size of one second. We see that the average error for the averaging window size varies between 1 to 6 s and is approximately 0.12 RMSE of the path loss exponent, on average, among the three regions. In addition, we observe that by increasing the averaging window size, the absolute error in all the three regions increases. However, the variation of the error decreases because of the fluctuations of the signal is flattened by applying a large averaging window size.

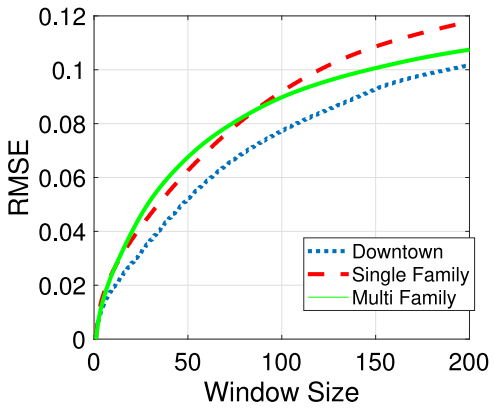


Fig. 5. Averaging impact on path loss exponent (γ) prediction.

4.3.2. Non-Continuous measurement periods

When crowdsourcing information from willing participants, we must be sensitive to their data usage and battery consumption issues, precluding prolonged, continuous measurements of detailed signal strength values. One option may be to uniformly reduce the number of samples per unit time for a given user over an extended period. Another option could be to aggregate small numbers of samples at different time periods and space from one or more users to compose an aggregate channel effect. We now study both the former (uniform downsampling) and latter (non-uniform downsampling).

Uniform downsampling impact. The channel scanner samples the channel quality at approximately 500 times per second as opposed to about 3 and 1 Hz with the Qualipoc and WiEye, respectively. In this scenario, as the mobile phone preserves energy and/or data usage the question becomes: how would the γ parameters further diverge from the results shown in Table 5? In other words, the previous result showed the extreme cases of either matching the same number of samples or having a very different number of samples.

To study the role of differing numbers of measurements on path loss estimation, we first examine the calculated γ parameter from a particular sector of a base station in each region, when using uniform and non-uniform down-sampling. We gradually reduce the number of samples obtained by channel scanner to eventually reach the same number of samples recorded by WiEye. At each step, we calculate the error in path loss exponent calculation with respect to our reference value, which is obtained by considering the highest resolution in channel scanner data set. To do so, we reduce the number of samples by i, where $i \in 1, \ldots, n$ and $n = \frac{\#Channel\ scanner\ records}{\#Phone\ records}$. As we reduce the data set by i samples, we are able to leverage i data sets for a given i to increase the confidence in the result and study the variation of error.

Fig. 7a shows the error in path loss estimates by reducing the signal samples received from a cell sector of a base station in the downtown area. By increasing the time interval between samples, the γ and resulting variation thereof are affected. We observe that the error caused by uniformly downsampling can reach up to 0.03 in this specific cell, which means the predicted value is very close to the reference γ. Although the RMSE over each 10 steps has some variation, it does not increase the error dramatically. Furthermore, by decreasing the number of samples, the variation of channel characteristic estimation is not as stable as when we have more data points.

Fig. 7b shows the impact of uniformly downsampling the channel characteristics on each of the three different regions (single-family residential, multi-family residential, and downtown). The maximum variation over all three regions is depicted as the variation of the RMSE at each point. Of particular note in this result is that downtown shows more sensitivity to downsampling, and the single-family residential region shows the least sensitivity.

Non-uniform downsampling impact. In a second scenario, perhaps the crowdsourced measurements are not coming from a single user which has uniformly throttled the number of measurements recorded or reported but from multiple users in the same area. Controlling for device differences for now (we will study this issue in Section 4.5), the newly composed data set for mobile phone measurements Y has a non-uniform sampling period in time and space compared to drive testing the region with a channel scanner. As before, we quantify the accuracy of the estimate of γ_Y to the estimated γ when mobile phone signal strength readings are dispersed through time and space. We assume that a sufficiently large number of users in a similar area have crowdsourced measurements. We also assume that the number of measurements from the non-uniformly sampled data set matches that of the uniformly sampled data set.

The non-uniform distributed measurements are studied in two domains: *(a)* temporal and *(b)* spatial. For non-uniform temporal downsampling, we reduce the number of samples randomly based on the time stamp of the received signal measurements from the channel scanner data set. Fig. 8a depicts the impact of the non-uniform temporal downsampling on the path loss exponent from a cell sector in downtown and shows that by increasing the number of samples, the error with respect to the reference value decreases. However, in general, the non-uniform temporal downsampling has caused a higher value in terms of RMSE for the same number of measurements as compared to uniform downsampling.

For non-uniform spatial downsampling, we select the most populated sector in each region. Then, we chose the measurements based on three clusters which are randomly distributed over the region. Then, we increase the number of the selected measurements in each cluster. Finally, we compare the γ of the aggregated samples from non-uniformly distributed clusters with the γ computed from all measurements from the channel scanner in the same region. A comparison between the uniform downsampling and non-uniform distributed measurements in space for three regions is depicted in Fig. 8b. The clustered scenario shows a higher error than the uniformly-distributed one. In addition,

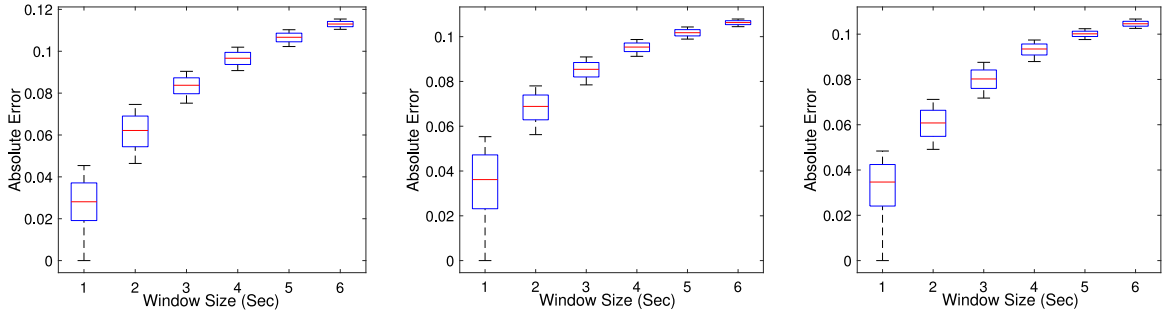
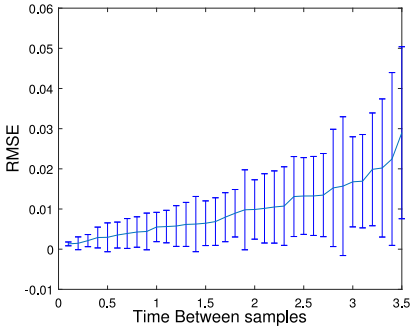
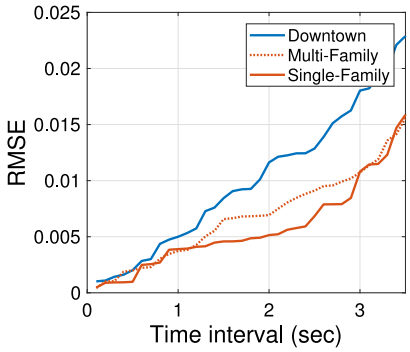


Fig. 6. Impact of averaging on the γ estimation in terms of mean and standard deviation in a period of 1 to 6 s.

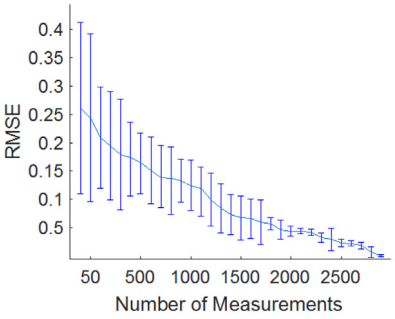


(a) Uniformly Downsampling (Downtown).

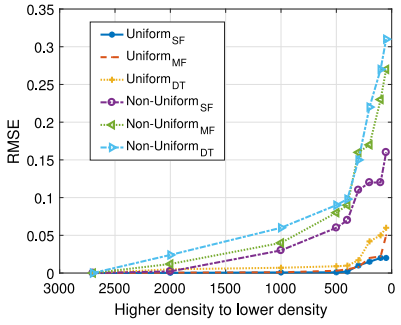


(b) Uniform Downsampling from a Sector in Three Regions.

Fig. 7. Uniformly downsampling the measurements of a sector in downtown (left) and across all three regions (right).



(a) Non-Uniformly downsampling selection in time.



(b) Non-Uniform (space) vs. Uniform (space).

Fig. 8. Non-uniform downsampling of a sector in downtown (left) and non-uniformly downsampling in space compared to uniformly (right)..

we observe that the error in the downtown region is higher than two other regions.

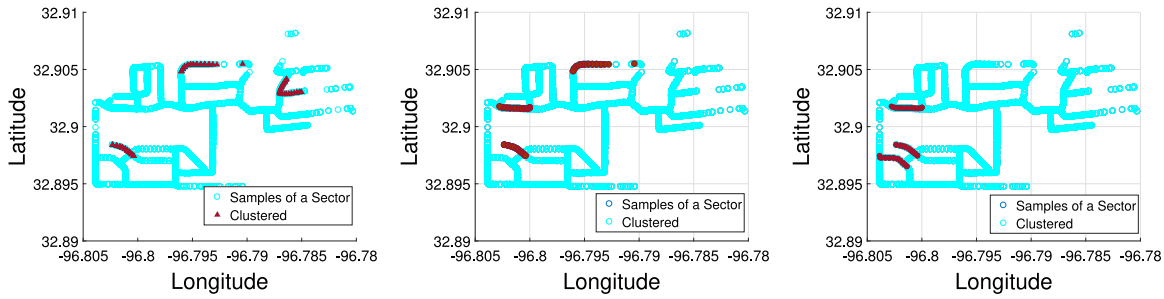
We have found that the location of the selected clusters in the non-uniform scenario is significant, as depicted in Fig. 9. To do so, we again select the most populated sector in a region. Then, we determine the location of three clusters of measurements based on Fig. 9. We start by choosing 50 measurements in each cluster and we increase it by 50 until we have 3000 measurements. The left figure shows a model which is more dispersed through a sector. The middle scenario covers the left and top left area of the sector. In the right scenario, all measurements have a grouping on the left of the sector. We measured the average of the RMSE for each scenario. The results show the 0.083, 0.15, and 0.3 as the average of the RMSE for each aforementioned scenario. In other words, a spatially well-distributed group of user measurements would contribute to a better result to predict the path loss exponent. Also, the type of cluster distribution has an impact on the number of measurements that are needed to estimate the channel condition. With this result and the current developments in the LTE standard (10) about the Minimization of Drive Test function [20], a carrier could more strategically poll users in a given area and/or at a certain time to reduce the resources necessary for their users to crowdsource and increase the likelihood of success of such an effort.

What is the required number of measurements? The number of measurements plays an important role in path loss prediction accuracy. Hence, we seek to find a sufficient number of measurements to provide a certain level of accuracy in channel characteristics prediction among three regions. To do so, we repeated the same procedure as before with our analysis with the following exception: we perform uniform downsampling, but his time, we consider more than one base station in each region and we select the sectors that contain the same number of signal measurements (about 4000 to 5000). We reduced the number of signal quality measurements and compared the path loss exponent results obtained from the new data set with the reference value. As we can see in Fig. 10, by decreasing the sampling size the averaged error is increased with greater fluctuations.

We depict three areas in each figure, where each area shows a certain level of accuracy in path loss prediction. The area on the far left of each graph shows the number of measurements that provides poor accuracy. The area towards the middle of each graph shows the range of the required number of measurements to obtain an acceptable error corresponding to the γ estimation. Finally, the area on the far right of each graph represents a range of measurements where the error is monotonically decreasing with each additional measurement providing improved accuracy. As depicted across all the graphs of Fig. 10, the required number of measurements to provide an accurate estimation of channel characteristics is between 700 and 1500. We explain in Section 5 that an error in path loss exponent estimation, result in overestimation or underestimation in the probability of coverage of a targeted region. Overestimating and underestimating in network performance prediction result in gaps in coverage area and redundancy or even unwanted self-interference within the same network deployment, respectively.

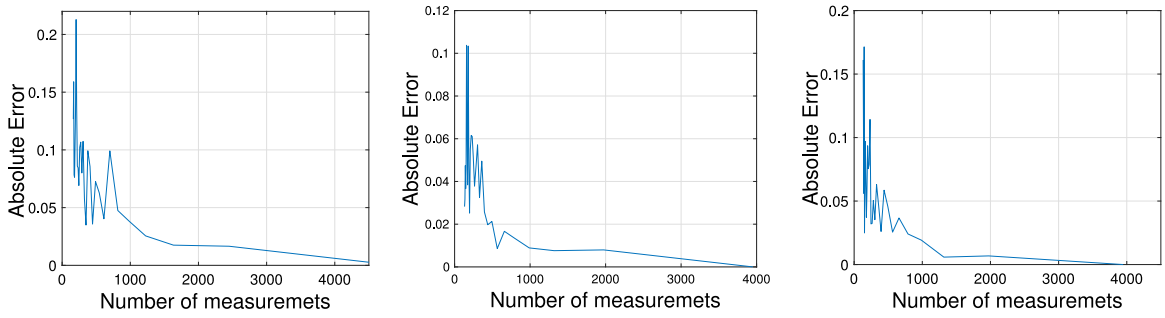
4.3.3. *Quantization of the received signal power*

Android reports the quality of the common pilot channel received signal quality for LTE in terms of Arbitrary Strength Units (ASU) with 98 quantized levels. The received signal level has a range of −44 dBm to −140 dBm and is mapped to "0 to 97" with the resolution of 1 dBm. Since the obtained signal strength by a channel scanner has much greater granularity, the question becomes: what role does quantization have on the path loss exponent? We have considered the quantization impact on path loss estimation as defined as the difference between the estimated γ compared to the highest resolution setting as measured by



(a) RMSE = 0.083 (b) RMSE = 0.15 (c) RMSE = 0.39

Fig. 9. Impact of spatial non-uniformly downsampling.



(a) Downtown (b) Multi-Family Residential (c) Single-Family Residential

Fig. 10. Required number of measurements in uniformly and non-uniformly downsampling case.

the channel scanner. To do so, we round each element of the received signal strength from the channel scanner to its upper bound or lower bound value. By comparing the result with the reference γ, we found the absolute error to be negligible (*e.g.*, 0.0003). We show this effect in Fig. 11a of the following subsection, which considers the joint effect of all of these imperfections.

4.4. Joint analysis of mobile phone factors on path loss

Up to this point, we applied each of the challenges with phone measurements individually. We now jointly consider the mobile phone imperfections impact (averaging, uniform and non-uniform downsampling in space and time, and quantization) on the γ estimation. To do so, we extract the collected data by the channel scanner obtained from a specific cell sector from three regions. Then, we apply the averaging on signal samples which are quantized already. Then, we downsampled (uniformly and non-uniformly in time and space) from the averaged and quantized values. At each step, we obtain the RMSE from the path loss exponent calculated from the channel scanner's samples with the highest resolution. Fig. 11a depicts the relative error caused by each shortcoming in comparison with the other studied issues for all 11 of our base stations. Fig. 11b shows the percentage of RMSE caused by each individual issue with respect to the reference γ for data from all base stations. Here, we observe that non-uniformly downsampling has a dramatic effect on the results. However, each base station has a diverse measurement number, which could contribute to these results. Hence, we analyze the impact of each imperfection with mobile phone measurements individually on a data set for a single sector in each region with a comparable number of measurements (4000 to 5000). As before, we applied averaging, uniform and non-uniform (spatial and temporal) downsampling, and quantization to the data. Fig. 12a shows the RMSE of the path loss prediction due to each effect as compared to the prediction with all measurements of that sector.

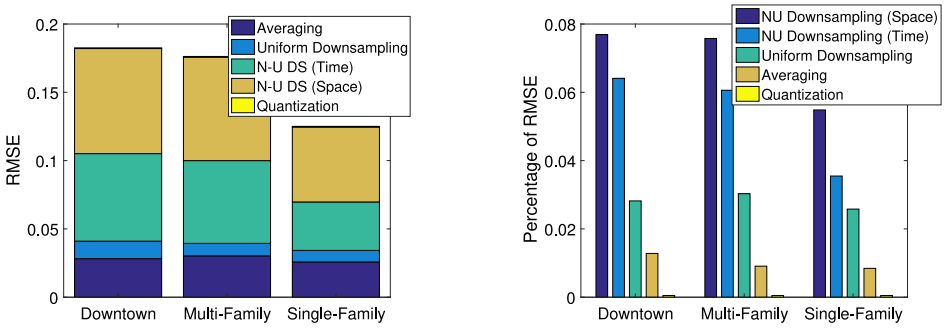
There are two interesting findings from these result: *(i)* either form of non-uniformly downsampling is clearly the most dominant effects considered when predicting the path loss exponent, and *(ii)* the two non-uniform downsampling techniques (time and space) have approximately equivalent performance (despite the noise of non-uniformly downsampling noted earlier). The latter finding offers great hope for crowdsourced data sets to be influential in characterizing the path loss characteristics of an environment.

4.5. Impact of heterogeneous mobile phones and users on path loss characterization

When crowdsourcing signal quality from mobile phone users, there is a diversity in hardware and software of the devices. Even two co-located mobile phones at the same time may report very different signal qualities due to different RF front ends. In this section, we study the impact of heterogeneous devices on the estimated path loss exponent. Up to this point, we have considered a single type of mobile phone, Samsung Galaxy 5S, due to its ability to support both Qualipoc and WiEye. Here, we use WiEye across three other mobile phones (4 total) with a two-phase approach. First, we consider the signal strength samples from all the devices to calculate the path loss exponent and evaluate the accuracy compared to the path loss exponent from the channel scanner signal quality samples. Second, we consider the differences in reported signal strength from each device introduced by each mobile phone in terms of dBm as compared to the raw measurements of the channel scanner. Lastly, we calculate the path loss exponent based on strictly crowdsourced data from WiEye users in different regions around the world and examine the geographical features of these areas.

4.5.1. Calibrating diverse phone models and setup

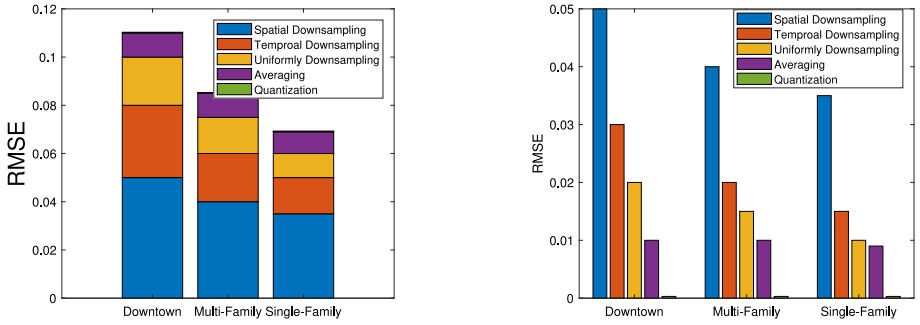
In this experiment, four Android phones described in Table 6 are used to collect signal strength data from the three aforementioned areas in Dallas (single-family residential, multi-family residential, and downtown). We installed our development version of WiEye, which logs signal strength samples at 1 Hz, on the following four phones: Samsung



(a) Relative impact of mobile phone factors on RMSE of γ.

(b) Impact of each factor on percentage of RMSE of γ.

Fig. 11. Joint impact of mobile phone imperfections relatively (left) and per effect (right).



(a) Impact of mobile phone factors on RMSE of γ.

(b) Impact of each factor on RMSE of γ prediction.

Fig. 12. Impact of each mobile phone imperfection.

Table 6
Measurement tools configuration and field-tested range of reported signal quality (dBm) from channel scanner (TSMW) and WiEye of four phones.

Tool	Model/OS	Chipset	Min	Max	Range
Channel scanner	TSMW/-	–	−130	−52	78
W_1	Samsung GS5/A5	MSM8974AC	−118	−54	64
W_2	Nexus 5X/A5	MSM8974	−119	−58	61
W_3	Google Pixel/A7	MSM8996	−120	−57	63
W_4	Samsung GS8/A7	MSM8996	−121	−54	67

GS5, Nexus 5X, Samsung S8, and Google Pixel. Each phone was co-located alongside the channel scanner on the roof of a car. The duration of the experiment was 360 min.

We first analyze the RSRP differences of the four phones in terms of the minimum, maximum, and resulting range of dBm reported across all measurements to understand the relative sensitivities. While a few hours of driving does not guarantee the full range of signal strengths, during this time, we observe that the greatest range of values is achieved by the Samsung S8 (67 dBm) as reported by WiEye and the least range of values belonged to the Nexus 5X (61 dBm). As a point of comparison, the TSMW Channel Scanner achieved a range of 78 dBm for the temporally-matched samples.

4.5.2. *Inferring path loss across devices*

We now use each phone to predict γ for four observed base stations in aforementioned regions. The dBm offset bias between the average received signal level by each phone and the channel scanner is shown in Table 7 per region.

We observe that on average the difference in reported received signal level by the scanner is 3 dBm higher versus the phones across the three regions with a range of 1.46 to 4.1 dBm. As we depicted before, the biases directly affect the path loss characterization. The lower reported channel quality corresponds to a higher value in obtained path loss exponent, while a higher reported channel quality corresponds to a lower path loss exponent. We now consider the calculated path loss exponent from the signal strength samples of each of the four phones, the calculated path loss exponent from the aggregated data set of the reported signal strength samples from all phones, and then the calculated γ from the compensated signal strength samples of all phones, considering the bias.

Table 8 shows the obtained path loss characteristics of one specific sector in three different regions, when we consider only a single phone's RSRP and all phones' RSRP. As a point of reference, we also include the γ from the channel scanner RSRP data. We observe that the obtained γ using the data set of each phone are relatively close to one another. We see that the Samsung S8 phone has the closest γ value between all four phones to the channel scanner. In other words, the device that receives the larger range is more accurate in terms of the γ estimation. The comparison shows that considering all RSRP data across device types actually increases the accuracy as compared to any given phone against the path loss exponent calculated from the channel scanner RSRP. Hence, we find that γ is predicted by using the RSRP from a *diverse* set of mobile phones. In addition, we compensated the signal strength of the aggregated data set by using the 3 dBm obtained in the previous section. We find that the compensated results in terms of γ are extremely close (with 2.7%, 0.3%, and 0.6% error for single-family residential, multi-family residential, and downtown, respectively) to the obtained results by the channel scanner.

4.5.3. *Inferring the path loss from crowdsourcing*

We now use crowdsourced measurements taken from our widely-distributed WiEye application on the Google Play store. We estimate the

Table 7
Average signal quality bias reported from heterogeneous phones as reported by WiEye with matched channel scanner measurement.

Device	W_1 (GS5)	W_2 (N5X)	W_3 (Pixel)	W_4 (GS8)
Location	dBm Diff. (Mean)	dBm Diff. (Mean)	dBm Diff. (Mean)	dBm Diff. (Mean)
Downtown	4.4 (−78.5)	2.1 (−76.2)	3.2 (−77.3)	1.6 (−75.7)
Single family	3.8 (−85.0)	2.4 (−83.6)	2.5 (−83.7)	1.7 (−82.9)
Multi family	4.1 (−80.3)	2.7 (−79.1)	3.5 (−80)	1.1 (−77.6)

Table 8
Path loss characteristics obtained by four devices in three modes: matched, aggregated, compensated mode.

Device	Single family	Multi-family	Downtown
Channel scanner	3.01	3.33	3.61
W_1 (GS5)	3.21	3.50	3.80
W_2 (N5X)	3.18	3.54	3.78
W_3 (Pixel)	3.38	3.58	3.90
W_4 (GS8)	3.19	3.47	3.75
Aggregated	3.27	3.53	3.83
Compensated	3.09	3.34	3.63

path loss exponent of regions around the world without physically drive testing those areas. Based on some of our highest user density, we have selected four environments with diverse geographical features: (i) tall buildings and trees in Dresden, Germany, (ii) low buildings and no trees in Artesia, New Mexico, (iii) mostly trees with a few homes in Macon, Georgia, and (iv) mostly free space in Thiersheim, Germany. The aerial view of each of these environments can be seen in the top figures of Fig. 13.

In Fig. 13, the bottom figures show the number of crowdsourced signal strength samples and their spatial location as captured by our Android application overlayed on a more basic map of the same area displayed in the aerial view on the top. Using these signal quality measurements from each region, we have computed the path loss exponent γ, which can be seen in the caption of each subfigure. We have ordered the figures from left to right where we see the path loss exponent is decreasing from left to right. In particular, γ_a equals 3.3 with the most diverse and complex environment with tall buildings and trees, γ_b equals 2.7 with an environment that has similar, small building types but no trees, γ_c equals 2.5 with mostly trees and a few homes, and γ_d equals 2.1 with mostly free space.

Therefore, the geographical features and complexity in the environment match the γ behavior we would expect, and the channel factors were derived strictly using crowdsourced measurements. Of particular note that in these measurements alone we saw a fairly dramatic change in the γ. In fact, we observed a range of 2.1 to 4.0 of the path loss exponent throughout this paper, which would constitute extremely different network designs across this range of propagation scenarios.

We find that a well-distributed signal measurement throughout a region would provide an accurate γ. Yet, we seek to achieve an acceptable accuracy level with less total measurements. In this experiment, we considered signal quality measurements obtained from various locations within a sector corresponding to a base station. The minimum and maximum distances from captured measurements regarding to a base station are 84 m and 2.5 km respectively. We divide the signal quality measurements into 3 groups for a particular sector based on distance: near, middle, and far (Fig. 14). We then perform the analysis on all combinations of two different regions of the three available, studying the role of distance away from the base station and its impact on our path loss prediction.

Table 9 shows the path loss slope for each cluster. As we expect, the variation of the γ over a homogeneous region is less than a heterogeneous environment due to the difference between the geographical characteristics of each region. In addition, we aggregate signal measurements from different clusters and compare the results with our reference path loss slope. The results show that aggregating the signal measurements from the near and edge region results in a path loss exponent closest to our reference γ.

In next step, we apply the same approach on our crowdsourced data set, depicted in Fig. 13. We select the case with a slope $\gamma_a = 3.3$ because of its complex environment with tall buildings and trees. The results show that the aggregated data sets from near and far regions ($\gamma = 3.29$) are the closest result to the reference slope ($\gamma = 3.3$).

5. Coverage estimation impact from prediction error

In the previous sections, we evaluated the path loss exponent prediction accuracy using RMSE and the total difference in the γ value. However, it is hard to interpret such error in terms of operational network performance. To build such intuition, we examine the role of this prediction error on coverage estimation and Bit Error Rate (BER). Since the received signal level randomly varies due to shadowing effects, network operators must determine the probability that the received signal strength crosses the specified threshold at the cell edge, which is calculated according to [28]:

$$P_{x_0(R)} = Prob[x > x_0] = \int_{x_0}^{\inf} P(x)dx = \frac{1}{2} - \frac{1}{2}erf\left(\frac{x_0 - \bar{x}}{\sigma\sqrt{2}}\right) \quad (1)$$

Here, x_0, is the receiver sensitivity, which is a function of the UE hardware design and the required service quality. x is the receive signal level at distance d, which can be obtained by applying a log-distance propagation model. It is a common approach to estimate the mean of the received signal level over a specific distance d from the base station. The variation of the received signal level due to shadowing, represented here by σ.

Knowing $Prob[x > x_0]$, we can determine the percentage that the received signal level exceeds a certain threshold in an area with radius R around the center of a base station, as shown in [42]:

$$F_u^{\gamma} = \frac{1}{\pi R^2}\int P(x_0)dA \quad (2)$$

The simplified version of the previous equation could lead to:

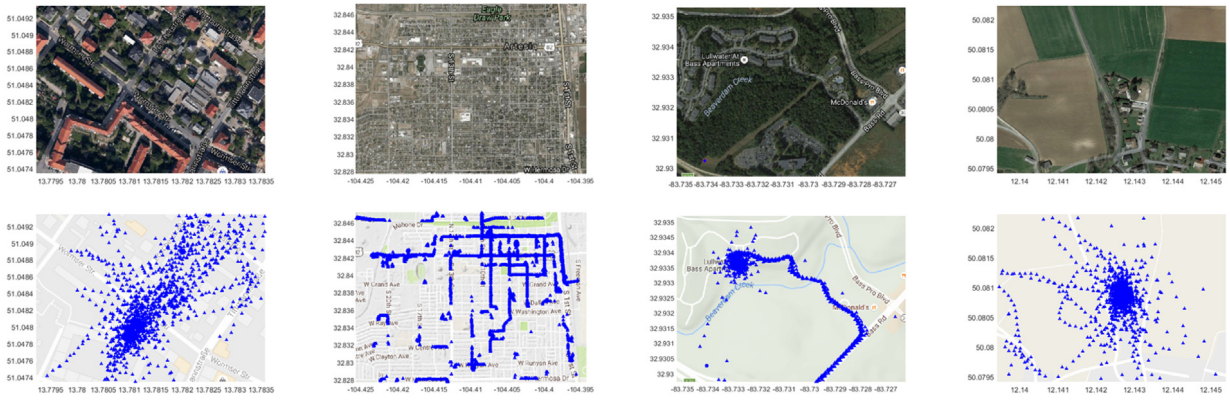
$$F_u^{\gamma} = \frac{1}{2}\left[1 - \mathrm{erf}(a) + \exp\left(\frac{1-2ab}{b^2}\right)\left(1 - \mathrm{erf}\,\frac{1-ab}{b}\right)\right] \quad (3)$$

Here, a and b are obtained from (4):

$$a = \frac{x_0 - \alpha}{\sqrt{2}\sigma}, \qquad b = \frac{10\gamma\log_{10}(e)}{\sqrt{2}\sigma} \quad (4)$$

where α is determined from the transmitter power, antenna heights and gains. To investigate the impact of the path loss error prediction on the probability of cell area coverage estimation, we assume that the mobile network provider has a restriction on the receiver sensitivity x_0 at a certain distance R to provide a particular service level. With respect to the above assumptions, we achieve the probability of exceeding the sensitivity level x_0 with probability $Prob_{x_0} = prob[x_R > x_0]$. Knowing this, we obtain the percentage of the useful area covered with a cell boundary of R and the received signal level x_0.

We consider the case where the actual path loss exponent is 3.0. Then, we consider a error in path loss exponent value from −1.0 to +1.0 for a range between 2.0 and 4.0 for the predicted value. We assume the receiver sensitivity is fixed due while providing a particular service level at the cell edge. Then, we obtain F_u corresponding to each γ,



(a) Location: Dresden, Germany with γ_a=3.3.

(b) Location: Artesia, New Mexico with γ_b=2.7.

(c) Location: Macon, Georgia with γ_c=2.5.

(d) Location: Thiersheim, Germany with γ_d=2.1.

Fig. 13. Path loss analysis for crowdsourced data sets in four different regions.

(a) Close Measurements

(b) Measurements at the middle

(c) Farthest Measurements

Fig. 14. Variation of γ in homogeneous region (single-family residential) at varying distance from base station.

Table 9
Field-tested path-loss exponent per cell from channel scanner (TSMW) and corresponding geographical features.

Region	Ref_γ	$NearCell_\gamma$	$MiddleCell_\gamma$	$EdgeCell_\gamma$	$Average_\gamma$	$Agg_{\gamma_{1,3}}$	$Agg_{\gamma_{1,2}}$	$Agg_{\gamma_{2,3}}$
Single-family	3.55	3.58	3.59	3.5	3.55	3.52	3.6	3.61
Downtown	3.57	3.63	3.54	3.46	3.54	3.56	3.6	3.5
Dresden (Germany)	3.3	3.4	3.34	3.3	3.34	3.29	3.37	3.27

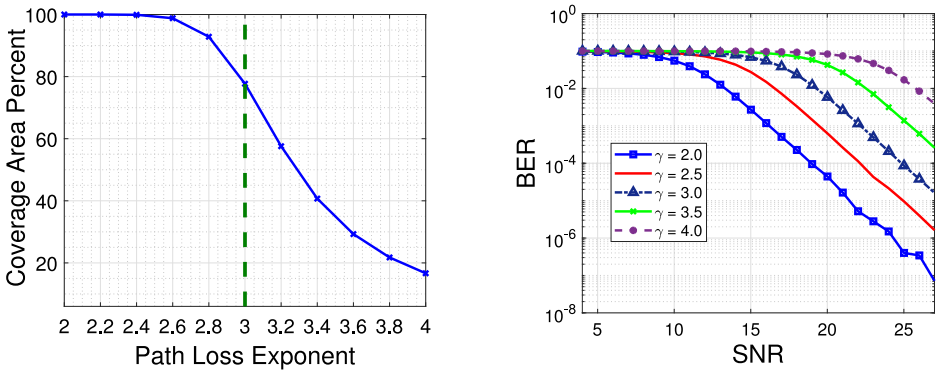
and we compare the estimated probability of coverage over the cell area with the reference γ. Fig. 15a shows the impact of the error in path loss exponent prediction on the probability of cell area coverage estimation when we are overestimating and underestimating the path loss exponent.

As we can see, in the case that a greater path loss exponent than actual is predicted ($\gamma \geq 3.0$), the coverage area probability drops from 79% to 20% by increasing the predicted γ from 3.0 to 4.0. In this case, with a fixed transmission power, the network operator would cover 59% more than expected, thereby creating unwanted redundancy and self interference. In the case that a lower path loss exponent than actual is predicted ($\gamma \leq 3.0$), the coverage area probability rises from 79% to 99% as early as 2.6 and remains at 100% to 2.0. While this might seem like a positive effect for network operators, it could be an even greater problem. Namely, the network operator will think that the propagation environment is better than actual and so increase the spacing between nodes, thereby creating coverage holes. For example, in the environment where an actual path loss exponent is 3.0 and the predicted path loss exponent is 2.6, there will be around 20% of the network that is not covered from the targeted area.

Lastly, we study the impact of the error in path loss exponent prediction on the network performance in terms of BER. Fig. 15b shows the variation of BER by changing the predicted path loss exponent from 2.0 to 4.0 with a step of 0.5 while the SNR is in a range of 8 to 24. We compare the BER of the estimated path loss exponents with our reference one ($\gamma = 3.0$). We observe that an error about 0.5 in path loss exponent prediction causes an order of magnitude change in the BER at an SNR of 20.

6. Conclusion

In this work, we take a first step towards crowdsourcing wireless channel characteristics in LTE cellular networks (and later generations of cellular technology) by considering the relationship between received signal strength measurements of diverse mobile phones at the firmware and API level versus advanced drive testing equipment. In particular, we performed extensive experimentation across four mobile phone types, two pieces of software, and a channel scanner in three representative geographical regions: single-family residential, multi-family residential, and downtown. With these devices and in-field measurements, we evaluated the effects of averaging over multiple samples, uniform and non-uniform downsampling (in time and space), quantization, and crowdsourcing on the path loss exponent estimation. We showed that both types of non-uniform downsampling have the most dramatic effects on path loss calculation. Conversely, we showed the quantization impact can largely be ignored since it showed a negligible influence on our estimation. One key result of note stems from the spatial non-uniformity of clusters of measurements observed within our crowdsourcing database, which required far more measurements than more uniformly spaced measurements. Also, we addressed the required number of measurements to have a sufficient understanding about the average of the signal attenuation in a specific environment.



(a) Underestimating the coverage area probability.

(b) BER estimation regarding to γ variations.

Fig. 15. Impact of the γ estimation's error on the cell area coverage probability estimation.

Using the MDT specification of LTE release 11, carriers could request specific measurement locations and times from users to be far more efficient in polling signal quality. Furthermore, we showed four regions around the globe and predicted the channel characteristics of these regions from our crowdsourced data. In summary, we lay a strong foundation for intuitively understanding a large majority of the issues involved with crowdsourcing channel characteristics. For example, we found that even a prediction error of 0.4 for the path loss exponent would cause a 40% redundancy in the covered area or coverage holes for 25% of the targeted area based on whether the error was above or below the actual value, respectively. In summary, we lay a strong foundation for understanding a large majority of the issues involved with crowdsourcing channel characteristics. In future work, we will study the impact of the various contexts on the received signal quality and path loss estimation to precisely characterize the role of each geographical feature on the large- and small-scale fading effects.

Acknowledgments

This work was in part supported by National Science Foundation (NSF) grants: CNS-1150215, CNS-1320442, and CNS-1526269. Also, we would like to thank Rhode & Schwarz for their extensive support in this measurement campaign.

References

[1] Utkarsh Goel, Mike P Wittie, Kimberly C Claffy, Andrew Le, Survey of end-to-end mobile network measurement testbeds, tools, and services, IEEE Commun. Surv. Tutor. 18 (1) (2016) 105–123.

[2] SeungJune Yi, SungDuck Chun, YoungDae Lee, SungJun Park, SungHoon Jung, Radio Protocols for LTE and LTE-advanced, John Wiley & Sons, 2012.

[3] Rohde & Schwarz GmbH & Co.KG. [n. d.]. Radio network analyzer operating manual.

[4] Universal Terrestrial Radio Access and Evolved Universal Terrestrial Radio Access (E-UTRA); Radio measurement collection for Minimaztion of Drive Test (MDT), 3GPP, Technical Report, Technical Specification, Technical report TS 37.320, 2014.

[5] Wuri A Hapsari, Anil Umesh, Mikio Iwamura, Malgorzata Tomala, Bódog Gyula, Benoist Sebire, Minimization of drive tests solution in 3GPP, IEEE Commun. Mag. 50 (6) (2012) 28–36.

[6] Alessandro Checco, Carlo Lancia, Douglas J Leith, Using crowdsourcing for local topology discovery in wireless networks, arXiv preprint arXiv:1401.1551, 2014.

[7] Jinghao Shi, Zhangyu Guan, Chunming Qiao, Tommaso Melodia, Dimitrios Koutsonikolas, Geoffrey Challen, Crowdsourcing access network spectrum allocation using smartphones, in: Proc. of ACM Hot Topics in Networks, 2014.

[8] Yuval Shavitt, Eran Shir, DIMES: let the internet measure itself, ACM SIGCOMM CCR 35 (5) (2005) 71–74.

[9] Zachary S Bischof, John S Otto, Mario A Sánchez, John P Rula, David R Choffnes, Fabián E Bustamante, Crowdsourcing ISP characterization to the network edge, in: Proc. of ACM SIGCOMM Measurements Up the Stack, 2011.

[10] Anshul Rai, Krishna Kant Chintalapudi, Venkata N Padmanabhan, Rijurekha Sen, Zee: zero-effort crowdsourcing for indoor localization, in: Proc. of ACM MobiCom, 2012.

[11] Tobias Hoßfeld, Michael Seufert, Matthias Hirth, Thomas Zinner, Phuoc Tran-Gia, Raimund Schatz, Quantification of YouTube QoE via crowdsourcing, in: Proc. of IEEE Multimedia (ISM), 2011.

[12] Marcel Dischinger, Massimiliano Marcon, Saikat Guha, P Krishna Gummadi, Ratul Mahajan, Stefan Saroiu, Glasnost: enabling end users to detect traffic differentiation, in: Proc. of USENIX NSDI, 2010.

[13] Sebastian Sonntag, Jukka Manner, Lennart Schulte, Netradar-Measuring the wireless world, in: Modeling & Optimization in Mobile, Ad Hoc & Wireless Networks (WiOpt), 2013 11th International Symposium on, IEEE, 2013, pp. 29–34.

[14] Sanae Rosen, Sung-ju Lee, Jeongkeun Lee, Paul Congdon, Z Morley Mao, Ken Burden, MCNet: Crowdsourcing wireless performance measurements through the eyes of mobile devices, IEEE Commun. Mag. 52 (10) (2014) 86–91.

[15] Aaron Gember, Aditya Akella, Jeffrey Pang, Alexander Varshavsky, Ramon Caceres, Obtaining incontext measurements of cellular network performance, in: Proceedings of the 2012 ACM conference on Internet measurement conference, ACM, 2012, pp. 287–300.

[16] Jongwon Yoon, Sayandeep Sen, Joshua Hare, Suman Banerjee, WiScape: A framework for measuring the performance of wide-area wireless networks, IEEE Trans. Mob. Comput. 14 (8) (2015) 1751–1764.

[17] Joe Cainey, Brendan Gill, Samuel Johnston, James Robinson, Sam Westwood, Modelling download throughput of LTE networks, in: Local Computer Networks Workshops (LCN Workshops), 2014 IEEE 39th Conference on, IEEE, 2014, pp. 623–628.

[18] Mads Lauridsen, Ignacio Rodriguez, Lucas Chavarria Gimenez, Preben Mogensen, et al., Verification of 3G and 4G received power measurements in a crowdsourcing Android app, in: Wireless Communications and Networking Conference (WCNC), 2016 IEEE, IEEE, 2016, pp. 1–6.

[19] William C Jakes, Donald C Cox, Microwave Mobile Communications, Wiley-IEEE Press, 1994.

[20] 3GPP, ETSI TS 137 320 "Radio measurement collection for minimization of drive Tests (MDT)", 2011.

[21] Ashkan Nikravesh, Hongyi Yao, Shichang Xu, David Choffnes, Z Morley Mao, Mobilyzer: An open platform for controllable mobile network measurements, in: Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services, ACM, 2015, pp. 389–404.

[22] Nicolas Haderer, Fawaz Paraiso, Christophe Ribeiro, Philippe Merle, Romain Rouvoy, Lionel Seinturier, A cloud-based infrastructure for crowdsourcing data from mobile devices, in: Crowdsourcing, Springer, 2015, pp. 243–265.

[23] Jaymin D Mankowitz, Andrew J Paverd, Mobile device-based cellular network coverage analysis using crowd sourcing, in: EUROCON-International Conference on Computer as a Tool (EUROCON), 2011 IEEE, IEEE, 2011, pp. 1–6.

[24] Massimiliano Molinari, Mah-Rukh Fida, Mahesh K Marina, Antonio Pescape, Spatial interpolation based cellular coverage prediction with crowdsourced measurements, in: Proceedings of the 2015 ACM SIGCOMM Workshop on Crowdsourcing and Crowdsharing of Big (Internet) Data, ACM, 2015, pp. 33–38.

[25] Andra Lutu, Yuba Raj Siwakoti, Ozgü Alay, Dziugas Baltrunas, Ahmed Elmokashfi, Profiling mobile broadband coverage, 2016.

[26] Berna Sayrac, Janne Riihijärvi, Petri Mähönen, Sana Ben Jemaa, Eric Moulines, Sébastien Grimoud, Improving coverage estimation for cellular networks with spatial bayesian prediction based on measurements, in: Proceedings of the 2012 ACM SIGCOMM workshop on Cellular Networks: Operations, Challenges, and Future Design, ACM, 2012, pp. 43–48.

[27] Michaela Neuland, Thomas Kurner, Mehdi Amirijoo, Influence of different factors on X-map estimation in LTE, in: Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd, IEEE, 2011, pp. 1–5.

[28] Christine Jardak, Petri Mahonen, Janne Riihijarvi, Spatial big data and wireless networks: experiences, applications, and research challenges, IEEE Netw. 28 (4) (2014) 26–31.

[29] Omer Faruk Celebi, Engin Zeydan, Omer Faruk Kurt, Omer Dedeoglu, Omer Iieri, Burak AykutSungur, Ahmet Akan, Salih Ergut, On use of big data for enhancing network coverage analysis, in: Telecommunications (ICT), 2013 20th International Conference on, IEEE, 2013, pp. 1–5.
[30] Theodore S Rappaport, et al., Wireless Communications: Principles and Practice, Vol. 2, Prentice Hall, 1996.
[31] Masaharu Hata, Empirical formula for propagation loss in land mobile radio services, IEEE Trans. Vehicular Technol. 29 (3) (1980) 317–325.
[32] Yoshihisa Okumura, Eiji Ohmori, Tomihiko Kawano, Kaneharu Fukuda, Field strength and its variability in VHF and UHF land-mobile radio service, Rev. Elec. Commun. Lab 16 (9) (1968) 825–873.
[33] Fumio Ikegami, Susumu Yoshida, Analysis of multipath propagation structure in urban mobile radio environments, IEEE Trans. Antennas Propag. 28 (4) (1980) 531–537.
[34] Joram Walfisch, Henry L Bertoni, A theoretical model of UHF propagation in urban environments, IEEE Trans. Antennas Propag. 36 (12) (1988) 1788–1796.
[35] Gerald K Chan, Propagation and coverage prediction for cellular radio systems, IEEE Trans. Vehicular Technol. 40 (4) (1991) 665–670.
[36] U Dersch, WR Braun, A physical radio channel model, Technical Report. IEEE CH2944-7/91/0000/0289, 1991.
[37] R. Enami, D. Rajan, J. Camp, RAIK: Regional analysis with geodata and crowdsourcing to infer key performance indicators, in: Proceedings of IEEE WCNC, 2017.
[38] Y. Shi, R. Enami, J. Wensowitch, J. Camp, Measurement-based characterization of LOS and NLOS drone-to-ground channels, in: 2018 IEEE Wireless Communications and Networking Conference (WCNC), 2018, pp. 1–6.
[39] Erik ÖstlinOstlin, Hajime Suzuki, Hans-Jürgen Zepernick, Evaluation of the propagation model recommendation ITU-R P. 1546 for mobile services in rural Australia, IEEE Trans. Veh. Technol. 57 (1) (2008) 38–51.
[40] 3GPP, ETSI TS 136 304 "Evolved universal terrestrial radio access (E-UTRA); User Equipment (UE) procedures in idle mode", 2011.
[41] 3GPP, ETSI TS 136 133 "Evolved universal terrestrial radio access (E-UTRA); Requirements for support of radio resource management ", 2011.
[42] Pete Bernardin, Meng F Yee, Thomas Ellis, Cell radius inaccuracy: a new measure of coverage reliability, IEEE Trans. Veh. Technol. 47 (4) (1998) 1215–1226.