GeoRIPE: Efficiently Harvesting Field Measurements for Map-Based Path Loss Modeling

Matthew Tonnemacher
Southern Methodist University
Dallas, Texas
mtonnemach@smu.edu

Dinesh Rajan
Southern Methodist University
Dallas, Texas
rajand@smu.edu

Joseph Camp
Southern Methodist University
Dallas, Texas
camp@smu.edu

ABSTRACT

Ensuring cellular coverage is an important and costly concern for carriers due to the expense of in-field experimentation (i.e., drive testing). With the ubiquity of smartphones, apps, and social media, there has been an explosion of crowdsourcing to understand a vast array of trends and topics at a minimal cost to the organization. While cellular carriers might seek to replace the expensive act of drive testing with the nearly cost-free crowdsourcing, questions remain as to: *(i)* the accuracy of crowdsourcing, considering the lack of user control, *(ii)* the detection of when drive testing might still be required, and *(iii)* the quantification of how many additional in-field measurements to perform for a certain accuracy level. In this work, we use geographical features of a region to reduce in-field propagation experimentation by predicting the number of measurements required to accurately characterize its path loss. In particular, we study the path loss prediction accuracy of drive testing and crowdsourcing by taking millions of measurements in a suburban and downtown region. We then use statistical learning to build a relationship between these geographical features and the measurements required. In doing so, we find that the number of measurements collected to achieve a certain path loss accuracy over the entire region can be reduced by up to 58% in a high density drive testing scenario.

1 INTRODUCTION

To address multi-fold increases in cellular demand, carrier cell sizes are shifting downwards to maximize network capacity. In doing so, the accurate and fine-grained coverage estimation of coverage becomes a critical issue for spatial reuse, intercell interference, and smooth handoffs between cells. Historically, in-field experimentation (i.e., drive testing) has been used to estimate the cellular propagation of a given region, which is costly for network operators due to the manpower and equipment required. The emergence of smartphones and their apps have offered a far cheaper alternative, recording in-field network observations directly from the cellular users (i.e., crowdsourcing). Crowdsourcing can allow network observations to be recorded in areas to which in-field testers may not have access.

Despite the availability of crowdsourced measurements, network providers continue to use extensive drive-testing to validate network coverage and quality of service metrics. One of the main goals in supplementing crowdsourced measurements with drive testing measurements is to fill in the gaps of crowdsourcing. However, the drive testing process can be costly and time consuming, creating a market for network drive testing reports. Thus, collecting these measurements in an efficient manner is a high priority. The most direct approach would be to minimize the number of measurements that needs to be collected to achieve the accuracy requirements set. In this paper, we attempt to reduce the number of measurements required for accurate path loss characterization throughout a given region by understanding and exploiting the variation in geographical features. In order for this to happen, several questions need to be answered:

(1) What are the relationships between geographical features and signal strength measurement requirements?
(2) Can the number of measurements required to achieve a particular accuracy level be determined for a specific area?
(3) How can these spatial differences be exploited to reduce the total number of measurements required to meet a particular regional path loss characterization accuracy?

In this work, we use geographical features of a region to more efficiently collect signal strength measurements, thereby reducing the amount of time spent on in-field propagation experimentation. To do this, we first introduce GeoRIPE, a statistical learning framework to situationally predict the number of measurements required to meet a specified path loss characterization precision. With this framework, geographical feature distribution input is used to suggest measurement collection requirements in a grid-like fashion over the target region. Then, we developed and deployed an Android application to gather signal strength measurements from real users throughout the world. We use a specially modified version of this application to collect high-density drive testing measurements from two distinct region types in a major metropolitan area. By using commercially available smartphones, measurements gathered are comparable to those gathered using purely crowdsourcing. Next, we explore the effect of land use on path loss characterization, showing how geographical feature diversity plays a large role in determining regional measurement requirements. We show especially strong correlation between the number of measurements required to accurately characterize the path loss in a region and the geographical feature ratio of small, medium, and large buildings, foliage and free space in an area. Finally, we validate the framework by

Figure 1: GeoRIPE Framework: using geographical features of a region (left) to infer the number of measurements required to predict path loss with a given accuracy level. The middle figure shows minimum measurements required (light is high, dark is low) of the same spatial distribution as the left-most figure. The right figure is a 3-D version of the middle figure.

comparing GeoRIPE to uniform measurement collection approach. We show that when fixing the accuracy requirement in path loss evaluation over the entire region, using the GeoRIPE framework can significantly reduce the required number of measurements that need to be taken to meet it.

2 GEORIPE FRAMEWORK

To illustrate the GeoRIPE[1] framework of using geographical features to reduce in-field propagation experimentation, we have depicted an aerial view of a region's terrain in the left-most image in Figure 1 with north on the top of the image. The terrain is classified into the following geographical features: buildings, dense foliage, and free space. Since in-field testing is expensive, our goal is to predict the least number of propagation measurements required to characterize the path loss in the region according to a specified level of accuracy. For example, the path loss of a region which has entirely free space (e.g., a desert) could be characterized with very few measurements. However, a diverse metropolitan region would require far more measurements to characterize.

The middle and right images of Figure 1 depict the end result of the GeoRIPE framework. The middle image depicts a two-dimensional overlay of the measurement density required to characterize the region shown in the left-most image. The southwest portion of the region is the lightest color, which means that it requires the greatest number of measurements to characterize due to the high concentration of buildings. In contrast, the northwest portion of the region is the darkest, which means that it requires the least number of measurements to characterize due to sparse building placement and less foliage. A three-dimensional version of the same figure can be seen in the right-most image to show the quantity of measurements required in each portion of the region.

A similar analysis could be done for a given region by an excessive amount of in-field testing and finding when each portion of the terrain converged to a particular level of path loss prediction accuracy. However, such an approach, by definition precludes any in-field testing reduction. While we take a very large number of measurements in certain regions in our work, we do so to train a statistical learning decision structure to infer the number of in-field measurements required. As a result, for any mix of such terrain

features, we can avoid the two in-field testing extremes of: *(i)* too few measurements, resulting in an inaccurate path loss estimate, or *(ii)* too many measurements, resulting in excessive experimentation costs. We can then evaluate the viability of using crowdsourcing to lower the drive testing cost.

The GeoRIPE framework's measurement distribution prediction is made with path loss accuracy in mind. So, before we evaluate the framework itself, we first need to give some background on path loss models in general as well as what model we use for our analysis. Path loss models attempt to predict the electromagnetic propagation as a function of distance. Many of these models rely on *a priori* information, using environmental details, a theoretical foundation, empirical findings, or some combination of the three for their prediction [6, 11, 21]. Other techniques operate under the assumption that *a priori* information is insufficient. These models supplement an existing model with a correction factor or factors based on measurements collected throughout a region to be modeled and tend to be more accurate than their *a priori* counterparts [17]. These active measurement models consist of two fundamental components: *(i)* how the measurements are gathered, and *(ii)* how they are incorporated into the model.

2.1 Path Loss Measurement and Supplemented Models

W. C. Lee studied the initial theoretical methodology of gathering active measurements for modeling path loss [13]. Lee proposed arced measurements at incremental distances from the transmitter while averaging measurements that fall within 20 to 40 wavelengths of each other, a claim corroborated by Shin using IEEE 802.11b measurements some years later [20]. In practice, it is often difficult to collect measurements strictly following the theoretically ideal guidelines due to environmental inaccessibility. This can be due to permission limitations, such as access restricted buildings or construction sites, or infrastructure limitations, where equipment setups are subject to the same mobility freedoms as the vehicles in which they operate. With a crowdsourced approach, a greater access diversity can be achieved with the limitations of a lack of control over data validity and input distribution. Due to these practical considerations, our work considers geographical complexity and its role in characterizing a region, both with vehicle-based drive testing and app-enabled crowdsourcing.

---

[1]As a mnemonic for this work, consider that fruit should be in the field the appropriate amount of time before harvested (i.e., to be *ripe*). Similarly, we seek to find the minimum amount of time necessary for in-field experimentation to accurately predict the path loss of a region.

One of the more recent path loss models utilizing collected measurements is one proposed by Robinson et al. [18]. Using the Technology For All (TFA) network in Houston, TX, they utilized a modified Flexible Path Loss Exponent model with a terrain correction factor derived iteratively from collected measurements. The model is an extension of Friis' fundamental study [6] and can be written as:

$$P_{rx} = P_{tx} + 10\alpha \log(d) - 20 \log(f)$$
$$- 20 \log\left(\frac{4\pi}{c}\right) \tag{1}$$

Here, P_{tx} and P_{rx} are the transmitted and received signal powers, respectively, α is the path loss exponent, f is the transmit frequency, d is the distance from the transmitter, and c is the speed of light. In their work, the authors use existing wireless mesh nodes and detailed terrain information to determine sections that are likely to share a similar path loss exponent. They then incrementally gather measurements around the borders of these sections in a push-pull algorithm to refine the coverage estimate of the mesh node.

2.2 Obtaining a Path Loss Exponent

In our statistical learning approach, it is necessary to train a classifier with path loss exponent observations derived from existing measurements to motivate predictions in areas that lack those same measurements. We borrow the idea of a spatially-dependent path loss exponent from Robinson et al. without the push-pull measurement adjustment algorithm, a reference node, and detailed terrain information (including material loss estimations). Instead, we use (1) in a square-shaped moving window over the region, using linear regression to obtain a path loss exponent for each window. Since the measurements are obtained from many different towers distributed over the area, each using potentially different transmit powers at different heights, we rely on a larger quantity of data to average out these inconsistencies. However, the accuracy (which we define as inversely proportional the standard deviation of obtained path loss exponents over several calculations using orthogonal measurements) is increased, which we rely on more heavily for our statistical learning framework. To calculate the metric of standard deviation on the path loss exponent, we divide the data considered into several independent sets, calculate path loss exponent for each independent set, and compute the standard deviation of the exponents derived. Again, this gives us a solid metric for path loss precision, even if the exponents themselves are biased by the data collection limitations.

3 IN-FIELD WIRELESS AND GEOGRAPHICAL DATA

In this section, we present our Android-based measurement gathering platform, which will be leveraged locally by us to gather a dense measurement set of wireless signal strengths in both a downtown region and suburban region. We also introduce the geographical feature data set that we use from the drive tested regions to establish a relationship between geodata and the attenuation of wireless signals. By using a smartphone based collection platform, we can gather Received Signal Strength Indicator (RSSI) measurements

that relate more directly to user experience than measurements collected with traditional network analyzing hardware.

3.1 Local Measurement Collection

Over the span of two weeks (over 30 in-car hours), we collected 6.7 million drive testing measurements by placing LG Nexus 4 smartphones in a vehicle and thoroughly driving throughout two regions in a snake like pattern, covering all available roads in each region. Since we are using the measurements for studying region-based path loss characteristics, the specific cellular technology used is less important. Therefore, the measurements were collected on GSM networks as they are still the most prevalent. The measurements were obtained at a relatively constant speed of 30 mph in two different areas of the Dallas metropolitan area. The first area is a suburban region several miles north of the city center with lush greenery prevalent throughout and is predominantly residential. The second area is in downtown, where there is far less vegetation, and the buildings are far taller than the suburban structures with non-uniform heights. Our goal is to use these two distinct regions to examine how differences in feature distribution affect the number of required measurements to characterize path loss to a certain degree of accuracy in each region.

3.2 Received Signal Strength in Android API

Each cellular measurement contains an RSSI field for each visible cellular tower, a GPS location, an accuracy reading, and physical speed of the device. While we now obtain RSSI readings in terms of dBm, most of our measurements were taken when the API reported RSSI in terms of Arbitrary Strength Units (ASU), an android specific quantized signal strength metric, which quantizes obtained RSSI values for GSM to 32 different levels shown in the equation below from the Android API [9].

$$P_{rx}(dBm) = 2 * P_{rx}(ASU) - 113 \tag{2}$$

$$P_{rx}(ASU) = [0, 31] \tag{3}$$

We consider $P_{rx}(ASU) = 0$ and 31 unusable since they correspond to SNR in an unlimited range; an ASU value of 31 includes any RSSI value above -51 dBm. Not including these measurements, however, clips the natural distribution of RSSI readings at locations with measurements near the quantization limits. The lower and upper bounds set by omitting measurements where $P_{rx}(ASU) = 31$ and $P_{rx}(ASU) = 0$, respectively, move the average RSSI at certain distances from the tower. Distances closer to the tower that generally have higher RSSI measurements near the upper bound may have a lowered average RSSI. Conversely, distances farther from that tower that generally have lower RSSI measurements near the lower bound may have a heightened average RSSI. The bias in the movement of average RSSI near the boundaries could end up changing (likely reducing) the value of the obtained path loss exponent. While the exact values of RSSI and path loss exponents are likely affected by the quantization error, we are not evaluating absolute path loss accuracy, only relative accuracy in our experiments, so the bias does not affect our results.

3.3 Geographical Feature Data

In order to obtain geographical feature information, we utilized an open-access online resource, Open Street Maps (OSM) [1], to identify, outline, and label specific regional features and output them to an easily accessible data structure for parsing. To this end, we mapped hundreds of offices, parks, houses, and other features in both the suburban and downtown regions and grouped them into the feature category classes for our statistical learning system. With statistical learning, the number of training observations necessary for accurate divisions scales up proportionally with the number of features used in the training. Due to this so-called curse of dimensionality, as well as the limited number of possible features to label in each region considered, it is necessary to divide all possible geographical features into relatively few feature categories for processing. With this in mind, we selected five feature categories under which all features were labeled: small buildings, medium buildings, large buildings, high foliage, and open space. In this system, we define small buildings consist of buildings that are under 5 stories tall (ground footprint is not considered for the category, but is implicitly considered when calculating feature distributions in a region). Similarly, we define the range of medium buildings as being between 5 and 15 stories tall and large buildings as being over 15 stories tall. These building height tiers were chosen to give each feature type non-trivial representation in the learning algorithm. Finally, we consider high foliage areas in the regions are areas with a large number of trees, and open space is the area defined by the complement to the set of all other features combined and includes structures such as roads, parking lots, etc. It is important to note that the feature set we consider is far from ideal; with more detailed geo-spatial feature data that is currently unavailable to us (such as exact building and foliage canopy heights), the GeoRIPE framework's accuracy will only improve.

(a) Downtown   **(b) Suburban**
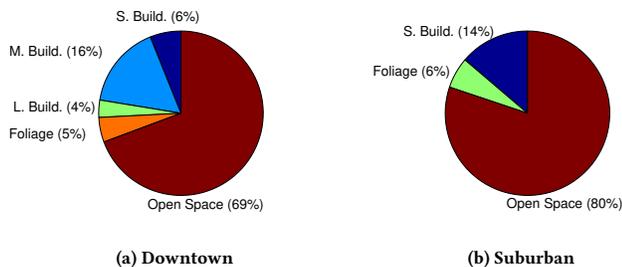
Figure 2: Regional Feature Distributions

The overall ratio of features in the downtown and suburban regions we examine are shown in Figures 2a and 2b, respectively. These ratios represent the relative space occupied by each feature according the following equations:

$$s = [f_1, f_2, ..., f_5] \tag{4}$$

$$\sum_i s_i = 1 \tag{5}$$

where s is a weighted vector for the normalized occupancy of each of the 5 features in the full region in terms of total feature area. From this figure, we can see that the suburban area lacks medium

and large buildings and has a higher percentage of open space than the downtown region, as anticipated. Ideally, we will be able to further differentiate and parse members of the open space set to derive additional feature categories in the future.

(a) Downtown   **(b) Suburban**
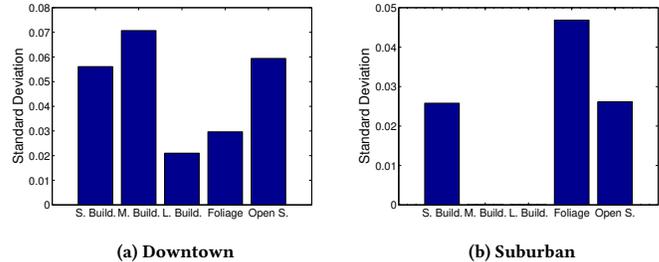
Figure 3: Regional Feature Distribution Deviation

Primarily, we want to examine how the features change over each region. To do this, we calculated the ratio of features in a moving window over each region. Treating the percentage of each feature in the windows as a random variable, we then compute the standard deviation, giving us a picture of the feature variability as we move across each region. The standard deviation of features can be seen for the downtown and suburban regions in Figures 3a and 3b, respectively. From these figures, we can see that the variance of features in the suburban region is, overall, significantly smaller than in the downtown region. This is because, in the suburban region, the grouping of features are polarized (ex: houses in half the region, foliage in the other half), while the in the downtown region, neighboring areas have a higher diversity in their feature composition.

4 DENSITY-DEPENDENT TILING OF IN-SITU DATA

While we have examined the differences in geographical feature distribution of the two areas, we have yet to explore the impact of changing the size of the subregion, or tile, used to group measurements spatially. In this section, we explore the differences in path loss exponent changes between these regions, their relative subregions, and the trade off between tile size, measurement density, and measurement error in evaluating path loss.

4.1 Extreme 1: Highly-Sparse Crowdsourced Data

The first scenario is one in which the data set has very few measurements. In such a situation, we need all the measurements we have to assign a single path loss exponent to characterize an entire region, similar to the traditional approach. In other words, dividing the region into smaller areas to have more path loss precision cannot occur because there is a lack of a sufficient number of measurements to compose a path loss exponent estimate. The result here can be considered a rough average of path loss over the entire region; however, accuracy at any given area depends on the variability of the region itself. While a single exponent over an entire region may create a simpler coverage calculation, it may not be accurate, especially in more diverse region types such as

large cities. When enough measurements are available, we can divide the region into independent tiles for characterization based on measurement density and region type instead. Figure 4 shows the suburban region is divided into 6 and 24 tiles for path loss evaluation, demonstrating the disparity in derived path loss exponents for the same area using different tile sizes. While some smaller tiles match the their large tile counterparts, others are different, alluding to diverse environmental characteristics.

(a) 6 Tiles **(b) 24 Tiles**
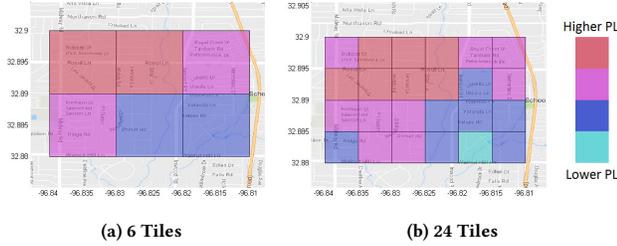
Figure 4: Suburban region path loss divided into a) 6 and b) 24 tiles.

4.2 Extreme 2: Highly-Dense Crowdsourced Data

We now examine the other extreme, when a very large number of measurements are available. In this case, the tile size is not limited by measurement density or acceptable error. With highly-dense measurements (*e.g.*, as the number of measurements approach infinity), tile size is virtually unbounded, and the standard deviation of path loss approaches zero (orthogonal subsets measurements would regress to the same exponent when evaluating as the subset size approaches infinity). Instead, the variability of the terrain determines the effective lower bound on the terrain characterization resolution, preventing the tile size from going to zero. In other words, decreasing the title size resolution after a certain point does not provide any additional information about path loss in the region.

Region	Tiles	Diff. Mean	Diff. Variance
Downtown	6	0.0668	0.0040
	24	0.0944	0.0059
Suburban	6	0.0742	0.0029
	24	0.0464	0.0014

Table 1: Mean and variance of differences between neighboring tiles' path loss.

Depending on the nature of the region being analyzed, the mean and variance of the difference between neighboring tiles changes with the tile size. Table 1 shows the mean and the variance in path loss exponent calculation differences between neighboring tiles for both the downtown and suburban regions of different tile sizes. As the tile size decreases, we observe different behavior from the two region types. In the downtown region, the differential mean and variance increase with a smaller tile size, while in the suburban region, the opposite is true. For more diverse regions like downtown,

using smaller tile sizes has a larger benefit in characterizing the spatial diversity of path loss. In less diverse regions like homogeneous neighborhoods, the differential path loss throughout the region does not require as high of a resolution; the path loss variability seen from smaller tile sizes is below the noise floor in generating the path loss exponents. Thus, the measurement density available along with the region type's path loss variability must be jointly considered in determining a minimum tile size for characterization.

4.3 General Case: In-Situ Tile Size Adaptation

In most cellular networks, it is likely that the set of available measurements is neither highly sparse nor infinitely dense. Instead, the system is generally in a state between these two extremes. Hence, choosing the tile size of the region becomes a critical issue since it is not initially clear if the measurement density or the terrain heterogeneity will drive the tile size.

(a) Suburban **(b) Downtown**
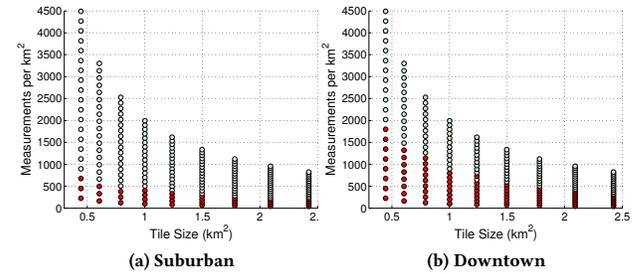
Figure 5: Examining tile size versus measurement density over different land uses for a given acceptable error.

Given a set of measurements with a specific measurement density, a minimum tile size exists that remains below the reliability threshold. In Figures 5a and 5b we set a standard deviation threshold of 0.0125 for our path loss evaluation and examined the minimum tile size for different measurement densities available. In these figures the lighter points are below the threshold and darker, red dots are above the threshold. For both land use types, we see that as the measurement density increases, the minimum tile size achievable under the threshold decreases, enabling a finer grain resolution while maintaining the reliability we desired. However, the suburban region consistently requires a lower measurement density to be below the error threshold than the downtown region because the suburban region has a lower terrain variability. From these figures, we can see that to achieve a tile size of 1 km^2 under the error threshold, the suburban region requires a measurement density of 500 measurements per km^2 while the downtown region requires 800. This relationship between the different land uses holds for each other tile size as well. The measurement density required for a certain resolution of path loss increases with the heterogeneity of the region.

5 EXPERIMENTALLY EVALUATING MAP-BASED MODELING

Despite that fact that there have been several works that suggest measurement distribution and geographical features play an enormous role in the resulting path loss characterization of a region,

there has not been a study showing how these metrics can be used to quantify the number of measurements required to characterize an area. In this section, we take a critical look at the impact of measurement distribution and geographical feature components on path loss precision. More specifically, we compare measurement distributions obtained from crowdsourcing versus drive testing measurements, examine geographical feature components of our two metropolitan region types, and correlate these feature distributions with both path loss exponents as well as the number of measurements required to obtain a certain precision in characterization. Our goal is to use regional geographical features to learn how to properly collect measurements, ensuring a predetermined precision in path loss characterization.

5.1 Path Loss Metric and Geographical Feature Correlation

Using geographical features as a region specific identifier, we want to understand how specific geographical features can be used to characterize path loss throughout a region. We now explore four different path loss related metrics to determine which had the closest relationship, and therefore the highest suitability, to be used as the target for our geographical feature based statistical learning approach. The four metrics we examine are path loss exponent (PLE), differential path loss exponent (DPLE), number of measurements required (MR) for path loss convergence, and the differential number of measurements required (DMR) for convergence. The MR and PLE metrics are calculated for a given region using Algorithm 1, which is initialized with parameters listed in Table 2. Algorithm 1 can be visualized as a sliding window filter moving across the region as illustrated in Figure 6. In this algorithm, the first two loops control the moving window as it shifts vertically and horizontally, respectively. For a given window at position v, h, we compute the path loss exponent directly with all available data, giving the PLE metric. Following that, we divide the data into G separate groups, calculate the path loss exponent in each group, and take the standard deviation over all exponents. We increase the number of measurements in each of these groups by S until the standard deviation is under a certain threshold (chosen to be whatever accuracy is acceptable, we chose 0.03 because that was about the point that an a linear increase in the number of measurements started to have diminishing returns). Additionally, When the standard deviation falls under this threshold, we record the measurements in each group as the MR metric.
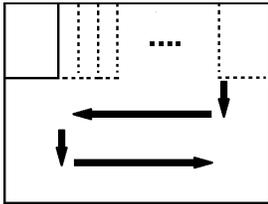
Figure 6: Visualization of Algorithm 1.

We use this algorithm to determine a map of the measurements required and path loss exponent metrics over a region. With our feature data for the region, we can derive a similar map of feature distributions using the same windowing method. The differential

Parameter	Setting	Description
W	$1 km^2$	Moving window area
V	20	Number of vertical shifts
H	40	Number of horizontal shifts
σ	0.03	Desired std. dev.
S	20	Measurement step size
G	30	Number of orthogonal groups

Table 2: Spatial feature and path loss metric algorithm parameters.

Data: measurements (M)
Result: V by H PLE and MR matrices
Initialize Parameters;
for $v \leftarrow 1$ **to** V **do**
  for $h \leftarrow 1$ **to** H **do**
    PLE$(v, h) \leftarrow ComputeExponent(\forall M \in W)$;
    $group(1..G) \leftarrow \forall M \in W$ split into G sets;
    $P_{size} \leftarrow 0$;
    while $\sigma_{temp} \geq \sigma$ **do**
      $\sigma_{temp} \leftarrow \infty$;
      $P_{size} \leftarrow P_{size} + S$;
      for $g \leftarrow 1$ **to** G **do**
        $P \leftarrow P_{size}$ elements $\in group(g)$;
        $exponent(g) \leftarrow ComputeExponent(P)$;
      end
      $\sigma_{temp} \leftarrow ComputeStdDev(exponent(1..G))$;
    end
    MR$(v, h) \leftarrow P_{size}$;
    $W \leftarrow W$ horizontally shifted by 1;
  end
  $W \leftarrow W$ vertically shifted by 1;
end
Algorithm 1: Algorithm for computing PLE and MR metrics.

---

metrics, differential path loss exponent and differential measurements required, can be easily derived from column and row differentiation of the PLE and MR matrices, respectively. A corresponding differential feature distribution map can be derived in the same manner. With matching metric and feature maps, we can correlate each metric with the corresponding feature map to obtain a sample Pearson correlation coefficient (the standard equation for correlating discreet groups) using Equation 6.

$$r_{ij} = \frac{\sum_{k=1}^{n}(x_{ik} - \bar{x}_i)(y_{jk} - \bar{y}_j)}{\sqrt{\sum_{k=1}^{n}(x_{ik} - \bar{x}_i)^2 \sum_{k=1}^{n}(y_{jk} - \bar{y}_j)^2}} \tag{6}$$

In this equation, n is the number of samples, x_{ik} is sample k of feature i, y_{jk} is the sample k of path loss metric j, and $\bar{x}_i$ and $\bar{y}_j$ are the average distribution of feature i and the average of path loss metric j, respectively.

We want to select a path loss metric to use as a training class for the statistical learning framework that has the highest correlation coefficients with the feature set to provide clear decision boundaries. The correlation coefficients for each of the path loss metrics in

the downtown and suburban regions are shown Tables 3 and 4, respectively.

Metric	S. Building	M. Building	L. Building	Foliage	Open Space
PLE	-0.23	0.02	-0.05	0.17	0.18
DPLE	-0.05	-0.03	0.02	-0.04	0.07
MR	**-0.32**	**0.34**	**0.49**	-0.10	-0.23
DMR	-0.06	0.05	0.03	0.01	-0.02

Table 3: Downtown metric-feature correlation coefficients.

Metric	S. Building	M. Building	L. Building	Foliage	Open Space
PLE	**0.36**	NA	NA	**0.31**	**-0.38**
DPLE	-0.15	NA	NA	0.05	0.05
MR	**-0.53**	NA	NA	**0.44**	-0.27
DMR	-0.06	NA	NA	0.01	0.04

Table 4: Suburban metric-feature correlation coefficients.

We can see that the MR metric has the highest overall correlation coefficient magnitude and is likely the best contender for a simple decision tree based learning algorithm. Interestingly, we see that for the suburban region in particular, the correlation coefficient for the MR metric are negative with small buildings and positive with foliage, while positive with both for the PLE metric. This suggests that while increased buildings and foliage contribute to a larger path loss exponent (as expected), the number of measurements required to drop below the 0.03 path loss exponent standard deviation increases only with the percentage of foliage.

(a) Downtown

(b) Suburban
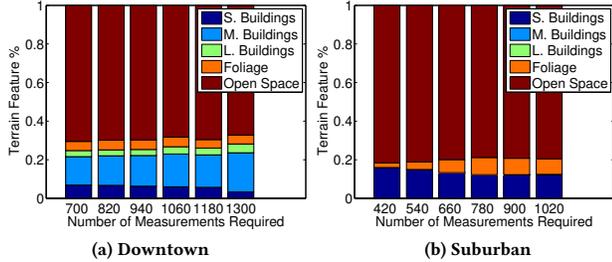
Figure 7: Average feature distributions for different MR tiers.

From this result, we can see that while the small buildings feature increases the path loss exponent, it decreases received power variability, while the foliage feature increases received power variability. This trend is visualized for the downtown and suburban areas in Figures 7a and 7b.

To understand how each terrain feature affects the measurement requirements individually, we examined the trends of each feature distribution as the number of required measurements increases in Figure 8. From this figure, we can see that in the suburban region, the individual feature impact is quite clear; increases in the percentage of foliage and decreases in the percentage of small buildings increases the number of measurements required, while the open space component doesn't fluctuate much at all. Conversely, we

(a) Downtown
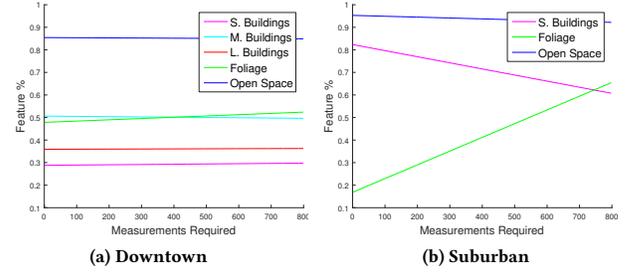
(b) Suburban

Figure 8: Feature trends for increasing measurement requirements.

cannot induce distinct trends from the downtown region. We see in the suburban region, there are only two features driving the increase in measurement requirements, thus trends can be easily seen. In the downtown region, however, each feature apart from open space has an effect on the measurements, thus trends from individual factors cannot be so easily derived.

5.2 Classifier Training for MR Prediction

To validate the GeoRIPE framework, we divide the MR results for the downtown and suburban regions into 6 same-sized class bins. As seen in Figure 9, the class groupings are not homogeneous for either of the regions. Unsurprisingly, the downtown region class distribution has a higher mean number of required measurements than the suburban region due to its higher geographical complexity.

(a) Downtown

(b) Suburban
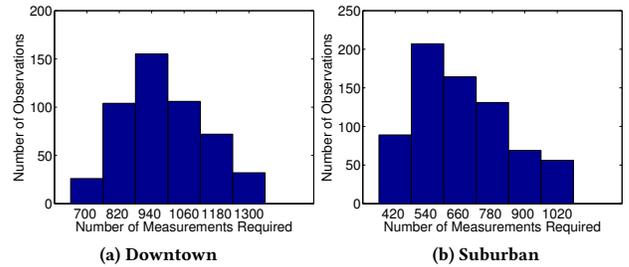
Figure 9: Regional MR Class Distributions

Under these class groupings, the input terrain feature distributions used each measurement class grouping are shown in Figure 7. In reference to Figure 9, we see that the majority of the regional features fall into groups centered around 940 and 540 measurements for the downtown and suburban regions, respectively. Thus, these bins will have a higher weight under the learning framework.

To train each the decision tree classifier, it is important to allow equal training weights per class as much as possible to balance the tree and not over-fit the data. For this, we randomly selected an equal number of observations for resulting in each class to balance the observations per bin. We further divided this set of observations into two separate training and validation observation sets, again being sure to have equal class representation in each set. We then trained the decision tree classifier with the training set and predicted MR classes with the validation set.

Due to the linear relationship between the MR metric and class, it is important to look at not only the prediction accuracy in choosing the correct class, but also the distribution of predicted class offsets (how many classes away from the correct class) when the correct class is not chosen. This is because a lower average offset between the predicted and correct MR class is nearly as important as the accuracy in choosing the correct class. For example, predicting the adjacent class is not as detrimental to the measurement number estimation as predicting multiple classes away.

(a) Downtown **(b) Suburban**
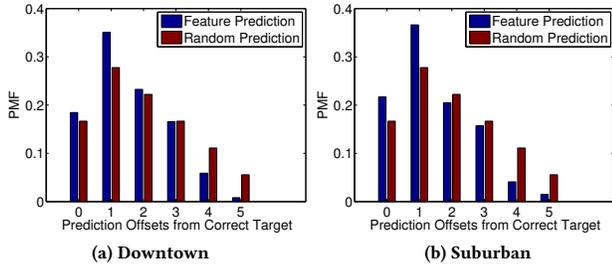
Figure 10: Regional Feature Versus Random Prediction Off-sets

Figure 10 shows the class prediction offset magnitudes for using feature prediction to choose a class and choosing a random class (according to the frequency of occurrence). This result shows that for both the downtown and suburban region, the average predicted class offset is significantly lower using the feature prediction than choosing a class at random, even if there may not be a very high accuracy in actually predicting the correct class.

5.3 Uniform Drive Testing Comparison

Using our trained and validated decision tree classifier, we wanted to compare GeoRIPE to uniform drive testing in two scenarios. First is a dense uniform drive testing scenario, in which measurements are gathered according to the requirements of the subregion (window from Algorithm 1) with the highest geographical variability. More specifically, we experimentally found the number of measurements required in the worst case subregion to meet the accuracy threshold and uniformly take that number of measurements over every subregion. The other is a sparse uniform drive testing scenario, in which measurements are gathered according to the measurement requirements of the subregion with the lowest geographical variability. In this experiment, the goal is to stay under a predetermined path loss exponent standard deviation (corresponding to an accuracy that a network provider would require) while using the lowest amount of measurements. To do this, we divided the regions into several uniformly sized tiles and gathered several orthogonal sets of measurements from each tile according to the sparse, dense, and GeoRipe predicted number of measurements. For each orthogonal set in each scheme, we calculated the path loss exponent and took the standard deviation over all exponents for each of the three techniques. By doing this, we can compare the path loss exponent accuracy and the number of measurements required for each technique. For this experiment, we trained the GeoRIPE classifier to predict the measurements required for the

standard deviation of 0.03 using half of the region, and predicted the number of measurements required for the for the other half. We repeated this experiment for both the downtown and suburban regions, and the results can be seen in Table 5.

Technique	Region	Average σ	Avg. Meas. per km^2
GeoRIPE	Downtown	0.0286	194
GeoRIPE	Suburban	0.0284	186
Sparse	Downtown	0.0631	40
Sparse	Suburban	0.0440	80
Dense	Downtown	0.0205	400
Dense	Suburban	0.0188	440

Table 5: GeoRIPE Standard Deviation and Measurements Compared to Sparse and Dense Uniform Drive Testing Scenarios

From this table, we can see that the sparse drive testing does not meet the required standard deviation of below 0.03 that we set at the start of the experiment. The dense drive testing does stay under the standard deviation requirement, using the minimum number of measurements to do so over all areas. Using GeoRIPE, the standard deviation requirement is also met, but it requires 58% fewer measurements than uniform dense drive testing to get all subregions below the threshold. From the GeoRIPE results, the effect of geographical complexity can be clearly seen; a lower standard deviation of path loss exponents is obtained using fewer measurements in the suburban region than the geographically more complex downtown region.

In addition to meeting the accuracy requirements using the least number of measurements, we wanted to evaluate the benefits of using the GeoRIPE framework over a uniform distribution that uses the same number of total measurements. To do this, we used measurements from the GeoRIPE distribution given by Equation 7.

$$p_x = \frac{M_x}{\sum_{\hat{x}}^{X} M_{\hat{x}}} \quad (7)$$

Here, x is a single section in the set of all tiles X, p_x is the fraction of measurements to be collected in section x, and M_x is the set of predicted MR values of tile x. We collected several orthogonal subsets of measurements in each tile for an increasing number of total measurements in each region and compared the accuracy of the two techniques. For each orthogonal subset in each tile, we calculated a path loss exponent and computed the standard deviation of the path loss exponents in each tile. The standard deviation for all tiles was averaged at each number of total measurements and the results were organized by standard deviation. For selected standard deviation, each of the techniques required a different number of measurements per km^2, resulting in Table 6.

From this table, we see that as the threshold for standard deviation is lowered, measurements required increases approximately 10% 'faster' using uniform drive testing. So, while the true value in using the GeoRIPE framework is predicting the number of measurements required over a region to meet a certain path loss exponent accuracy, the normalized GeoRIPE distribution also achieves the desired path loss accuracy with proportionally fewer measurements

σ	GeoRIPE # Meas.	Uniform # Meas.
0.050	64	64
0.045	80	84
0.040	96	108
0.035	128	138
0.030	178	196
0.025	252	276
0.020	400	436

Table 6: GeoRIPE versus Uniform Drive Testing Measurements to Achieve a Fixed σ

than the uniform counterpart. This result, however, only analyzes the average path loss over the entire region. Using a similar windowing method previously described, we wanted to see how many measurements it took to bring the standard deviation of the path loss exponent in all the windows to fall below these thresholds. We found that while the number of measurements for the GeoRIPE framework to accomplish this is similar to the numbers in Table 6, uniform drive testing required an average of 20% more measurements than the listed numbers. The biggest difference in this experiment is alluded to in Table 5, wherein GeoRIPE requires 58% fewer measurements to go below the standard deviation threshold than uniform drive testing.

6 RELATED WORK

Measurement Collection Approaches. Due to the low cost of crowdsourcing from smartphones, the technique has been used by many other groups to collect data about wireless networks. In a study by Huang et al. [12], LTE performance data was collected by creating an Android application named 4GTest. This application gained 3,000 users during 2 months of data collection and collected data that focused on media streaming by mobile clients. With this data, [12] was able to show that with the download speed increase seen with LTE networks, the traffic bottleneck shifted from the network to the processing power of the mobile devices. In [5], an Android application was again used to capture network speed data. This study focused on comparing the speeds of 802.11 networks to the speeds of LTE networks in major cities around the globe. Neidhardt et al. used a crowdsourced infrastructure to provide an open source and more accurate base station location and coverage estimation system [15]. While they had promising results on the base station localization aspect, they concede that cellular coverage estimation was lacking with their purely crowdsourced measurements, especially in urban environments with diverse terrain features. Our work focuses on the minimum measurement requirements according to different geographical features of a given region.

Measurement-Driven Path Loss Evaluation. There have been several measurement studies that strive to more accurately characterize path loss in specific region types. Hata et al. [11] and Okumura et al. [16] specifically focus on accurate characterization in urban regions. Using measurements gathered by [16] in Tokyo, Japan, Hata et al. empirically derived a path loss prediction formula with correction factors for various region types such as large-city urban, small-city urban, suburban, and open areas. Additionally,

the Hata model considers base station transmitter height. Similar to the path loss prediction curves found by [16] in Japan, Allsebrook et al. [3] evaluated path loss prediction curves for three British cities: Birmingham, Bath, and Bradford. Akimoto et al. [2] derived a model based on gathered measurements in a rural area using the 2 and 5 GHz bands. Similarly, [8] studied measurements collected in a suburban neighborhood at 5.7 GHz as did [10] with measurements taken in Istanbul in the GSM-900 band. More recently, Robinson et al. sought to minimize the number of measurements necessary to accurately characterize mesh node coverage in the TFA network in Houston [18]. Their work uses an online push-pull measurement gathering approach, taking very few active measurements on an existing deployment based on terrain features in the area. Additionally, Sayrac et al. [19] and several others [4, 7, 14] try to reduce the number of drive testing measurements required for coverage evaluation via Baysian kriging, showing how their techniques can be used to detect coverage holes. However, their analysis replies on the spatial correlation between the measurements themselves to detect coverage holes from existing transmission infrastructure. In contrast, our approach aims to analyze geographical features of a region and predict the number of measurements required to obtain an accurate estimate of path loss throughout, including from transmission sources that do not yet exist, by tying the measurement requirements to the terrain itself.

7 CONCLUSION

In this paper, we built the GeoRIPE framework which predicts the minimum number of in-field measurements required to accurately characterize the path loss of a region according to that region's geographical features. To find if such measurements would be sufficient for a given area, we gathered millions of signal strength measurements along with geographical feature ratios in both a downtown and suburban region. Using this data, we correlated several distinct geographical features with different metrics for path loss evaluation complexity. We found that, together, these features are correlated with the number of measurements required to achieve a fixed path loss accuracy. We also evaluated the merit of using area bounded path loss metrics. By abstracting propagation loss parameters away from specific paths and binding them to a specific area, we are able to evaluate path loss for arbitrary paths through the area. We found that the size of the individual path loss evaluation areas should be selected based on the complexity of the terrain features residing in each area. In general, the more complex the area, the smaller the evaluation area should be. Finally, to validate our work, we compared drive testing using our GeoRIPE framework to uniform drive testing in each region. We found that our technique, as opposed to spatially uniform drive testing, required fewer measurements to achieve a similar path loss characterization accuracy.

ACKNOWLEDGMENTS

REFERENCES

[1] [n. d.]. Open Street Maps. www.openstreetmap.org. ([n. d.]). License: http://www.openstreetmap.org/copyright.

[2] Mamoru Akimoto, Tatsuya Shimizu, and Masashi Nakatsugawa. 2006. Path loss estimation of 2 GHz and 5 GHz band FWA within 20 km in rural area. *Proc. of ISAP* (2006).

[3] K. Allsebrook and J.D. Parsons. 1977. Mobile radio propagation in British cities at frequencies in the VHF and UHF bands. *Vehicular Technology, IEEE Transactions on* 26, 4 (1977), 313–323.

[4] Hajer Braham, Sana Ben Jemaa, Gersende Fort, Eric Moulines, and Berna Sayrac. 2017. Fixed Rank Kriging for Cellular Coverage Analysis. *IEEE Transactions on Vehicular Technology* 66, 5 (2017), 4212–4222.

[5] Shuo Deng, Ravi Netravali, Anirudh Sivaraman, and Hari Balakrishnan. 2014. WiFi, LTE, or Both?: Measuring Multi-Homed Wireless Internet Performance. In *Proceedings of the 2014 Conference on Internet Measurement Conference.* ACM, 181–194.

[6] Harald T Friis. 1946. A note on a simple transmission formula. *proc. IRE* 34, 5 (1946), 254–256.

[7] Ana Galindo-Serrano, Berna Sayrac, Sana Ben Jemaa, Janne Riihijärvi, and Petri Mähönen. 2013. Automated coverage hole detection for cellular networks using radio environment maps. In *Modeling & Optimization in Mobile, Ad Hoc & Wireless Networks (WiOpt), 2013 11th International Symposium on.* IEEE, 35–40.

[8] S.S. Ghassemzadeh, H.R. Worstell, and R.R. Miller. 2010. Wireless Neighborhood Area Network Path Loss Characterization at 5.7 GHz. In *Proc. of IEEE VTC-Fall.*

[9] Google. 2017. Android API Description. (2017). https://developer.android.com/reference/android/telephony/CellSignalStrength.html

[10] B.Y. Hanci and I.H. Cavdar. 2004. Mobile radio propagation measurements and tuning the path loss model in urban areas at GSM-900 band in Istanbul-Turkey. In *Proc. of IEEE VTC-Fall.*

[11] M. Hata. 1980. Empirical formula for propagation loss in land mobile radio services. *Vehicular Technology, IEEE Transactions on* 29, 3 (1980), 317–325.

[12] Junxian Huang, Feng Qian, Alexandre Gerber, Z Morley Mao, Subhabrata Sen, and Oliver Spatscheck. 2012. A close examination of performance and power characteristics of 4G LTE networks. In *Proceedings of the 10th international conference on Mobile systems, applications, and services.* ACM, 225–238.

[13] William CY Lee. 1985. Estimate of local average power of a mobile radio signal. *Vehicular Technology, IEEE Transactions on* 34, 1 (1985), 22–27.

[14] Han-Wen Liang, Chih-Hsiang Ho, Li-Sheng Chen, Wei-Ho Chung, Shih-Yi Yuan, and Sy-Yen Kuo. 2016. Coverage hole detection in cellular networks with deterministic propagation model. In *Intelligent Green Building and Smart Grid (IGBSG), 2016 2nd International Conference on.* IEEE, 1–6.

[15] Eric Neidhardt, Abdulbaki Uzun, Ulrich Bareth, and Axel Kupper. 2013. Estimating locations and coverage areas of mobile network cells based on crowdsourced data. In *Wireless and Mobile Networking Conference (WMNC), 2013 6th Joint IFIP.* IEEE, 1–8.

[16] Yoshihisa Okumura, Eiji Ohmori, Tomihiko Kawano, and Kaneharu Fukuda. 1968. Field strength and its variability in VHF and UHF land-mobile radio service. *Rev. Elec. Commun. Lab* 16, 9 (1968), 825–73.

[17] Chris Phillips, Douglas Sicker, and Dirk Grunwald. 2013. A survey of wireless path loss prediction and coverage mapping methods. *Communications Surveys & Tutorials, IEEE* 15, 1 (2013), 255–270.

[18] Joshua Robinson, Ram Swaminathan, and Edward W. Knightly. 2008. Assessment of urban-scale wireless networks with a small number of measurements. In *Proc. of ACM MobiCom.*

[19] Berna Sayrac, Janne Riihijärvi, Petri Mähönen, Sana Ben Jemaa, Eric Moulines, and Sébastien Grimoud. 2012. 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, 43–48.

[20] Hweechul Shin. 2010. *Measurements and models of 802.11 b signal strength variation over small distances.* Ph.D. Dissertation. University of Delaware.

[21] Bernard Sklar. 1997. Rayleigh fading channels in mobile digital communication systems. I. Characterization. *Communications Magazine, IEEE* 35, 7 (1997), 90–100.