Appendix A

EDAM: The Effective Directivity Antenna Model

Increasingly, wireless networks are using directional antennas to improve the throughput, reach of networks [181], or to reduce interference between adjacent networks and other noise sources. A more recent development is the use of electronically steerable directional or phase array antennas [149, 33, 215]. These antennas provide better network performance by dynamically controlling the radiation pattern of the antenna. Networks that utilize these antennas can reap substantial improvements in efficiency at all layers of the networking stack.

Different network simulators model such antennas with different degrees of fidelity. This chapter argues that the models in the most common network simulators make such simplifying assumptions that it is often difficult to draw strong conclusions from the simulations derived using those models. This is demonstrated this using a series of measurements with several different and widely used directional antenna configurations. A more accurate model is developed based on measurements and intuitions about radio propagation[1]. This model captures more about the uncertainty of the environment than the specifics of the antenna and that our results should be generally applicable to many different directional antenna patterns with similar gain characteristics.

The measurement study described here uses sophisticated measurement equipment, including a vector signal analyzer (VSA) and signal generator (VSG). Since the costs of such equipment are prohibitive, a method that uses inexpensive equipment (such as standard networking cards) is also developed to produce

[0] Work in this appendix has also been published in [26, 25, 27, 28, 30].

[1] All of the measurements collected for this research are available publicly at [6]. An implementation of our model for the Qualnet 4.5.1 simulator is available at `http://systems.cs.colorado.edu`.

the data needed for the derived models.

A.1 Directional Models

The simulators commonly used in networking research do not consider antenna directionality and radio propagation as interacting variables. This paper considers three widely used simulators, *OpNet*, *QualNet*, and *NS-2*. Each one supports several models of radio propagation, but they all follow the same general model with regard to antenna gain: For any two stations i and j, the received signal strength is computed according to the general form of equation A.1:

$$\text{Received Power} = P_{tx} * G_{tx} * |PL(i,j)| * G_{rx} \tag{A.1}$$

The received power P_{rx} is the product of the transmitted power P_{tx}, the transmitter's gain G_{tx}, the magnitude of path loss between the two stations $|PL(i,j)|$, and the receiver's gain G_{rx}.

The transmitter and receiver gains are treated as constants in the case of omnidirectional (effectively isotropic in the azimuth plane) antennas. For directional antennas, however, gain is an antenna-specific function of the direction of interest. The orientation of an antenna can be modeled in terms of its zenith (ϕ) and azimuth (θ). Then, for a given antenna a, characterization function $f_a(\phi, \theta)$ can be defined:

$$\text{Gain in direction } (\phi, \theta) = f_a(\phi, \theta) \tag{A.2}$$

$$\text{Combined gain} = f_a(\phi, \theta) * f_b(\phi', \theta') \tag{A.3}$$

Correspondingly, the receiver gain is modeled by a (potentially different) function of the direction from which the signal is received. Besides being a source of interference for a dominant signal, the energy traveling along secondary paths also carries signal. If one of the weaker signals for a transmitter happens to be aligned with a high gain direction of a receiving antenna, the received power from that path can be greater than that of the primary path.

The above models describe the power emitted in, or received from, a single direction. In reality, the transmitter's power is radiated in all directions, and the receiver aggregates power (be it signal or noise) from all directions. Although the simulators considered here assume that the single direction of interest for each

station is precisely toward the other station, equations A.1 and A.3 can be generalized to the case where there are multiple significant signal paths:

$$P_{rx} = \sum_{l \in \text{paths}} P_{tx} * f_a(\phi_l, \theta_l) * PL_l(i,j) * f_b(\phi'_l, \theta'_l) \tag{A.4}$$

In Equation A.4, note that that P_{rx} is not necessarily all "signal". It may be the case that only one signal is decodable and the others destructively interfere. In this case equation A.5 is a better model:

$$P_{rx} = \max_{l \in \text{paths}} P_{tx} * f_a(\phi_l, \theta_l) * PL_l(i,j) * f_b(\phi'_l, \theta'_l) \tag{A.5}$$

Both of these models assume that there is some way to describe available paths that a signal may take. As with the Rayleigh and Rician fading models, it may be possible to build a parameterized model of those paths for "cluttered" and "uncluttered" environments. This is the approach taken here, using measured data to determine the model.

With any of the three simulators we consider, the user has the freedom to provide any type of mapping between gain and angle. This means that the user could conceivably make measurements with their desired hardware in their desired environment, much as we have done, and then install this as the pattern. However, even though the antenna can conceivably be modeled arbitrarily well, the *directionality of the signal* is an effect of the interaction between antenna and environment and that modeling both in isolation, however well, misses significant effects. This chapter proposes a combined empirical model that attempts to account for both the pattern of the antenna and the deviation from this pattern due to environmental effects.

A.2 Method

This section describes the method devised for deriving empirical models for antenna patterns using commodity hardware and address any reservations about their accuracy by providing a means for equipment calibration.

A.2.1 Data Collection Procedure

Two laptops are used, one configured as a receiver and the other as a transmitter. Each is equipped with an Atheros-based MiniPCI-Express radio that is connected to an external antenna using a U.Fl to N pigtail adapter and a length of LMR-400 low loss antenna cable. The receiver laptop is connected to a 7 dBi omnidirectional antenna on a tripod approximately two meters off the ground. The transmitter laptop is connected to the antenna we intend to model on a tripod 30.5 m from the receiver, also two meters off the ground. The transmitter tripod features a geared triaxial head, which allows precise rotation.

The transmitter radio is put in 802.11x ad hoc mode on the least congested channel. The transmitter's ARP table is manually hacked to allow it to send UDP packets to a nonexistent receiver. The receiver is put in monitor mode on the same channel and logs packets with tcpdump. Finally, both the receiver and transmitter must have antenna diversity disabled. With the equipment in place, the procedure is as follows: For each 5 degree position about the azimuth, send 500 unacknowledged UDP packets. Without intervention otherwise, due to MAC-layer retransmits, each will be retried 8 times, resulting in 4000 distinct measurements.

During the experiment, the researchers themselves must be careful to stay well out of the nearfield of the antennas and to move to the same location during runs (so that they, in effect, become a static part of the environment). If additional data is desired for a given location, multiple receivers can be used, provided the data from them is treated separately (as each unique path describes a unique environment).

In the process of collection, some packets will be dropped due to interference or poor signal. In practice, the percentage of dropped frames *per angle* is very small: the maximum lost frames per angle in the data sets is on the order of 5%, with less than 1% lost being more common (the mean is 0.01675%). Moreover, the correlation coefficient between angle and loss percentage is -0.0451, suggesting that losses are uniformly distributed across angles. Given that 4000 samples have been taken in each direction, noise in the measurements due to packet loss is negligible.

A.3 Measurements

This section explains the data sets collected, discusses the normalization procedure developed, and gives some high level statistical characterization of the data.

A.3.1 Experiments Performed

In order to derive an empirical model that better fits real world behavior, data was collected in several disparate environments with three different antennas. A summary of these data sets is provided in table A.1. With the exception of the reference patterns, all of the measurements were made with commodity hardware by sending many measurement packets between two antennas and logging received signal strength (RSS) at the receiver. The three antenna configurations used include: (1) a HyperLink 24dBi parabolic dish with an 8 degree horizontal beamwidth, (2) a HyperLink 14dBi patch with a 30 degree horizontal beamwidth, and (3) a Fidelity Comtech Phocus 3000 8-element uniform circular phased array with a main lobe beamwidth of approximately 52 degrees. This phased array functions as a switched beam antenna and can form this beam in one of 16 directions (on 22.5 degree increments around the azimuth). For the HyperLink antennas, the same antenna was used in all experiments of a particular type to avoid intra-antenna variation due to manufacturing differences.

In addition to the *in situ* experiments, a "reference" data set is available for each configuration. The Array-Reference data set was provided to us by the antenna manufacturer. Because HyperLink could not provide us with data on their antennas, Parabolic-Reference and Patch-Reference were derived using an Agilent 89600S VSA and an Agilent E4438C VSG in a remote floodplain[2].

Following is a brief description of each of the experiments:

Parabolic-Outdoor-A, Patch-Outdoor-A: A large open field on the University of Colorado campus was used for these experiments. The field is roughly 150m on a side and is surrounded by brick buildings on two of the four sides. Although there is line-of-sight and little obstruction, the surrounding structures make this

[2] We were unable to aquire access to an anechoic chamber in time for this study, but would like to make use of one in future work, for even cleaner reference measurements.

location most representative of an urban outdoor deployment.

Parabolic-Outdoor-B, Patch-Outdoor-B: A large University-owned floodplain on the edge of town was used for the most isolated data sets. The floodplain is flat, recessed, and is free from obstruction for nearly a quarter mile in all directions. This location is most representative of a rural backhaul link.

Array-Outdoor-A: The same open field is used as in the Parabolic-Outdoor-A and Patch-Outdoor-A data sets. The collection method here differs from that described in section A.2. A single phased array antenna is placed approximately 30 m away from an omnidirectional transmitter. The transmitter sends a volley of packets from its fixed position as the phased array antenna electronically steers its antenna across each of its 16 states, spending 20 ms in each state. Several packets are received in each directional state. The phased array antenna is then manually rotated in 10 degree increments while the omnidirectional transmitter remains fixed. The same procedure is repeated for each of 36 increments. Moving the transmitter changes not only the angle relative to the antenna but also the nodes' positions relative to their environment. To address this confound, each physical position is treated as a separate experiment. This means that the number of angles *relative to the steered antenna pattern* is limited to the number of distinct antenna states (16). The transmission power of the radio attached to the directional antenna was turned down to 10dBm to produce more tractable noise effects (the default EIRP is much too high to model small scale behavior).

Parabolic-Indoor-A and Patch-Indoor-A: This data set was collected in the University of Colorado Computer Science Systems Laboratory. The directional transmitter was positioned approximately 6 m from the receiver in a walkway between cubicles and desks. This is the most cluttered environment studied.

Parabolic-Indoor-B, Parabolic-Indoor-C, Patch-Indoor-B, and Patch-Indoor-C: An indoor office space was used for this set of tests. See figure A.3 for the floor-floorplan of this office space. Two receivers were used here: one with line of sight and one without line-of-sight, placed amidst desks and offices.

Array-Indoor-A and Array-Indoor-B: Seven phased array antennas are deployed in the same 25x30m indoor office space used for Parabolic-Indoor-B, Parabolic-Indoor-C, Patch-Indoor-B and Patch-Indoor-C. Data from two of the seven antennas are analyzed here. Each antenna electronically steers through its 16 directional states, spending 20 ms at each state. Two mobile omnidirectional transmitters move through the space and transmit 500 packets at 44 distinct positions. For each packet received by a phased array, the packet's transmission location and orientation is recorded (i.e., which of the four cardinal directions was the transmitter facing) along with the directional state in which the packet arrived and the RSSI value.

Parabolic-Reference and Patch-Reference: The large floodplain is used here. An Agilent VSA is connected to the omnidirectional receiver and makes a 10 second running average of power samples on a specific frequency (2.412 GHz was used). Three consecutive averages of both peak and band power are recorded for each direction. The directional transmitter is rotated in five degree increments and is connected to a VSG outputting a constant sinusoidal tone at 25 dBm on a specific frequency. Before, after, and between experiments, we make noise floor measurements, and as a postprocessing step, we subtract the mean of this value (-59.62 dBm or 1.1 nW) from the measurements.

A.3.2 Normalization

The task in comparing data sets is to come up with a scheme for normalization so that they can be compared to one another directly. For each data set, the mean peak value is determined, which is the maximum of the mean of samples for each discrete angle. This value is then subtracted from every value in the data set. The net effect is that the peak of the measurements in each data set will be shifted to zero, which allows comparison of measurements from diverse RF environments directly.

A.3.3 Error Relative to the Reference

Figure A.4 shows the normalized measured *in situ* patterns and their corresponding (also normalized) reference patterns. Recall that the reference pattern is generated and recorded by calibrated signal processing equipment and the measured data is collected using commodity 802.11 cards. There is much variation in

Label	Environment	LOS?	Dist. (m)	Samples	Loss (%)
Parabolic-Outdoor-A	Open Field on Campus	Yes	30.5	214471	24.81
Parabolic-Outdoor-B	Empty Floodplain	Yes	30.5	258876	7.05
Parabolic-Indoor-A	Laboratory	Yes	30.5	267092	2.21
Parabolic-Indoor-B	Office Building	Yes	≈ 60	268935	10.41
Parabolic-Indoor-C	Office Building	No	≈ 15	283104	5.12
Parabolic-Reference	Empty Floodplain	Yes	30.5	219	N/A
Patch-Outdoor-A	Open Field on Campus	Yes	30.5	455952	12.44
Patch-Outdoor-B	Empty Floodplain	Yes	30.5	278239	4.99
Patch-Indoor-A	Laboratory	Yes	30.5	290030	2.21
Patch-Indoor-B	Office Building	Yes	≈ 60	265593	7.40
Patch-Indoor-C	Office Building	No	≈ 15	278205	2.65
Patch-Reference	Empty Floodplain	Yes	30.5	219	N/A
Array-Outdoor-A	Open Field on Campus	Yes	≈ 30	475178	N/A
Array-Indoor-A	Office Building	Mixed	Varies	2672050	N/A
Array-Indoor-B	Office Building	Mixed	Varies	2708160	N/A
Array-Reference	Open Urban Area	Yes	≈ 5	360	N/A

Table A.1: Summary of data sets.

the measured patterns and in how much they differ from the reference (which would be typically classified as error). As might be expected, the measurements in outdoor environments exhibit less noise due to less clutter, but still deviate from the reference on occasion. As a further confirmation that the measurement process works well, notice how well Parabolic-Outdoor-B and Patch-Outdoor-B (figures A.4(b) and A.4(d)) correlate with the reference pattern (recall that these experiments were done in the same floodplain as the reference, indicating that the commodity hardware can compete with the expensive specialized equipment in a similar environment).

On inspection of this data, the first question is whether there a straightforward explanation for error in the measured patterns. Figure A.6 provides a CDF of all error for each antenna. The three antennas provide similar error distributions, although offset in the mean. The array data is the most offset from the others (presumably because its reference pattern is theoretical rather than measured) and exhibits some bimodal behavior. The patch measurements are closest to the reference, showing a large kurtosis about zero. Figure A.5 shows a PDF of error averaged at each angle—discarding outliers this way, the error between antennas begins to suggest similar distributions.

Clearly, the antennas have different error characteristics. However, for each antenna, and for each data set, it might be that the error in a given direction is correlated with that in other directions—if this were true, a single or small set of probability distributions could be used to describe the error process in a given environment with a given antenna.

A Shapiro-Wilkes test is used on the per angle error for each data set. The resulting p-values are well under the $\alpha = 0.05$ threshold, and in all cases the null hypothesis that the error is normally distributed can be rejected; this means that standard statistical tests (and regression models) that assume normality cannot be used. A pairwise Mann-Whitney U-test can be used to determine which pairs of samples grouped on some criterion (in this case, angle) are drawn from the same distribution. For each data set, a “heatmap” is generated where each cell corresponds to a pair of angles. The cell is colored by the p-value produced by the U-test when run pairwise, comparing the error for the reference pattern and the *in situ* pattern for those angles. Remarkably, all of the traces produce similar heatmaps: in the majority of pairs the null hypothesis that their error process is drawn from the same distribution can be rejected. However, for angles

near zero, this hypothesis cannot be rejected. This observation, that *measurements where the main lobe of the directional antenna is pointed at the receiver may exhibit correlated error processes*, motivated another series of tests.

To further explore "possibly well behaved" error processes about the main lobe, a Kruskal-Wallis rank-sum test was applied to two scenarios: (1) For angles near zero, are batches with the same antenna (but different environments) equivalent? (2) For angles near zero, are batches with the same environment (but different antennas) equivalent?

For (1), the null hypothesis is soundly rejected for all combinations (p-value $\ll 0.05$) For (2), the results still point strongly toward rejection (mean p-value = 0.0082), however there is one outlier—in the case of 355 degrees in the laboratory environment, a p-value of 0.2097 is achieved. One outlier, however, is not sufficient to overcome the evidence that neither antenna configuration nor environment alone is sufficient to account for intra-angle variation in error—even in the more seemingly well behaved cone of the antenna mainlobe.

A.3.4 Observations

There are several qualitative points that are worth bringing out of this data: (1) In the indoor environments, none of the measurements track the reference signal at all closely; (2) In all environments, there is significant variation between data sets; (3) The maximum signal strength is generally realized in *approximately* the direction of maximum antenna gain, but directions of low antenna gain often do not have correspondingly low signal strength. This means that *no system for interference mitigation can safely rely on predetermined antenna patterns*.

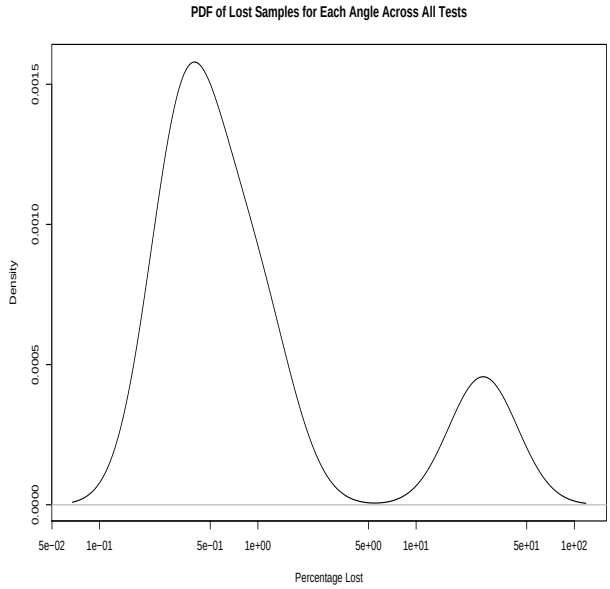


Figure A.1: Probability Density Function (PDF) of percentage of dropped measurement packets in a given angle for all angles and all data sets.

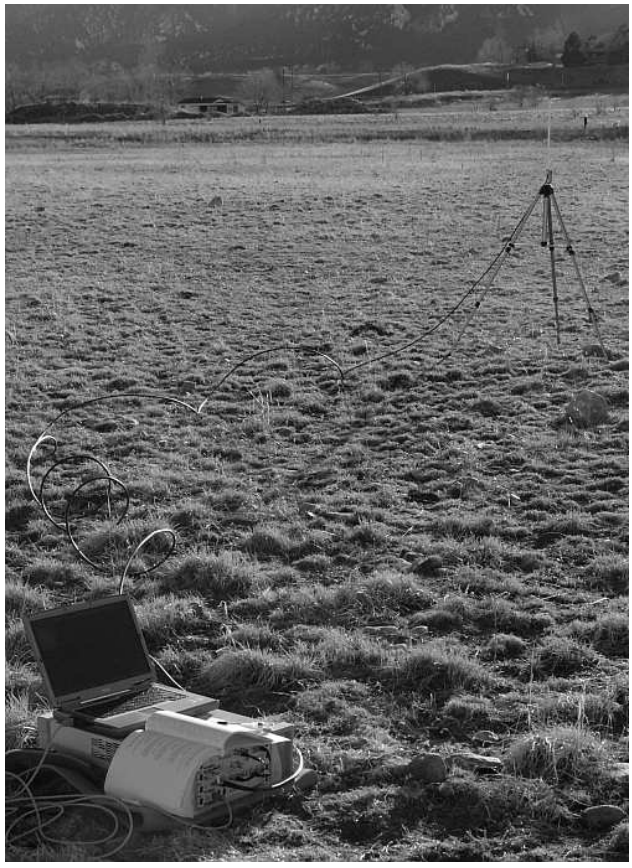
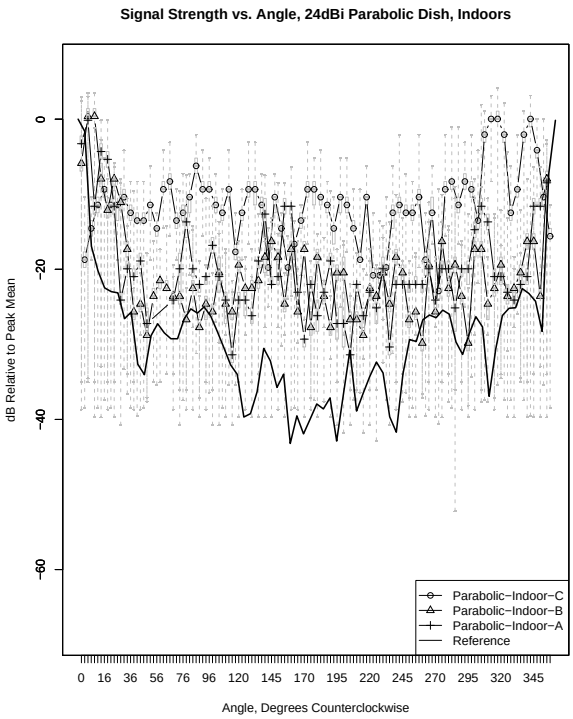
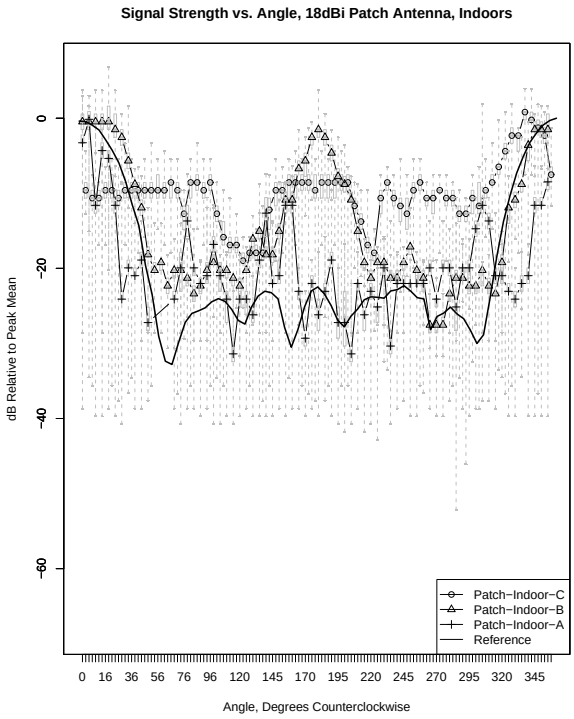


Figure A.2: Receiver side of measurement setup in floodplain

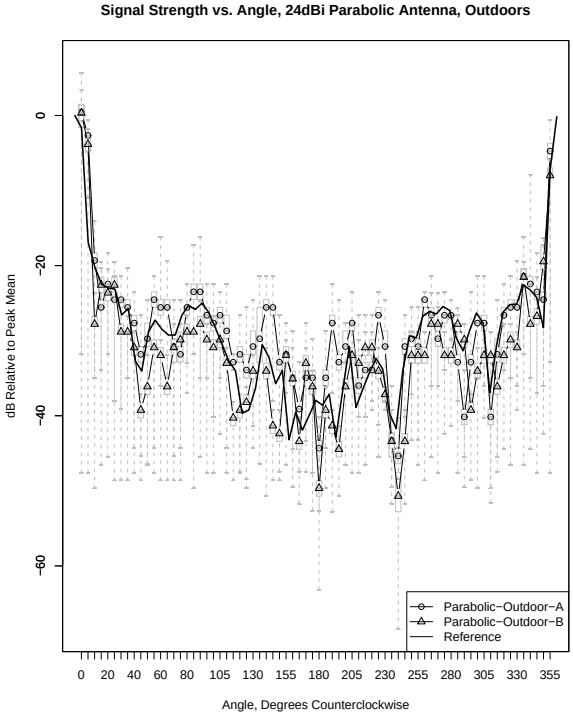
Figure A.3: Floorplan of office building used in Array-Indoor-A, Array-Indoor-B, Patch-Indoor-B, Patch-Indoor-C, Parabolic-Indoor-B, and Parabolic-IndoorC.



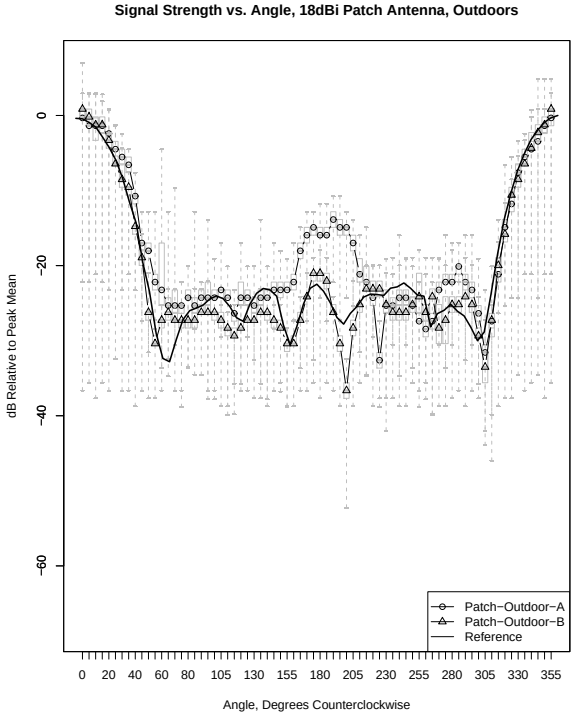
(a) Parabolic dish indoor environments



(c) Patch panel indoor environments



(b) Parabolic dish outdoor environments



(d) Patch panel outdoor environments

Figure A.4: Comparison of signal strength patterns across different environments and antennas.

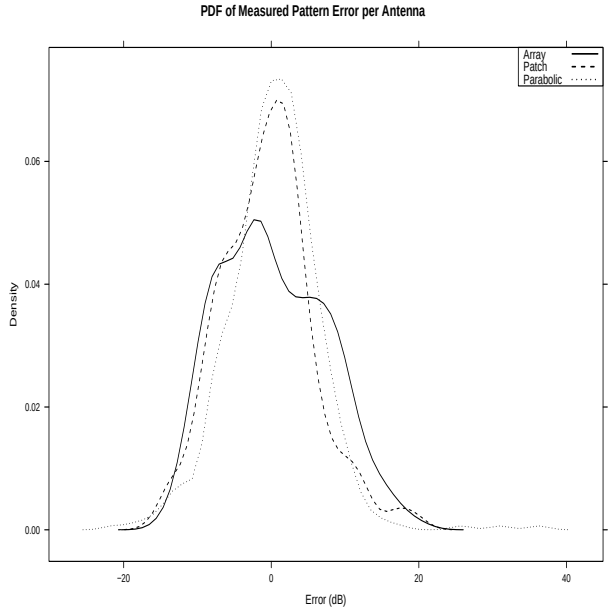


Figure A.5: Probability Density Functions (PDFs) for the averaged error process (combined across multiple traces) for each antenna type.

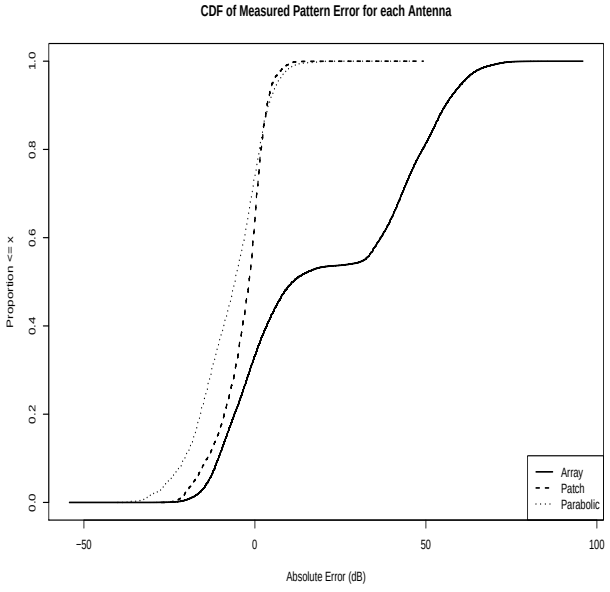


Figure A.6: Cumulative Density Functions (CDFs) for the error process (combined across multiple traces) for each antenna type.

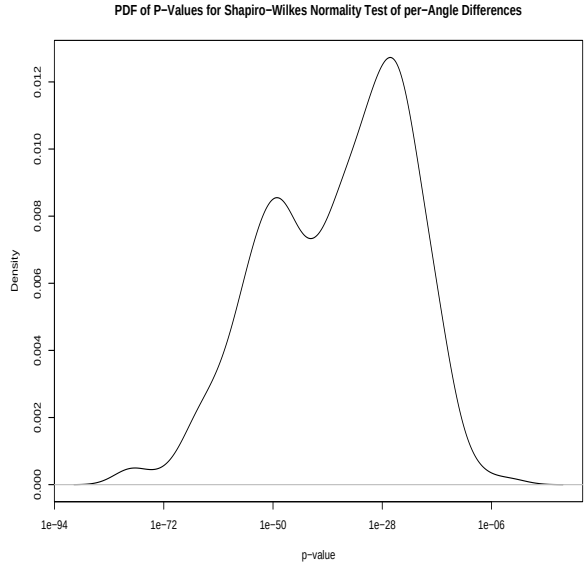


Figure A.7: Probability Density Function (PDF) of p-values from testing the normality of the error process in each direction for each data set. In all cases, the null hypothesis (that the samples are normally distributed) can be confidently rejected.

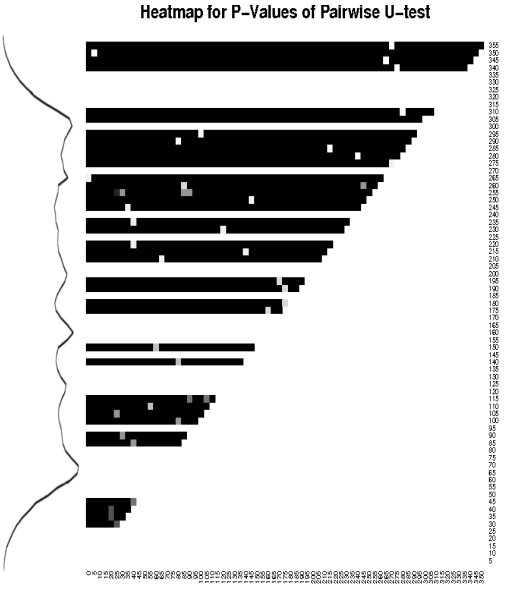


Figure A.8: Heatmap of p-values for the Mann-Whitney U-test which was run pairwise against the error from the reference pattern in each angle. This plot, which is for Patch-Indoor-A, was chosen as a representative. *All traces showed similar trends.* Darker values indicate very small p-values, meaning that the null hypothesis can be rejected with confidence. In this case, the null hypothesis is that the samples come from the same distribution. The Patch reference pattern is provided on the left for reference.

A.4 A New Model of Directionality

This chapter began with the observation that path loss and antenna gain are typically regarded as orthogonal components of the power loss between transmission and reception (equations A.1 – A.3). In this section, the *best case* accuracy of this approach is evaluated, and a new model based on the limitations identified is derived.

A.4.1 Limitations of Orthogonal Models

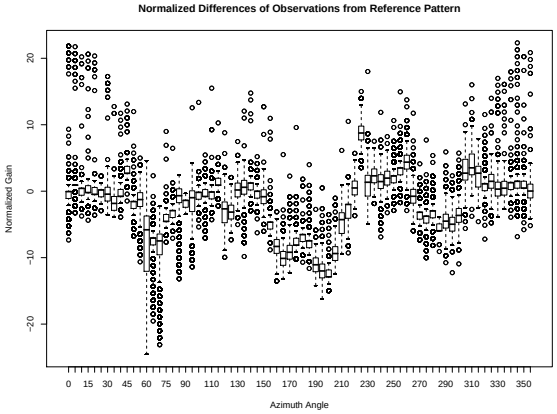
If transmit power and path loss do not vary with antenna angle, the received power relative to antenna angle can be modeled as:

$$\widehat{P_{rx}} = \beta_0 * f(\phi, \theta) \tag{A.6}$$

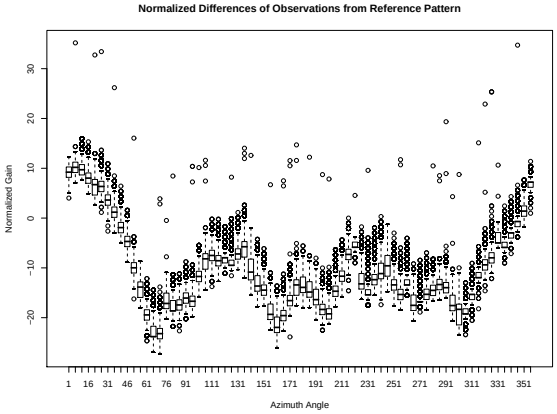
β_0 is a constant combining the path loss—however calculated—and the gain of the nonrotating antenna. $f(\phi, \theta)$ is a function describing the gain of the other antenna relative to the signal azimuth θ and zenith ϕ. Without loss of generality, an assumption is made that the antenna being varied is the receiver, and that the zenith, ϕ, is fixed.

To evaluate the accuracy of this model, the estimate b_0 for β_0 is found that minimizes the sum of squared error (SSE). In effect, this is assuming the *best possible path loss estimate*, without specifying how it is determined. If the function $f(\phi, \theta)$ correctly describes the antenna, and if path loss and antenna gain are in fact orthogonal components of the received signal strength, then the remaining error should be *randomly* distributed about 0.

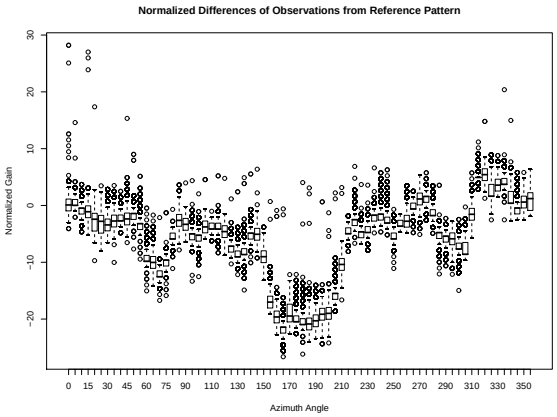
Figure A.9 depicts the error of this *orthogonal model* for several data sets. There are several qualitative observations to be made: First and most importantly, *the error value is not uniformly random, but rather correlated with direction.* The variability within any given direction is less than that for the data set as a whole. Second, the error is significant. In the worst states, the *mean error* is between 8 and 10 dB, in either direction. Third, the model overestimates signal strength in the directions where the gain is highest, and underestimates in the directions where the gain is lowest. That is, *the difference in actual signal strength between peaks and nulls is less than the antenna in isolation would produce.* This has significant



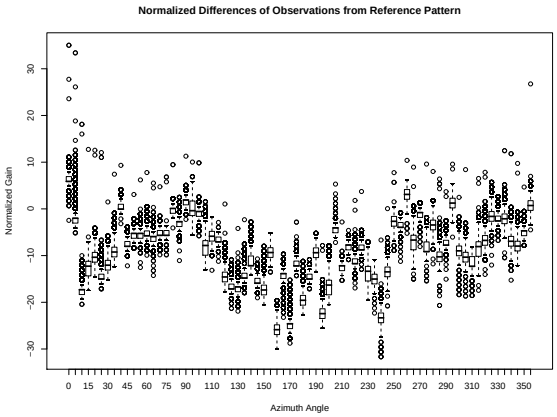
(a) Patch-Outdoor-A



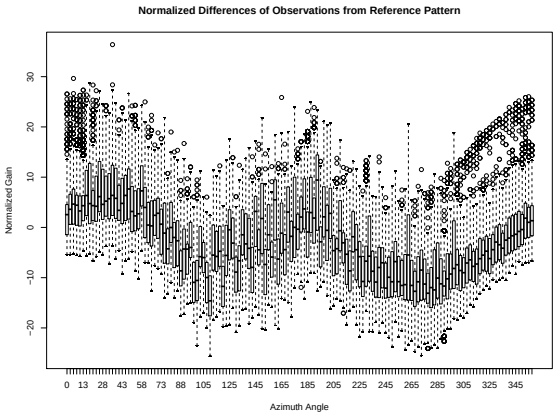
(b) Patch-Indoor-B



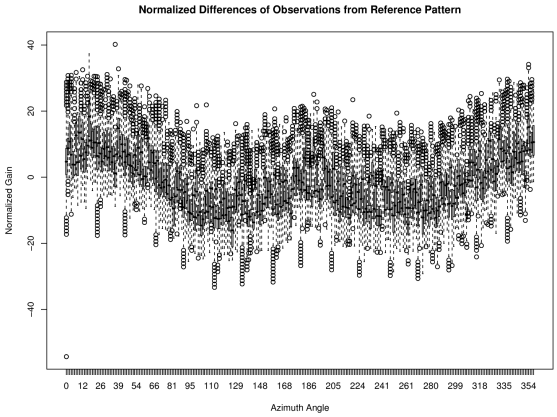
(c) Patch-Indoor-C



(d) Parabolic-Indoor-C



(e) Array-Outdoor-A



(f) Array-Indoor-A

Figure A.9: Differences between the orthogonal model and observed data in dB: $\hat{P_{rx}} - P_{rx}$.

implications for systems that use null steering to manage interference.

The data in figures A.9(e) and A.9(f) is aggregated from 36 distinct physical configurations. In each configuration, the directional receiver was (electronically) rotated in 22.5 degree increments, and between configurations, the omnidirectional transmitter was physically moved around the receiver by ten degrees. A consequence of this method is that these 10 degree changes represent not only a change of the angle between the transmitter and the antenna, but also a change of location with the attendant possibility of fading effects. To account for this, each of the 36 configurations is considered individually. This gives less angular resolution, but also fewer confounds. Figure A.10 displays each configuration as a separate line. The model accuracy is fairly consistent: The residual standard error of the aggregate is 8 dB, and the individual cases range from 5.74 dB to 11.4 dB with a mean of 7.6 dB.

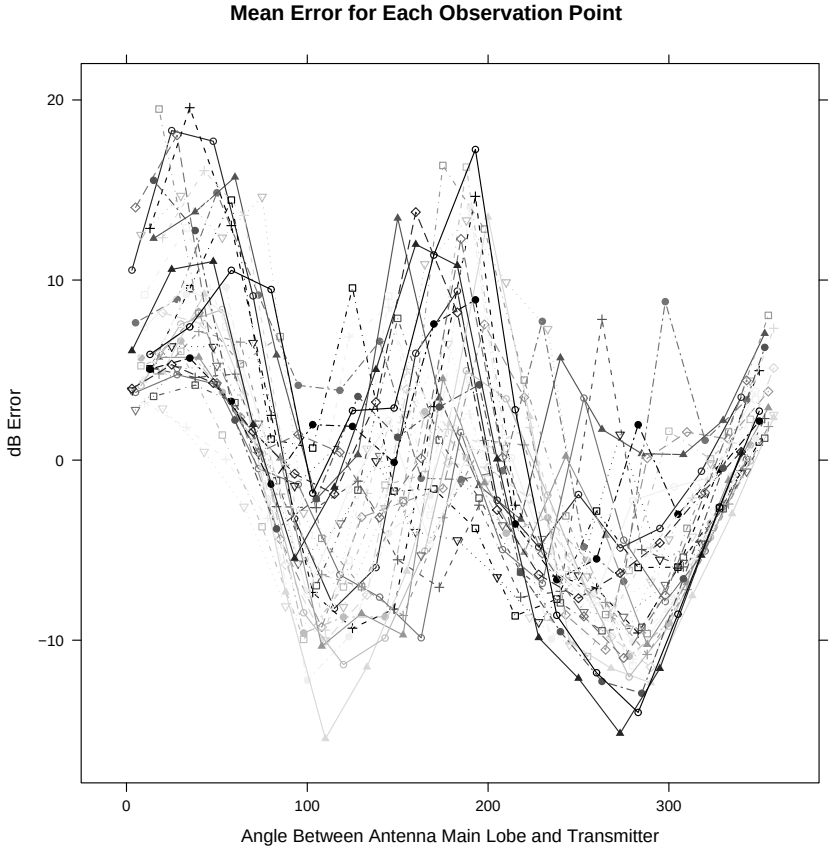


Figure A.10: Mean error of orthogonal model for each observation point in the Array-Outdoor-A data set. The format is the same as in figure A.9.

The path loss value used for each data set was the lowest error fit for that specific data, and the antenna patterns ($f(\theta)$) for the patch and parabolic antennas were measured using the specific individual antenna in question. Note also that error patterns differ from environment to environment: one could derive an *ex post facto* $f(\theta)$ to eliminate the error in a single data set, but it would not be applicable to any other.

The magnitude and *systematic nature* of the error suggest that the orthogonal model has inherent limitations that cannot be alleviated by improving either the antenna model or path loss model separately.

A.4.2 An Integrated Model

To address these limitations, an integrated model is derived that addresses the systematic errors discussed above, while remaining simple enough to use in analysis and simulations.

The *environment specific, direction specific* error shown in figure A.9 is addressed with the following environment aware model, given in equation A.7. The expected received power is given by a constant β_0, the antenna gain function $f(\phi, \theta)$, and a yet to be determined environmental offset function $x(\phi, \theta)$:

$$\widehat{P_{rx}} = \beta_0 * f(\phi, \theta) * x(\phi, \theta) \tag{A.7}$$

As with the orthogonal model, a constant zenith is assumed and $f(\phi, \theta)$ and $x(\phi, \theta)$ are considered with regard to the azimuth θ. Equation A.7 can be converted to a form that lends itself to least squares (linear regression) analysis in the following way: First, equation A.7 is rewritten as addition in a logarithmic domain, and second a discrete version of the general $x(\theta)$ is substituted in. In the discrete $x(\theta)$, the range of angles is partitioned into n bins such that bin i spans the range $[B_i, T_i)$. Each bin has associated with it a boxcar function $d_i(\theta)$ to be 1 if and only if the angle θ falls within bin i (equation A.8) and an unknown constant *offset value* β_i. These transformations yield the model given in equation A.10.

$$d_i(\theta) = \begin{cases} 1, & B_i \leq \theta < T_i \\ 0, & \text{otherwise} \end{cases} \tag{A.8}$$

$$x(\theta) = \sum_{i=1}^{n} d_i(\theta)\, \beta_i \tag{A.9}$$

$$f(\theta) - \widehat{P_{rx}} = \beta_0 + \beta_1 d_1(\theta) + \beta_2 d_2(\theta) + \cdots + \beta_n d_n(\theta) \tag{A.10}$$

If $x(\theta)$ is discretized into n bins, the model has $n + 1$ degrees of freedom: One for each bin and one for β_0, the signal strength without antenna gain. For any given signal direction, exactly one of the $d_i(\theta)$ functions will be 1, so each prediction is an interaction of two coefficients: β_0 and β_i. Consequently, β_0 could be eliminated and an equivalent model achieved by adding β_0's value to each β_i. Mathematically, this means that there are only n independent variables in the SSE fitting, and the full set is collinear. In practice, the constant β_n is dropped, but this does not mean that packets arriving in that bin are any less well modeled. Rather, one can think of bin n as being the "default" case.

The azimuth can be divided into arbitrarily many bins. The more finely it is divided, the more degrees of freedom the model offers, and thus the more closely it can be fitted to the environment. To investigate the effect of bin number, every data set is modeled using from two to twenty bins. Figure A.11 shows the residual standard error as a function of bin count. The grey box plot depicts the mean and interquartile range for all of the data collectively, and the foreground lines show values for links individually. In general, there appears to be a diminishing return as the number of bins increases, with the mean remaining nearly constant above 16 bins.

In discussing parameters for this model, we will use the 16-bin case specifically. We find the same patterns across other numbers, though the actual coefficients are bin count specific. One result of note with regard to bin count is this: Several environments exhibit a "sawtooth" pattern in which the odd bin counts do better than the even ones, or vice versa. This appears to be an effect of the *alignment* of the bins relative to environmental features, rather than the *number* of bins as such.

The model described here has significantly less error than the orthogonal model: Across all data sets, the mean residual standard error is 4.0 dB, *(4.4dB indoors)* compared to 6.15 dB *(7.312 dB indoors)* for the orthogonal model. More importantly, the error remaining in the discrete offset model is largely noise: The mean error is almost exactly zero for several ways of grouping the data. Figure A.12 depicts the error (predicted value minus observed value). While the outliers reveal some direction correlated effect that is not accounted for, this model is much better for the bulk of the traffic. Over 99.9% of the traffic *at every angle* falls within the whisker interval.

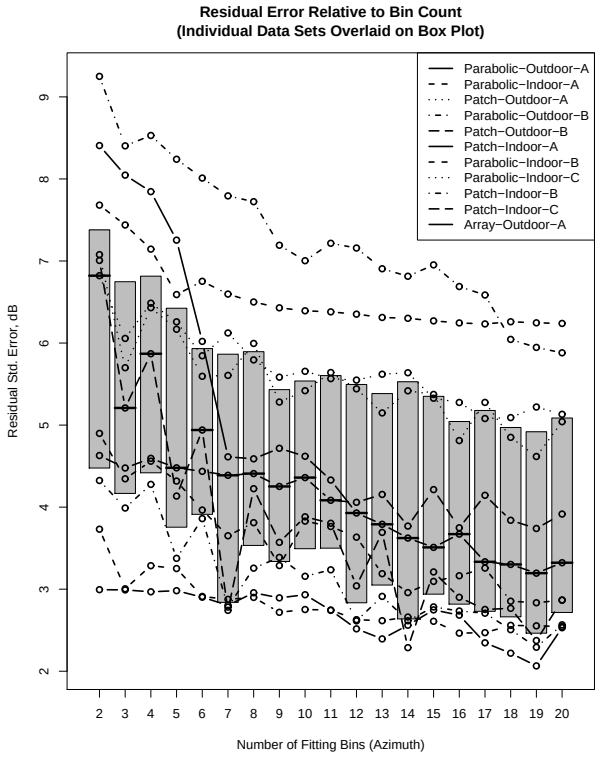


Figure A.11: Effect of increasing bin count (decreasing bin size) on modeling precision.

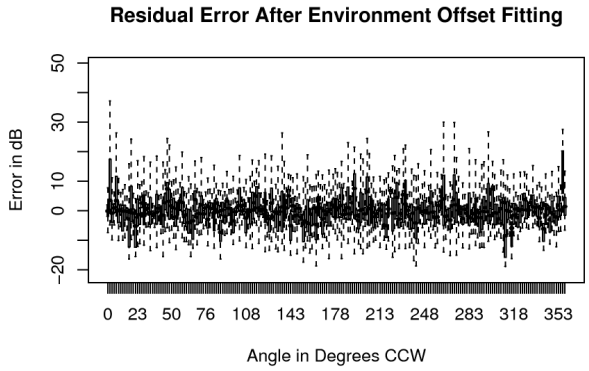


Figure A.12: Residual error of the discrete offset model with 16 bins.

A.4.3 Describing and Predicting Environments

The environmental offset function $x(\phi, \theta)$, or its bin offset counterpart, models the impact of a particular environment combined with a particular antenna. This can serve as an *ex post facto* description of the environment encountered, but it also has predictive value: If one knows the offset function for a given environment, it is possible to more accurately model wireless systems in that environment. There is no practical way to know the exact spatial RF characteristics of an environment—and thus its offsets—without actually measuring it. However, these results suggest that it is possible to identify parameters generating the *distribution* from which the offset values for a *class of environments* are drawn.

An analysis of possible determining factors for the fitted offsets was conducted across all traces and a range of bin counts. A linear regression fit and ANOVA test found significant correlation with two factors: The nominal antenna gain $f(\theta)$ and the observation point; none of the other factors examined were consistently significant. Table A.2 shows the regression coefficients and P-values for both factors for a variety of traces. The observation angle was always statistically significant, but the coefficient is constantly near zero. For each factor, the regression coefficient describes the correlation between the fitted offset and the factor. That is, the coefficient shows how much the actual signal strength can be expected to differ from the orthogonal model, for any value of that factor. For example, the antenna gain coefficients of .668 and .703 for Parabolic-Indoor-C and Patch-Indoor-C mean that in those data sets for every dB difference in antenna gain between two angles, *the best fit difference in actual signal strength is only* $\approx$ *0.3 dB.*

There are two key results pertaining to the antenna gain regression coefficient: First, the coefficients for different antennas in the same environment are very close. Second, the coefficients for distinct but similar environments are fairly close. This suggests that classes of environments can reasonably be characterized by their associated coefficients, which provides a compact representation of environment classes that lends itself easily to simulation. In this way, the task of the researcher is reduced to choosing amongst several *representative environment classes* when designing their experiment.

Data Set	Factor	Coefficient	P-value
Parabolic-Outdoor-A	Antenna Gain	0.185	1.02e-87
	Obs. Angle	0.00301	5.1e-06
Patch-Outdoor-A	Antenna Gain	0.146	6.4e-50
	Obs. Angle	0.00744	1.14e-17
Array-Outdoor-A	Antenna Gain	0.41	2.03e-206
	Obs. Angle	-0.0271	5.36e-188
Parabolic-Outdoor-B	Antenna Gain	0.0377	8.68e-05
	Obs. Angle	-0.00323	5.95e-05
Patch-Outdoor-B	Antenna Gain	0.00919	0.0492
	Obs. Angle	-0.00198	3.08e-06
Parabolic-Indoor-A	Antenna Gain	0.33	4.6e-102
	Obs. Angle	0.00463	1.91e-05
Patch-Indoor-A	Antenna Gain	0.258	1.22e-122
	Obs. Angle	0.00894	3.09e-24
Parabolic-Indoor-B	Antenna Gain	0.378	2.2e-134
	Obs. Angle	0.00971	1.97e-16
Patch-Indoor-B	Antenna Gain	0.372	1.1e-81
	Obs. Angle	0.014	3.87e-18
Parabolic-Indoor-C	Antenna Gain	0.668	1.39e-234
	Obs. Angle	-0.0146	4.15e-36
Patch-Indoor-C	Antenna Gain	0.703	0
	Obs. Angle	-0.0154	2.63e-48

Table A.2: Factors influencing fitted offset values, 16-bin case.

A.5 Simulation Process

The statistical model laid out above can be used as the basis for more realistic simulations. It has long been recognized that radio propagation involves very environment specific effects. Three major ways of addressing such effects in modeling and simulation have been identified: The first is to simply ignore the variability and use a single representative value in all cases. The second, which goes to the opposite extreme, is to model specific environments in great detail. A third approach is to randomly generate values according to a representative process and perform repeated experiments.

Each approach has its benefits, but this chapter advocates the repeated sample approach. Precisely modeling a specific environment probably has the greatest fidelity, but it provides no information as to how well results achieved in a single environment will generalize to others. Stochastic models have the advantage of being able to produce arbitrarily many "similar" instances, and parametric models make it possible to study the impact of varying a given attribute of the environment. Such approaches are frequently used to model channel conditions [151], network topology [247, 218], and traffic load [129].

The following algorithms produce signal strength values consistent with our statistical findings. The key parameters are the gain offset correlation coefficient K_{gain}, the offset residual error S_{off}, and the per packet signal strength residual error S_{ss}. These values were computed across many links for two types of environments in sections A.4.3 and A.4.2. Table A.3 summarizes these results.

Environment	K_{gain}	S_{off}	S_{ss}
Open Outdoor	0.01 - 0.04	1.326 - 2.675	2.68 - 3.75
Urban Outdoor	0.15 - 0.19	2.244 - 3.023	2.46 - 2.75
LOS Indoor	0.25 - 0.38	2.837 - 5.242	2.9 - 5.28
NLOS Indoor	0.67 - 0.70	3.17 - 3.566	3.67 - 6.69

Table A.3: Summary of Data Derived Simulation Parameters: Gain-offset regression coefficient (K_{gain}), offset residual std. error (S_{off}), and signal strength residual std. error (S_{ss}).

Algorithm 5 is a one time initialization procedure which computes the offsets between the antenna gain in any direction and the expected actual signal gain.

Algorithm 6 computes the expected end to end gain for a given packet, *not including fixed path loss*. Thus, the simulated signal strength would be determined by the transmit power, path loss, receiver gain,

Algorithm 5 Compute direction gain

1: $K_{gain} \leftarrow$ gain offset correlation coefficient
2: $S_{off} \leftarrow$ offset residual std. error
3: **procedure** DIRECT-GAIN
4: **for** Node $n \in$ all nodes **do**
5: $P \leftarrow$ partition of azimuth range $[-\pi, \pi)$
6: **for** $p_i \in P$ **do**
7: $\theta_i \leftarrow$ center angle of p_i
8: $X \leftarrow$ random value from $(\mu = 0, \sigma^2 = S_{off})$
9: $o_{n,p_i} \leftarrow K_{gain} * f_n(\theta_i) + X$
10: **end for**
11: **end for**
12: **end procedure**

fading model (if any) and the directional gain from algorithm 6. Note that a fading model that accounts for interpacket variation for stationary nodes might make the random error ϵ in line 9 redundant.

A.6 Summary and Conclusion

This chapter has presented an empirical study of the way different environments and antennas interact to affect the directionality of signal propagation. The three primary contributions of this work are:

(1) A well validated method for surveying propagation environments with inexpensive commodity hardware.

(2) A characterization of several specific environments ranging from the very cluttered to the very open.

(3) New, more accurate, techniques for modeling and simulating directional wireless networking.

Wireless signal—and interference—propagation is more complicated than common previous models have acknowledged. Because models of the physical layer guide the development and evaluation of higher layer systems, it is important that these models describe reality well enough. Indeed, [30] shows that application layer results reported by simulators can be affected dramatically by the way directional antenna models are simulated, producing results that deviate significantly from reality. The measurements described here, and the resulting model, bring to light several important aspects of the physical environment that previous models have failed to capture. The *effective* directionality of a system depends not only on the antenna,

Algorithm 6 Compute per-packet gain

$S_{pss} \leftarrow$ residual error of packet signal strengths
function DIRECTIONAL-PACKET-GAIN(src, dst)
    $\theta_{src} \leftarrow$ direction from src toward dst
    $\theta_{dst} \leftarrow$ direction from dst toward src
    $p_{src} \leftarrow$ partition at src containing θ_{src}
    $p_{dst} \leftarrow$ partition at dst containing θ_{dst}
    $G_{src} \leftarrow f_{src}(\theta_{src}) - o_{src,p_{src}}$
    $G_{dst} \leftarrow f_{dst}(\theta_{dst}) - o_{src,p_{dst}}$
    $\epsilon \leftarrow$ random value from $(\mu = 0, \sigma^2 = S_{pss})$
    return($G_{src} + G_{dst} + \epsilon$)
end function

but is influenced by the environment to such a large extent that many decisions cannot be made without *in situ* measurements.