The Efficacy of Path Loss Models for Fixed Rural Wireless Links

Caleb Phillips[1], Scott Raynel[2], Jamie Curtis[2], Sam Bartels[2], Douglas Sicker[1], Dirk Grunwald[1], and Tony McGregor[2]

[1] Computer Science Department
University of Colorado, Boulder, USA
[2] Computer Science Department
University of Waikato, New Zealand

Abstract. In this paper we make use of a large set of measurements from a production wireless network in rural New Zealand to analyze the performance of 28 path loss prediction models, published over the course of 60 years. We propose five metrics to determine the performance of each model. We show that the state of the art, even for the "simple" case of *rural* environments, is surprisingly ill-equipped to make accurate predictions. After combining the best elements of the best models and hand-tuning their parameters, we are unable to achieve an accuracy of better than 12 dB root mean squared error (RMSE)—four orders of magnitude away from ground truth.

1 Introduction

Modeling the propagation of a wireless transmitter in a complex environment has entertained scientists for at least sixty years. The result is a staggering number of proposals of just about every shape, size, and approach imaginable. The basis for this level of interest is solid—predicting the attenuation of transmitted signals with high precision has very important applications in the design, trouble-shooting, and simulation of wireless systems.

Despite the large quantity of work done, we recognize an important shortcoming: there have been relatively few comparative evaluations of path loss prediction models using a sufficiently representative dataset as a basis for evaluation. Those studies that do exist make comparisons between a small number of similar models. And, where there has been substantial work of serious rigor done, for instance in the VHF bands where solid work in the 1960's produced well validated results for analog television (TV) propagation, it is not clear how well these models work for making predictions outside their intended coverage (i.e., frequency, distance, environment type, etc.). The result is that wireless researchers are left without proper guidance in picking among dozens of propagation models from which it is not clear which is best or what the penalty is of using a model outside of its intended coverage. This work provides a first step towards solving that problem.

In this paper, we describe, implement, and analyze 28 propagation models spanning 60 years of publications using five metrics to gauge performance. Although many of these models are massively different from one another, they all make use of the same basic variables on which to base their predictions: position (including height and orientation) of the transmitter and receiver, carrier frequency, and digital elevation model and land cover classification along the main line-of-sight (LOS) transmit path. These models are a mix of approaches: empirical, (purely) analytical, stochastic or some combination thereof. In the present study, we are not including ray-tracing models (e.g., [11]) or partition based models (e.g., [5]) which require substantial knowledge of the environment which is seldom available at all, and rarely at the precision required to make useful predictions. We are also not considering active-measurement models (e.g., [8]) which make use of in-situ measurements to correct their predictions. We expect to analyze these more complex models in later work.

To perform our evaluation we use a large set of active measurements collected from a production wireless network on the northern isle of New Zealand. This network spans approximately 8300 square kilometers, containing more than 368 transceivers (with 1328 possible links, 1246 of which are under measurement), and provides Internet connectivity to more than 740 clients. The network is built using commercial off-the-shelf equipment (COTSE) and operates in the popular bands of unlicensed spectrum at 2.4 and 5.8 GHz. All of the measurements we use will be released to the community to enable comparative evaluations.

2 Related Work

The vast majority of existing work analyzing the efficacy of path loss models has been carried out by those authors who are proposing their own improved algorithm. In such cases, the authors collect data in an environment of interest and show that their model is better able to describe this data than one or two competing models. Unfortunately, this data is rarely published to the community, which makes comparative evaluations impossible. One noteworthy exception is the work of the COST-231 group in the early 1990's, which published a benchmark dataset (900 MHz measurements taken in European cities) [3]. This effort produced a number of well-validated models which are tuned for 900 MHz transmitters in urban environments. We consider all of the proposed COST-231 models in our analysis here. The COST-231 data, being collected in an urban environment, is inappropriate for our present work, but we expect to use it in future work.

There are several studies similar to our own that compare a number of models with respect to some data. In [4], the authors compare five models with respect to data collected in rural and suburban environments with a mobile receiver at 910 MHz. They discuss the abilities of each model, but abstain from picking a winner. In [1], the authors compare three popular models to measurements collected at 3.5 GHz. The authors highlight the best of the three, which turns out to be the ECC-33 model proposed in [6]. In [9], Sharma et al. do a very similar analysis, but instead focus on measurements made in India at 900 and 1800 MHz. In contrast to [1], they find that the SUI and COST-231 models perform best. We believe our work here is the first to do an in-depth and rigorous analysis of a *large number of diverse propagation models* using a *large and realistic dataset from a production network*. And, it is the first such comparative study looking at results for the widely used 2.4 and 5.8 GHz bands.

3 Measurement

The network used in our study is a large commercial network that provides Internet access to primarily rural segments of the Waikato region in New Zealand. Every two minutes, each device on the network transmits a measurement frame at each supported bit-rate. For this study we only use measurements from the lowest bit-rate for each protocol (1 Mbps for 802.11b/g and 6 Mbps for 802.11a). Meanwhile, each device uses a monitor mode interface to log these measurement frames.

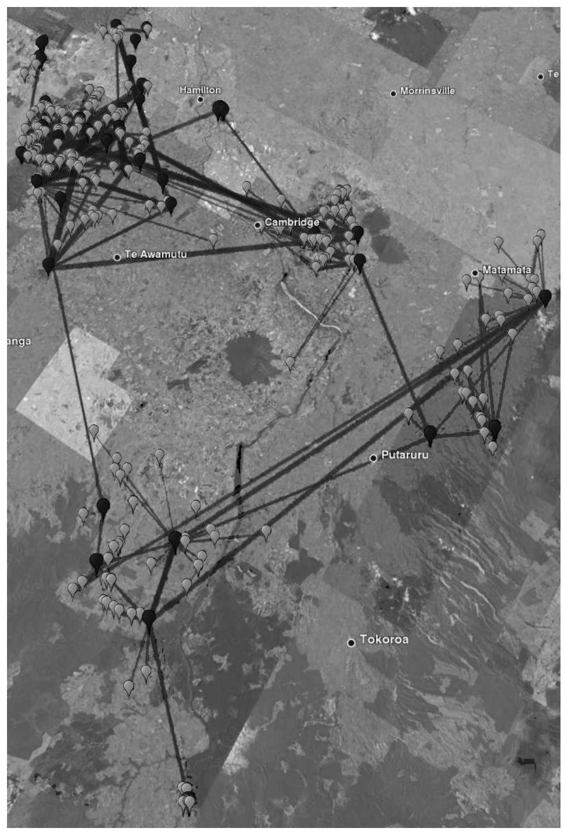

Fig. 1: The largest of three disconnected sections of the network (80x100km). Link width indicates strength. Back-haul nodes (mainly 5.8 GHz) are dark/black and CPEs are light/white.

The back-haul network is composed of long distance 802.11a links operating at 5.8 GHz[3]. These are commonly point-to-point links that use carefully steered highly directional antennas. The local access network is composed of predominantly 802.11b/g links which provide connectivity to client premise equipment (CPEs). Often, an 802.11g access point with an omnidirectional or sector antenna will provide access to a dozen or more CPE devices which have directional (patch panel) antennas pointing back to the access point. With few exceptions, each node in the network is an embedded computer running the Linux operating system which allows the use if standard open-source tools to perform measurement and monitoring. All nodes under measurement use an Atheros-brand radio and the MadWifi driver is used to collect frames in monitor mode and record received signal strengths. In [2], we showed that this hardware is able to measure signal strength at a sufficient accuracy for path loss modeling.

After collection, the data requires fairly substantial scrubbing. We discard any frame that arrives with its checksum in error or those from a source that produces less than 100 packets. The remaining packets are used as an oracle to analyze the performance of the propagation models. For this particular analysis we use one week of data collected between July 25th, 2010 and August 2nd, 2010. Because detailed documentation about each node simply did not exist prior to our study, some assumptions were made for analysis. The locations of nodes for which there is no specific GPS reading are either hand-coded, or in the case of some CPEs, geo-coded using a street address. Antenna orientations for directional antennas are assumed to be ideal—pointing in the exact bearing of their mate. All nodes are assumed to be positioned 3m off the ground, which is roughly correct for the vast majority of nodes. While these assumptions are not perfect, and are clearly a source of error, we feel that they are as accurate as is feasible for a network of this size and complexity. Certainly, any errors in antenna heights, locations, or orientations are on the same scale as those errors would be for anyone using one of the propagation models we analyze to make predictions about their own network.

In the end, our scrubbed data for a single week constitutes 19,235,611 measurements taken on 1328 links (1262 802.11b/g links at 2.4 GHz and 464 802.11a links at 5.8 GHz) from 368 participating nodes. Of these nodes, the vast majority are clients and hence many of the antennas are of the patch panel variety (70%). Of the remaining 30%, 21% are highly-directional point-to-point parabolic dishes, 4.5% are omnidirectional, and 4.5% are sector antennas. We believe this dataset is of sufficient scope and diversity to justify the claim that it is representative of a large class of wireless networks which have similar characteristics and operating frequency.

4 Models

Table 1 provides details of the models evaluated in this study. We subdivide models into five categories: *Foundational* models, which are purely theoretical and (often) form the core of more advanced models, *Basic* models, which are the majority and typically include empirical corrections from measurements and often require special tuning parameters for the environment type, *Terrain* models, which expand on the basic models by including terrain features into their calculations, and *Supplementary* models, which are not able to stand on their own but instead are used to make corrections to existing models.

At a high level, a model's task is to predict the value of $L_t + L_s$ in this log-domain equation:

[3] Atypically liberal power regulations in New Zealand and Australia around 5.8 GHz allow for much longer links than can be seen in most other places in the world.

Name	Short-Name	Category	Year
Friis' Freespace	friis	Foundational	1946
Egli	egli	Basic	1957
Hata-Okumura	hata	Basic	1968
Edwards-Durkin	edwards	Basic/Terrain	1969
Alsebrook-Parsons	alsebrook	Basic/Terrain	1977
Blomquist-Ladell	blomquist	Basic/Terrain	1977
Longley-Rice Irregular Terrain Model (ITM)	itm	Terrain	1982
Walfish-Bertoni	bertoni	Basic	1988
Flat-Edge	flatedge	Basic	1991
COST-Hata/Cost-231	cost231	Basic	1993
Walfish-Ikegami	walfish	Basic	1993
Two-Ray (Ground Reflection)	two.ray	Foundational	1994
Hata-Davidson	davidson	Basic	1997
Erceg-Greenstein	erceg	Basic	1998
Directional Gain Reduction Factor (GRF)	grf	Supplementary	1999
Rural Hata	rural.hata	Basic	2000
ITU Terrain	itu	Terrain	2001
Stanford University Interium (SUI)	sui	Basic	2001
Green-Obaidat	green	Basic	2002
ITU-R/CCIR	itur	Basic	2002
ECC-33	ecc33	Basic	2003
Riback-Medbo	fc	Supplementary	2006
ITU-R 452	itur452	Terrain	2007
IMT-2000	imt2000	Basic	2007
deSouza	desouza	Basic	2008
Effective Directivity Antenna Model (EDAM)	edam	Supplementary	2009
Herring Air-to-Ground	herring.atg	Basic	2010
Herring Ground-to-Ground	herring.gtg	Basic	2010

Table 1: Models Studied along with their categorization, citation, and year of (initial) publication.

$$P_r = P_t - (L_t + L_s + L_f(t)) \tag{1}$$

Where P_r and P_t are the received and transmitted power and the total path loss is the sum of L_t, the trivial free-space path loss, L_s, the loss due to shadowing/slow-fading from large unmoving obstacles like mountains and buildings, and $L_f(t)$, the small-scale/fast fading due to destructive interference from multipath effects and small scatterers (which varies with time t). Models cannot, without perfect knowledge of the environment, be expected to predict the quantity $L_f(t)$. In most applications, this additional error is computed "stochastically" using a probability distribution. For the protocols used in our study, however, this quantity tends to be small due to the averaging effect of wideband modulation schemes [10].

It is worth noting that among the models we study, only very few were designed with the exactly sort of network we are studying in mind. Indeed, some are very specific about the type of environment in which they are to be used. In this work, we pay little attention to these coverage requirements because

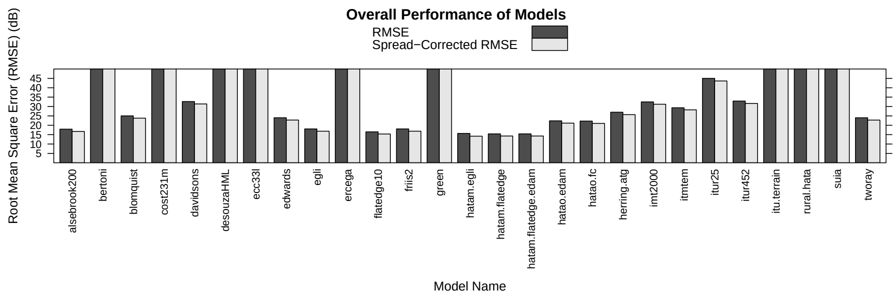

Fig. 2: Overall model performance as described by (residual) root mean squared error (RMSE) and spread-corrected RMSE (SC-RMSE). Spread corrected error is adjusted (reduced) by the expected measurement spread on a given link.

we observe that they are not largely followed in the literature (the Longley-Rice Irregular Terrain model, in particular, is frequently used well outside of its intended coverage). In this study both appropriate and "inappropriate" models are given an equal chance at making predictions for our network. We have no starting bias about which should perform best.

5 Results

To obtain results, we ask each model to offer a prediction of median path loss for each link in our network. The model produces an estimate of the loss $\hat{L}$ which we combine with known values to calculate the predicted received signal strength P_r:

$$P_r = P_t + G_t(\theta) + G_r(\phi) - \hat{L} \tag{2}$$

Where G_t is the antenna gain of the transmitter in the azimuthal direction (θ) of the receiver and G_r is the antenna gain of the receiver in the azimuthal direction (ϕ) of the transmitter. These gains are drawn from measured antenna patterns (one for each type of antenna)[2]. The transmit power (P_t) is set to 18 dBm for all nodes, which is the maximum transmit power of the Atheros radios our nodes use. For a given link, we calculate the median received signal strength value across all measurements ($\bar{P_r}$). Then, the prediction error, ϵ, is the difference between the prediction and the median measured value: $\epsilon = \bar{P_r} - P_r$.

Some models come with tunable parameters of varying esotericism. For these models, we try a range of reasonable parameter values without bias towards which we expect to be best. To conserve space, in the following discussion and figures we show results from only the 27 best performing models/configurations.

Figure 2 provides the overall performance of each algorithm in terms of its RMSE. To account for underlying variance in the measurements, we use

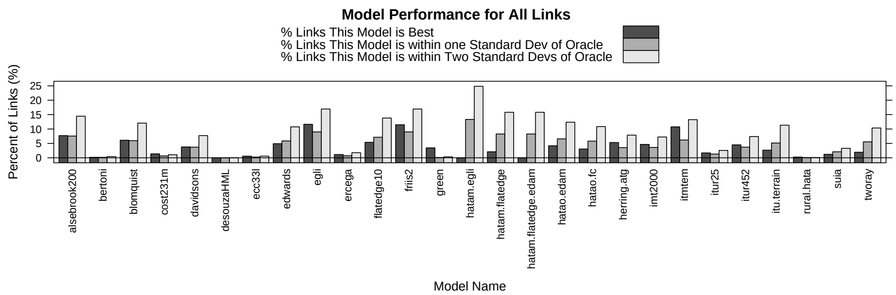

Fig. 3: Competitive and Individual Performance. Competetive performance is the percentage of links a given model is the best predictor for. Individual performance is the percentage of links a model makes a prediction within one (or two) standard deviations of the correct value.

a "spread corrected" RMSE ($\hat{\epsilon}$) where the link's measured standard deviation ($\bar{\sigma}$) is subtracted from the prediction error: $\hat{\epsilon} = |\epsilon| - \bar{\sigma}$. This corrected RMSE gives an idea of error in excess of expected variance due to temporal variation (i.e., fast-fading and intrinsic/diurnal periodicity)[4]. As we can see, the best performing models achieve an RMSE on the order of 15 dB. The best models are the Alsebrook model (with its terrain roughness parameter set to 200m) at just under 18 dB RMSE (16.7 dB when corrected), and the Flat-Edge model (with 10 "buildings" presumed) at 16.5 dB RMSE (15.3 dB when corrected)[5].

Figure 3 provides two domain-oriented metrics that describe models' competitive and individual "goodness". The competitive metric is the percentage of links that a given model produces the best prediction for (and hence sums to 100). We can see that no given model dominates the competition—the honor of best prediction is spread fairly evenly among half a dozen models that each achieve the best prediction between 10 and 15 percent of the time. The other metric is an individualistic definition of success—the percentage of links a given model's prediction is within the expected spread (measurement standard deviation). The best performing models are "correct" 10% of the time using this metric. If we lower the bar to making a prediction within two standard deviations of the measured median value, the best performing models (Egli, Friis (with $\alpha = 2$), Flat-Edge, ITM, ITU Terrain, and Two-Ray) achieve between 10 and 15% correct.

Figure 4 plots our next metric: ability to order links. In some applications it may be sufficient for a propagation model to order two or more links by

[4] Although we are careful to correct for this measurement variation, it is on the whole rather small: 1.31 dB median standard deviation and 1.67 dB at the third quantile.

[5] Some models perform substantially better when we consider only the fraction of cases that are in their intended coverage. The ITM, for instance, has a competitive spread-corrected RMSE of 17.3 dB when only error-free predictions are considered.

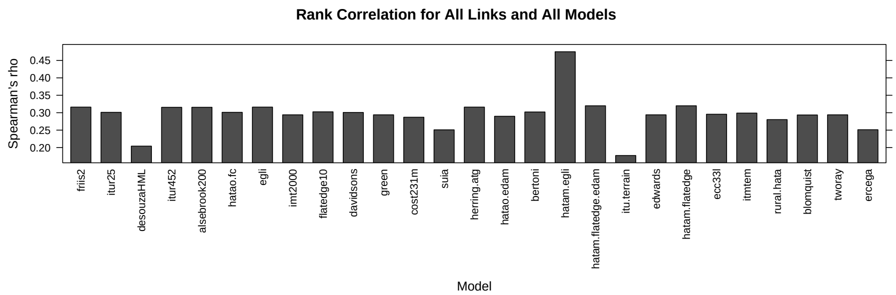

Fig. 4: Ability to order links, computed using Spearman's ρ. A value of 0 indicates a random ordering (relative to the oracle order) and a value of 1 would be a perfect ordering.

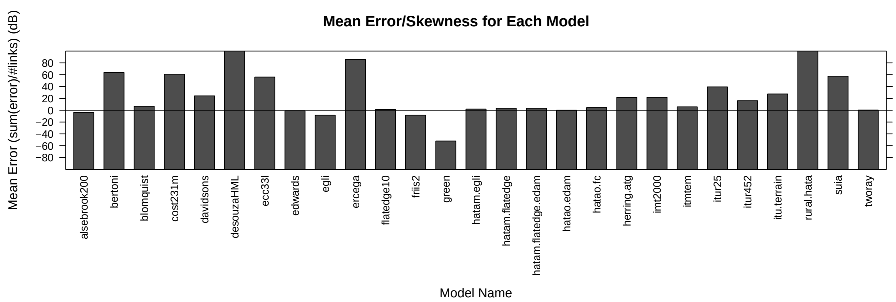

Fig. 5: Prediction Error Skewness, computed as the sum of error divided by the number of total links. Models that make an equal amount of over and under predictions acheive a value near zero. Models that make a majority of under or over predictions have a large negative or positive value respectively.

strength. In this scenario, we imagine that the predicted path loss isn't itself expected to be absolutely correct, but instead simply a relational performance compared to other links in the same network. In this figure, we plot Spearman's non-parametric rank order coefficient ρ for each model. For this metric, a value of zero indicates no correlation (random order) and a value of 1.0 or -1.0 indicates perfect positive or negative correlation. We can see that with few exceptions, all models score in the neighborhood of 0.25 to 0.30 indicating a small positive correlation. The best model (hatam.egli) performs around 0.45 and the worst model (itu.terrain) acheives less than 0.20 correlation.

Our final metric is skewness, which is shown in figure 5. For many applications an over or under estimate of path loss can come with a high price. This metric plots the sum of all residual error for each model. A model that makes an equal amount of over and under estimations should produce a skewness of

0. A model that systematically over-predicts path loss (i.e., under-predicts the received signal strength at sites) will have a large positive value and a model that systematically under-predicts path loss will have a large negative value.

We see that even in the mean case, the best models, with their best parameter settings cannot achieve an error of less than 15 dB—five orders of magnitude from the correct value! Even our more permissive performance metrics show the models are unable to widely succeed at seemingly simple tasks of rank-ordering links, or making predictions within two standard deviations of the measured value. This raises the question: is there some common source of error that is affecting all models?

To answer this question, we analyzed the covariance (correlation) between "best prediction error" (the error of the best prediction from all models) and various possible factors. We found no significant correlation between carrier frequency (and therefore neither modulation scheme nor protocol) or antenna geometry. We did however find that link distance is significantly correlated with error for a large number of models. This makes sense: many models were designed with particular lengths of links in mind and we are using them outside of their coverage in this study. It also raises the question: can a hybrid model which uses one of two or more other models at different link lengths produce a model which is better performing than any single model alone?

To answer this question, we implemented two hybrid models. The first uses the Hata model (for medium cities) for links under 500m (where it is well performing) and the Flat-edge model (with 10 "buildings") for longer links (hatam.flatedge10). This model performs marginally better than all other models, producing a corrected RMSE of 14.3 dB. Very slightly better performance is achieved by combining the Hata model with the Egli Model (14.2 dB RMSE).

It is interesting to note that in our analysis the best performing models would *not* typically be chosen for this environment. The two best performing individual models are Flat-Edge and Alsebrook. The Flat-Edge model attempts to calculate the path loss after the signal diffracts over some number of interfering "screens". Here, we pick 10 as the number of screens and obtain decent results, better in fact than the models which take the true terrain profile into account when they make predictions. The Alsebrook model is a simple plane-earth (two-ray) model with some corrections from measurements and an optional static correction for terrain "roughness". In the version that performs best for our measurements, we arbitrarily set the terrain "roughness" to 200m and the "street width" and average "building height" to the suggested default values of 5 and 20m. Perhaps comporting with Occam's Razor, the simplest models (Friis, Egli, Two-Ray) are often as well performing and in many cases better performing than the more complex models with respect to our metrics.

6 Conclusion

Overall our results show that even with the best models, hand-tuned for our environment, we can expect an RMSE in excess of 12 dB (4 orders of magnitude from correct and a far cry from the 3 dB repeated-measures variation which we treat as the gold standard [7])—a result that precludes use in all but the least demanding applications. More forgiving performance metrics show similarly bleak results: no model is able to obtain better than 25% of predictions within two standard deviations of the true value and the best models are typically 20% wrong when it comes to placing links in an order relative to all other links. We have also shown that picking a "good looking" model at random from the literature and applying it to a new (or even seemingly congruent) domain is a precarious task which can produce substantially wrong predictions. Given this, we believe attempts to model path loss in even more complex environments, such as indoors, are premature. Instead, we advocate a renewed focus on rigorous cross validation using publicly available data sets. We also caution users of these models to be wary of their predictions and to do in-situ validation whenever possible. In future work we expect to explore more complex models for path loss prediction such as those that make use of active correction from measurements (e.g., [8]).

References

1. Abhayawardhana, V., Wassell, I., Crosby, D., Sellars, M., Brown, M.: Comparison of empirical propagation path loss models for fixed wireless access systems. In: VTC 2005-Spring. vol. 1, pp. 73–77 (May 2005)
2. Anderson, E., Phillips, C., Sicker, D., Grunwald, D.: Modeling environmental effects on directionality in wireless networks. In: 5th International Workshop on Wireless Network Measurements (WiNMee) (2009)
3. Cichon, D.J., Kürner, T.: Digital mobile radio towards future generation systems: Cost 231 final report. Tech. rep., COST European Cooperation in the Field of Scientific and Technical Research - Action 231 (1993)
4. Delisle, G.Y., Lefévre, J.P., Lecours, M., Chouinard, J.Y.: Propagation loss prediction: A comparative study with application to the mobile radio channel. IEEE Trans on Ant and Prop VT-34, 86–96 (1985)
5. Durgin, G., Rappaport, T., Xu, H.: Measurements and models for radio path loss and penetration loss in and around homes and trees at 5.85 ghz. IEEE Trans on Comms 46(11), 1484 –1496 (nov 1998)
6. Erceg, V., Hari, K., et al.: Channel models for fixed wireless applications. Tech. rep., IEEE 802.16 Broadband Wireless Access Working Group (2001)
7. Rizk, K., Wagen, J.F., Gardiol, F.: Two-dimensional ray-tracing modeling for propagation prediction in microcellular environments. IEEE Trans on Veh Tech 46(2), 508 –518 (May 1997)
8. Robinson, J., Swaminathan, R., Knightly, E.W.: Assessment of urban-scale wirelesss networks with a small number of measurements. In: MobiCom (2008)
9. Sharma, P.K., Singh, R.: Comparative analysis of propagation path loss models with field measured databases. International Journal of Engineering Science and Technology 2, 2008–2013 (2010)
10. Shin, H.: Measurements and Models of 802.11B Signal Strength Variation Over Small Distances. Master's thesis, University of Delaware (2010)
11. Sridhara, V., Bohacek, S.: Realistic propagation simulation of urban mesh networks. Computer Networks 51(12), 3392 – 3412 (2007)