Research Article

A New Statistical Algorithm Based on the Conventional Lee’s Path Loss Model for the 900 and 1800 MHz

Liyth Nissirat, Mahamod Ismail, Mahdi A. Nisirat and Mandeep Singh

Department of Electrical, Electronics and System Engineering, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia

Department of Electrical Engineering, Faculty of Engineering Technology, Albalqa’ Applied University, Jordan

Abstract: The aim of this proposed technique is to optimize the prediction quality and to minimize the overall measured RMSE error of the original Lee’s path loss model. The new technique is a statistical algorithm based on the concept of mean value normalization that has a wide range of applications. Statistical path loss models, such as Lee’s model, are assumed as major prediction models used primarily in the pre-planning procedures to pre-estimate losses and minimize the overall cost. Lee’s model as a major prediction model is well known to accurately predict such losses in diverse landscape criteria’s. As compared to the original model, the new proposed technique has demonstrated better RMSE accuracy. Less RMSE, of an average of 3-4 dB’s is obtained in most macro-cell open areas in the area of Jiza town, south of Amman city, Jordan. Examples are provided in both the 900 MHz and the 1800 MHz to signify the enhancement of the prediction accuracy of the new proposed algorithm.

Keywords: Large scale path loss models, model optimization, statistical parameters

INTRODUCTION

Lee’s model is one of the major prediction models that are widely accepted as an accurate statistically based model. It was proposed in 1982 by Lee (1985) after conducting extensive experimental measurements in different USA cities with diverse variation in landscape structures. The model was originally proposed for the 850 MHz band by inserting the local averages for a stream of received signal strengths known as bins. Corrections for other bands were made possible by inserting a relative frequency modification fraction. Real collected received power data are used in order to fine tune the Lee’s model and to establish the required training procedure based on the provided original set. The general proposed model by Lee is of the form:

\[ PL = L_0 + \alpha \log \left( \frac{r}{h_{ref}} \right) - 10 \log(F_1F_2F_3F_4F_5) + 10 \log \left( \frac{P_{ref}}{P_{ref \text{ avg}}} \right) \]  

where,

\[ F_1 = \left( \frac{h_b}{30.46} \right)^2 \]  

\[ F_2 = \left( \frac{G_a}{4} \right) \]  

\[ F_3 = \left( \frac{h_m}{3} \right)^2 \text{ for } h_m > 3 \]  

\[ F_4 = G_m \]  

\[ F_5 = \left( \frac{f}{900} \right)^{-n} \]

The variables are as given in Table 1.

The set of data used in Lee’s experiments ranges from basic suburban areas into dense metropolitan areas in USA and other some major cities worldwide as shown in Table 2. For the case of the suburban areas a Path loss slope of 38.4 dB/dec. and an intercept point of 53.9 dB (or with the A parameter of 99.86 dB) were reported. Other city structure category embodies intense metropolitan areas such as the Tokyo city which was represented by a slope of 30.5 dB/dec. and an intercept point of 77.8 dB (or with the A parameter of 123.77 dB). The difference in these two parameters, the slope and the intercept point, between disparate areas landscape was related to the different impact of two major factors that would affect the propagation of signals (Lee, 1985, 1992, 1993, 1997; Lee and Lee, 2000):
Table 1: List of the variables used in Lee’s conventional model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_L$</td>
<td>Received signal level, the mean signal level received at the distance $R$ from the transmitter (dB)</td>
</tr>
<tr>
<td>$L_0$</td>
<td>The signal strength expected for reference conditions $R_{REF}h_{bREF}h_{mREF}$</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>The slope, the rate of decay in signal strength (dB/decade)</td>
</tr>
<tr>
<td>$R$</td>
<td>The distance from the transmitter</td>
</tr>
<tr>
<td>$R_{REF}$</td>
<td>The reference distance</td>
</tr>
<tr>
<td>$h_b$</td>
<td>The height of the base transmit antenna</td>
</tr>
<tr>
<td>$h_m$</td>
<td>The height of the mobile receive antenna</td>
</tr>
<tr>
<td>$h_{bREF}$</td>
<td>The reference height of the mobile receive antenna</td>
</tr>
<tr>
<td>$f$</td>
<td>Operating frequency</td>
</tr>
<tr>
<td>$G_b$ and $G_m$</td>
<td>Base antenna and mobile antenna gains, respectively</td>
</tr>
<tr>
<td>$n$</td>
<td>Frequency adjustment factor takes values between 2 and 3</td>
</tr>
<tr>
<td>$P_{tx}$</td>
<td>$P_{tx}$, the transmit ERP</td>
</tr>
<tr>
<td>$P_{REF}$</td>
<td>The reference transmit ERP</td>
</tr>
</tbody>
</table>

Table 2: Lee’s model parameters for different areas

<table>
<thead>
<tr>
<th>Area and type</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tokyo (urban)</td>
<td>123.77</td>
<td>30.5</td>
</tr>
<tr>
<td>New work (urban)</td>
<td>101.20</td>
<td>43.1</td>
</tr>
<tr>
<td>Philadelphia (urban)</td>
<td>108.49</td>
<td>36.8</td>
</tr>
<tr>
<td>Sub-urban</td>
<td>99.86</td>
<td>38.4</td>
</tr>
<tr>
<td>Rural</td>
<td>86.12</td>
<td>43.5</td>
</tr>
<tr>
<td>Free space</td>
<td>96.92</td>
<td>20.0</td>
</tr>
</tbody>
</table>

- Manmade structures and clutters
- Natural terrain

Thus in the original Lee’s model the effect of each of the above mentioned factors were dealt with separately (Lee and Lee, 2000).

In Lee and Lee (2000) claimed that the careful averaging process of path losses through the radio-path would wipe out terrain variations and thus would reveal the losses due to manmade structures. The drawback of this averaging process is that it would produce an area path loss pattern that may be assumed not desirable if the detailed effect of the terrain is preferred.

To accommodate the effect of the natural terrain, the normalization process described in Lee and Lee (1997) may be implemented to eliminate the effect of terrain landscape on the received signal. Though may be considered accurate as has been claimed, the insertion of the mentioned normalization process on each measured point is not an easy task. In addition, high resolution terrain data is required to perform such a procedure that may complicate the calculation process, if such data is assumed to be obtained.

In his basic suggestion of the model, Lee’s claimed that the proposed mathematical procedure would convey prediction accuracy within a ±3 dB. Major studies, though, revealed that the prediction accuracy of lee’s model lay below the expected level of precision in different landscape structures and in different frequency ranges (Xu et al., 2010; Evans et al., 1997). This may be attributed to different reasons. Among them is the terrain roughness extent that may cause additional losses and has been covered in later extension of the model as proposed in Lee and Lee (2000). The second cause is the manmade structure effect. The extent of such structures differs significantly from the areas where the original Lee’s model was conducted. Finally, the correction provided by Lee for different frequency ranges was also proved to be not as accurate as expected (Evans et al., 1997). Therefore many proposed studies have adopted different scale of optimization to suit Lee’s model in their area of measurements. The deviation of Lee’s model in the given area of studies has been reported in the range of 5-10 dB (Xu et al., 2010; Evans et al., 1997; Nissirat et al., 2012).

The optimization process proposed in Evans et al. (1997) may be found less complex. The technique proposed was meant to test the validity of Lee’s model for the 1900 MHz prediction. Results showed that predicted losses in flat areas deviates in an average of 5-6 dB from the actual measured losses. More deviation in the range of 9-10 dB is calculated for areas of more rough terrain topology. Concluded remarks signify the validity of Lee’s model in the given range.

The objective of this study is to introduce a new statistical algorithm, which relies on the original Lee’s path loss model and is based on data normalization over the mean (average) value. This new proposed technique uses the same mathematical structure originated by Lee; though depart from the original model by the statistical management used to perform the path loss calculations. Followed sections would first introduce the specifications of the measurement campaign along with the details of the area under test. Then the new statistical method of conduction is proposed and mathematically derived. Finally, examples are presented to verify the accuracy of the proposed model as compared to the original Lee’s model.

**DATA SET AND MEASUREMENT CAMPAIGN**

Principally, statistical analysis relies heavily on available data. Such data either be collected by the researchers in the field or obtained from original sources for analytical study. In this study, received power data obtained from mobile service providers working in the Hashemite Kingdom of Jordan. These data where obtained from two companies, as a result of
a drive test campaign, the first named Umniah, working
in the frequency band of 1800 MHz and the second
called Zain working in the frequency band of 900 MHz.
Data collected with the aid of a Global Positioning
System (GPS) device contained among others,
longitude, latitude and altitude readings of both of the
transmitter and the route taken by the receiver. The set
of data used in this study where mainly from two
locations. The first location is through the airport road
and had collected path-loss readings starting from the
west of Amman city down to Jiza town of a length of
about 30-40 km as shown in Fig. 1.

The data that suit the class of an open area was
extracted, were the open area category was assumed as
per the classification given in Table 3. The second
location was around the center of Jiza town which was
considered also as an open area as shown in Fig. 2.
Scattered building of a single to two floors, 8-12 m of
altitude is the most prevailed structures in the open area
case.

The transmitters are mainly of omni-directional
Kathrein antennas and are mounted on towers of about
20-30 m above ground levels as shown in Fig. 3.
During the drive test campaign, the calculated average
speed of the car was 30-40 km with an average
measuring sample rate of 3-4 samples/m.

**METHODOLOGY**

The measurements are conducted for macro-cell
scenarios in open areas with almost flat landscape
profile. Using datasets collected from open areas
facilitates focusing on the effect of the natural terrain
variation on the accuracy of estimating the path-loss
parameters. The need of terrain normalization described
in Lee (1992) and Lee and Lee (1997) is not needed in
such a case were the effect of the buildings and
manmade structures is assumed minimum. Assuming
an n of statistically independent path-loss datasets
arranged as columns of the matrix \( \mathbf{Y} = [Y_1, Y_2, ..., Y_n] \),
where each dataset \( Y_i \) contains \( r \) number of samples,
then \( \mathbf{Y} \) is \( r \times n \) matrix. The path-loss is measured with
respect to the travelled distance \( x \) producing a pair of
data for each point \( (x_m, Y_i^m) \), \( m = 1...r \) and \( i = 1...n \).

In the conventional Lee’s model, the intercept and
the slope of the path-loss curve are calculated for each
dataset \( Y_i \) using the least square method (Lee, 1985,
1992; Lee and Lee, 2000; Alotaibi et al., 2008).
Therefore, each pair of data \( (x_m, Y_i^m) \) is related
according to the linear fitting:

\[
Y_i^m = C_i + D_i x_m
\]
Fig. 2: Routes of Jiza town south of Amman city. Dimensions 8×8 km, Google Earth (2010)

Fig. 3: Picture of the transmitter used for the 900 MHz, airport road, Jiza town, Amman, Jordan

where, $C_i$ and $D_i$, are the intercept and the slope, respectively for the dataset $Y_i$. Then the intercept and the slope in Lee’s model would be calculated as the average of the intercept and the slope of every dataset:

$$C = \frac{\sum_{i=1}^{n} C_i}{n} \quad \text{and} \quad D = \frac{\sum_{i=1}^{n} D_i}{n}$$ (8)

In the proposed method, each dataset is first statistically scaled and shifted producing a transformed dataset $Y'_i$ presented as (Lee, 1985, 1992; Lee and Lee, 2000):

$$Y'_i = \frac{Y_i - \mu_{Y_i}}{\sigma_{Y_i}} \quad \text{for } i = 1, \ldots, n$$ (9)

where, $K_i$: A scaling factor for the dataset $Y_i$; $\mu_{Y_i}$ and $\sigma_{Y_i}$: The mean and the standard deviation of the dataset $Y_i$.

The average of the scaled and shifted datasets is then fitted using the least square method to calculate the normalized intercept and slope (Liping et al., 2009). For each point in the scaled and shifted dataset $Y'_i$:

$$\frac{Y'_1 + Y'_2 + \ldots + Y'_n}{n} \approx C_{\text{norm}} + D_{\text{norm}} x_m,$$

for $m = 1, \ldots, r$ (10)

Which is equivalent to:

$$\frac{\sum_{m=1}^{r} K_i Y'_m}{\sigma_{Y_1}} + \frac{\sum_{m=1}^{r} K_i Y'_m}{\sigma_{Y_2}} + \cdots + \frac{\sum_{m=1}^{r} K_i Y'_m}{\sigma_{Y_n}} \approx C_{\text{norm}} + D_{\text{norm}} x_m$$ (11)

Rearranging (11) we would have:

$$\frac{\sum_{m=1}^{r} K_i Y'_m}{\sigma_{Y_1}} + \frac{\sum_{m=1}^{r} K_i Y'_m}{\sigma_{Y_2}} + \cdots + \frac{\sum_{m=1}^{r} K_i Y'_m}{\sigma_{Y_n}} \approx C_{\text{norm}} \prod_{j=1}^{n} \sigma_j + D_{\text{norm}} \prod_{j=1}^{n} \sigma_j x_m + \left( \frac{\sum_{m=1}^{r} K_i Y'_m}{\sigma_{Y_1}} + \frac{\sum_{m=1}^{r} K_i Y'_m}{\sigma_{Y_2}} + \cdots + \frac{\sum_{m=1}^{r} K_i Y'_m}{\sigma_{Y_n}} \right)$$ (12)

For the fitting line to represent the datasets $Y_i$, the scaling factors should be:

$$K_i = \frac{1}{\prod_{j=i}^{n} \sigma_j}$$ (13)
Substituting $r_{G7} > d$ and $r_{Gbdc}$ into Eq. (9), the scaled and shifted data can be written as:

\[
Y_i^* = Y_i - \frac{\mu_i}{\sigma_i}
\]

(14)

This procedure is summarized in the following flowchart given in Fig. 4.

RESULTS AND EXAMPLES

Two locations, one is for the 900 MHz band and the other is for the 1800 MHz were used as examples to demonstrate the validity of this proposed method. The first location was through the airport road crossing the Jiza town with a measured length of about 3 km as shown in Fig. 5.

The calculated intercept value and slope applying Lee’s conventional path-loss method were 121 dB and 26.5 dB/dec, respectively. The calculated values obtained by the proposed method were 115.6 dB for the intercept value and 32.9 dB/dec for the slope. Obtained RMSE of 3.83 dB is calculated for the conventional method as compared to 6 dB for the Lee’s conventional method. The best fitting line RMSE was measured with a deviation of 3.35 dB as shown in Fig. 6.
The 1800 MHz route is located in the suburb of Jiza town. The drive test was along a road of 3 km length where scattered buildings not more than 10 m of altitude were present as shown in Fig. 7. An RMSE of 3.8 dB is measured for the proposed method as compared to 6.6 dB for the original Lee model as shown in Fig. 8. The improved RMSE values obtained signifies the importance and advantages of the proposed method as compared to the original Lee model.
Table 4 summarizes the obtained RMSE, of some other routes taken in this study, for both the 900 and 1800 MHz bands as compared to the conventional Lee model.

CONCLUSION

This paper emphasizes the effect of statistically scaling and shifting the obtained path-loss datasets on improving the estimation of Lee’s path-loss parameters. Statistical Scaling and shifting brought datasets collected from different areas into an absolute common average and deviation values. Such procedure would rather reduce the effects of terrain variations on the collected datasets. The proposed method has been applied on datasets collected from 20 different sites for both the 900 and 1800 MHz bands. In both frequency ranges the proposed method predicted the path-loss more accurately as compared to the conventional method. It is worth to note that the maximum obtained deviation of the proposed method prediction from the best fitting line is about 1 dB. Moreover, the proposed method prediction accuracy outperformed the conventional method by up to 4-5 dB.

ACKNOWLEDGMENT

The authors would express their appreciation to the support of this study under the governmental grant number Oup-2014-182.

REFERENCES


