Investigating the Correlation Between Gamma Distribution and Probability Density of Data Rates

Jeong Min Kong

Raviraj S. Adve

UNIVERSITY OF TORONTO

Introduction

What?

Recent Observations Have Shown That Gamma Distribution Correlates Very Well With Probability Density Graph of Data Rates at Microwave Frequencies. To Further Understand the Phenomenon, the Goal of This Investigation Was to Find **Relationships Between the Gamma Distribution Parameters and Various Environmental Factors**. Note: Please Go to the Last Slide (Appendix) to Read a Brief Explanation About Gamma Distribution, If Necessary.

How?

Python-Based 4G LTE and 5G mmWave Wireless Coverage Simulation Program Was Used to Obtain SINR at Each User; Additional Code Was Written to Approximate the Data Rates Using the Shannon-Hartley Theorem. **MATLAB** Was Used for Analyzing the Data Rates (ie. Finding the Best-Fit Gamma Distribution Parameters for Probability Density of Data Rates) and Plotting (ie. Gamma Distribution Parameters vs Values of the Environmental Factor Being Investigated).

UNIVERSITY OF TORONTO

Verification of the Simulation Program

As Mentioned in the Introduction, the Python-Based 4G LTE and 5G mmWave Wireless Coverage Simulation Program Was Very Important for This Investigation As It Was Used for Collecting Data Rates That are Used to Make Probability Density Graphs. Therefore, to Verify That Output Data Rates are Good Approximations of the Real World Outcomes, **Test Simulations** Were Ran to Check That Probability Density Graph of the Output Data Rates are Well Modeled With Gamma Distribution.

Test Simulation Examples:

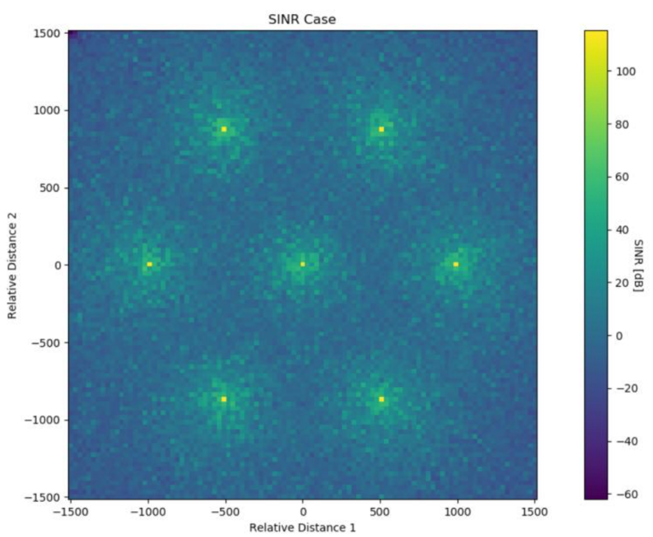


Fig 1. SINR Heat Map for Urban: 6 Cellular Networks Each Located in a Hex Cell, No Buildings

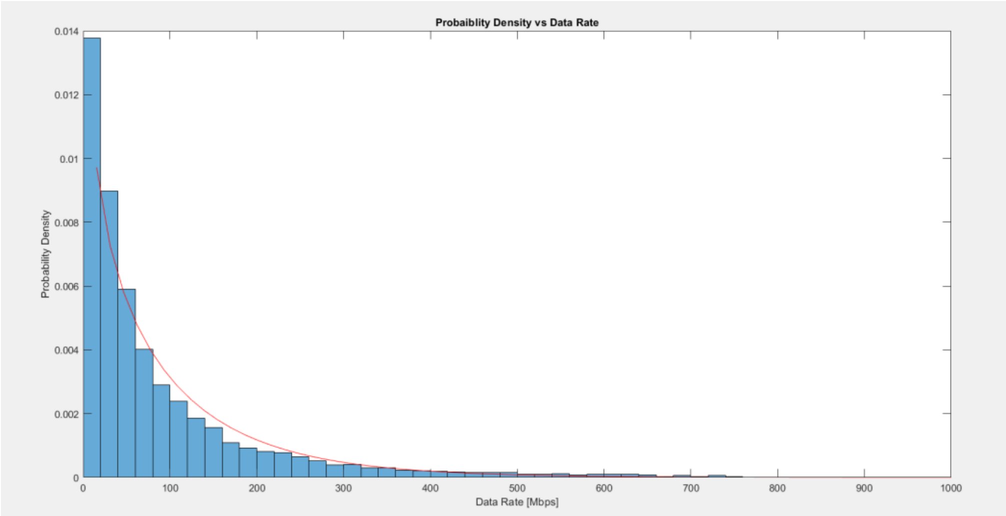


Fig 2. The Blue Histogram is the Probability Density of the Data Rates and the Red Line is the Best-Fit Gamma Distribution.

UNIVERSITY OF TORONTO

Verification of the Simulation Program

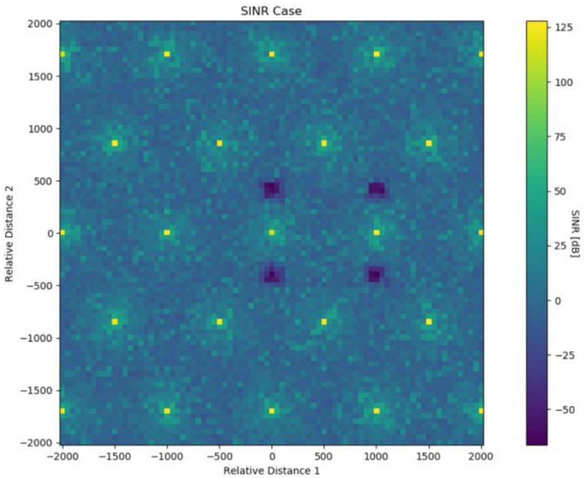


Fig 3. SINR Heat Map for Urban: 23 Cellular Networks Each Located in a Hex Cell, Few Buildings

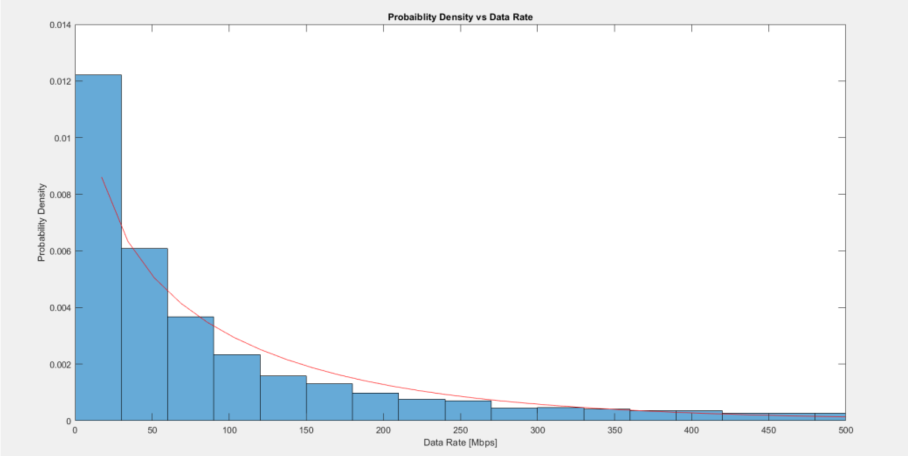


Fig 4. The Blue Histogram is the Probability Density of the Data Rates and the Red Line is the Best-Fit Gamma Distribution.

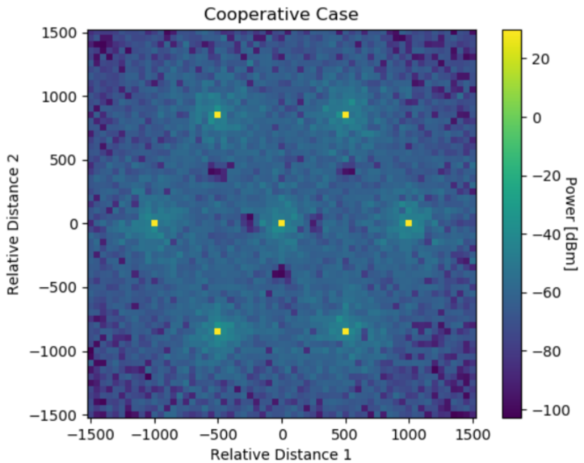


Fig 5. Cooperative Heat Map for Rural: 6 Cellular Networks Each Located in a Hex Cell, Few Buildings

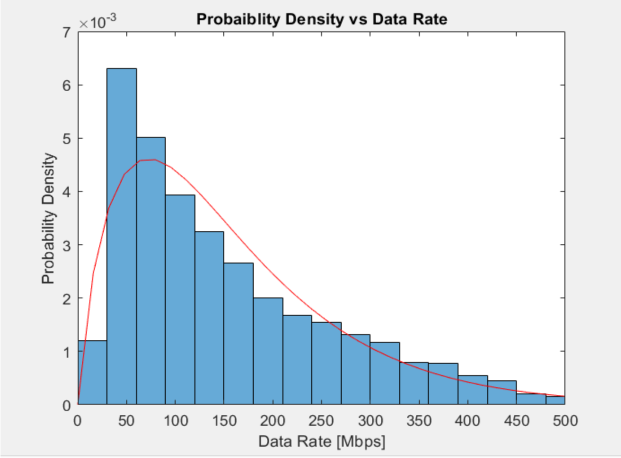


Fig 6. The Blue Histogram is the Probability Density of the Data Rates and the Red Line is the Best-Fit Gamma Distribution.

Environmental Factors

Specifically, This Investigation Focused on Finding Relationships Between the Gamma Distribution Parameters and **3 Environmental Factors:**

1. **# of Users:**
   Make Any Interesting Observations By Plotting Alpha and Beta (Gamma Distribution Parameters) vs # of Users in a Given Map.
2. **# of Access Points:**
   Make Any Interesting Observations By Plotting Alpha and Beta (Gamma Distribution Parameters) vs # of Access Points in a Give Map.
3. **Shadowing Values:**
   Even Though Shadowing is Typically Modeled Using Gaussian Normal Distribution, the Shadowing Values Will Be Manually Set for This Investigation to Make Any Interesting Observations in the Alpha and Beta (Gamma Distribution Parameters) vs Shadowing Values Plot.

UNIVERSITY OF TORONTO

Environmental Factor 1: # of Users

Observations:

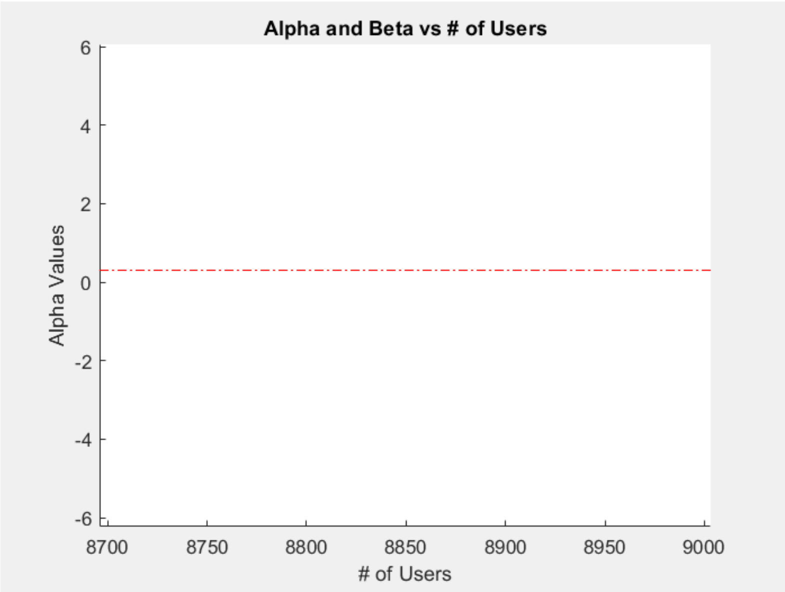


Fig 7. Alpha Values vs Number of Users

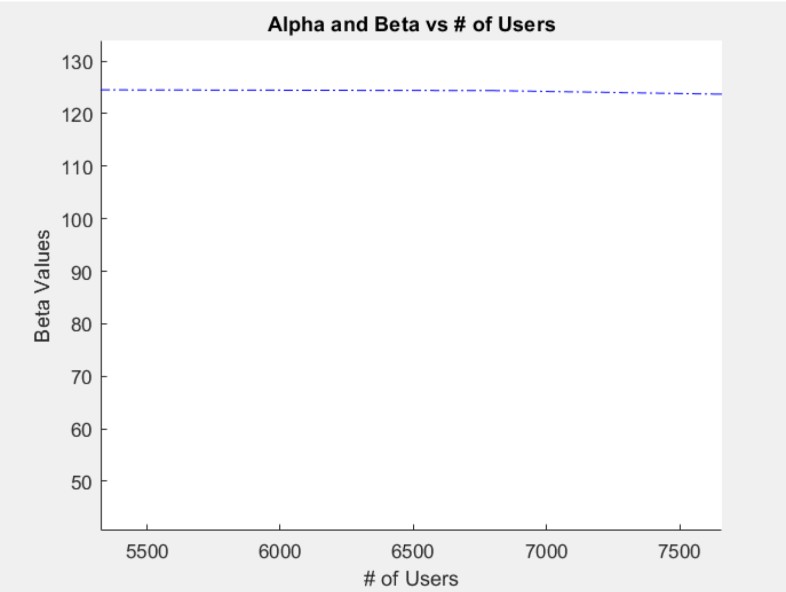


Fig 8. Beta Values vs Number of Users

Alpha and Beta vs # of Users

As It Can Be Seen On Figures 7 and 8, **Both** the Alpha and the Beta Values Appear to Be **Unaffected** By the Number of Users in a Given Map. The Differences Between the Maximum and the Minimum Values Were Very Small; For the Simulation in Figure 7 and 8, the Alpha Values Were Between 0.3763 and 0.3783 and the Beta Values Were Between 130.5 and 133.75. Because the Differences Between the Maximum and Minimum Boundaries Were Extremely Small, the Relationship Between Alpha/Beta Values and # of Users Can Be Approximated With **Constant** Values.

These Relationships Imply That the Number of Users in a Map Does **Not Affect** the Average Data Rate.

UNIVERSITY OF TORONTO

Environmental Factor 1: # of Users

Method Used :

1. Use the Python-Based Wireless Coverage Simulation Program to Obtain Cellular Data Rates for Every 'x' Meters (eg. Every 10 Meters). Store These Data Rates in an Array and Transfer It to MATLAB.
2. For Every i (Where i Ranges From 1 to the Size of the Array), Remove the Elements of the Data Array With Indexes Equal to the Multiples of i. The Purpose of This Step is to Remove Users in Such a Way That Number of Users/Hex Cell is Random for All Cells.
3. Use These Modified Data Arrays for Making Probability Density Graphs and Determining the Best-Fit Gamma Distribution Parameter Values. Store the Parameter Values for Each i in the New "Alpha" and "Beta" Arrays.
4. Plot Alpha and Beta vs # of Users Graph and Analyze the Results.

UNIVERSITY OF TORONTO

Environmental Factor 2: # of Access Points

Observations:

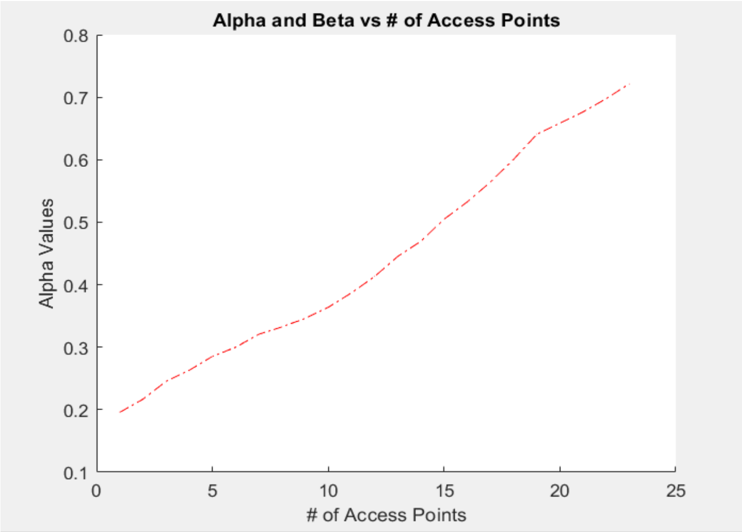


Fig 9. Alpha vs # of Access Points

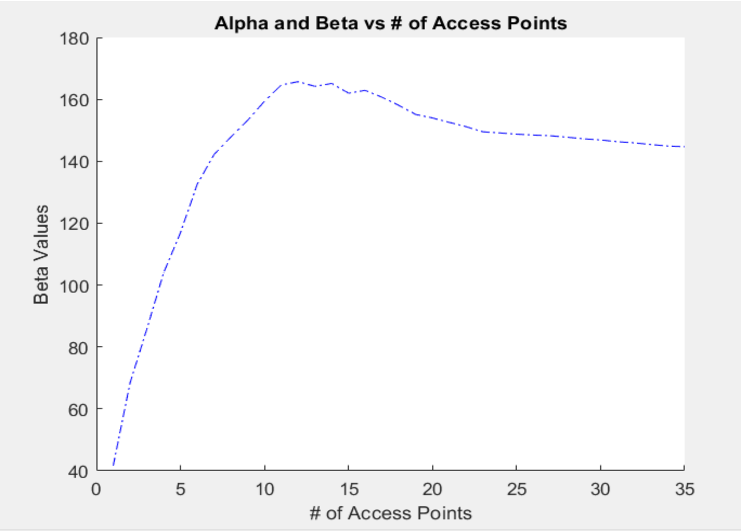


Fig 10. Beta vs # of Access Points

Alpha vs # of Access Points:

As It Can Be Seen On Figure 9, the Relationship Between the Alpha Parameter and the # of Access Points is Approximately +ve **Linear**. For the Simulation in Figure 9, the Alpha Values Increased Between 0.015-0.030 for Every Increase in the Number of Access Points; a Very Consistent Rate.

Beta vs # of Access Points:

As It Can Be Seen On Figure 10, the Relationship Between the Beta Parameter and the # of Access Points Appears to **Increase** Logarithmically Until It Reaches Its Maximum Value; It Then Slowly **Saturates** to Approach a Limit.

These Relationships Show That the Average Cellular Data Rate of the Users in the Map **Increases** As the Number of Access Points **Approaches** Its Saturation Value; When the Number of Access Points **Reaches** the Saturation Point, the Average Cellular Data Rate Begins to **Slowly Decrease**.

UNIVERSITY OF TORONTO

Environmental Factor 2: # of Access Points

Method Used:

1. Add a Cellular Network at the Intended Location In the Python-Based Wireless Coverage Simulation Program. Store the Resulting Data Rates for All Users in the Map in an Array and Transfer It to MATLAB.
2. Make a Probability Density Graph Using the Data Array and Store the Best-Fit Gamma Distribution Parameter Values in the New “Alpha” and “Beta” Arrays.
3. Repeat Steps 1-2 With Additional Cellular Networks Each in Different Locations. Make Sure That the Environment of the Map (eg. # of Users) **Other** Than the # of Access Points is the **Same** for All Simulations.
4. Plot the Alpha and Beta vs # of Access Points Graph and Analyze the Results.

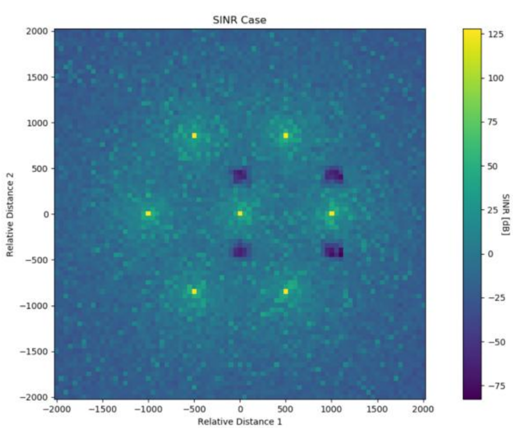


Fig 11. SINR Heat Map With 7 Cellular Networks for Simulation From Figures 9/10.

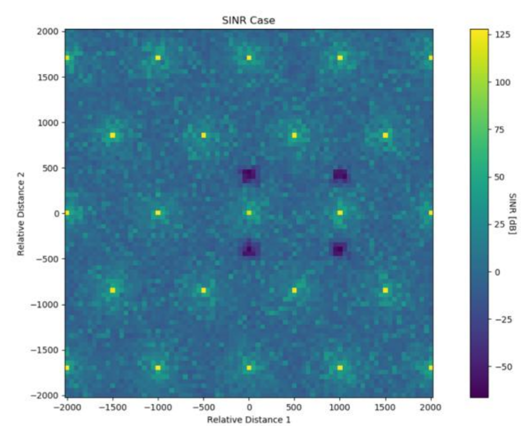


Fig 12. SINR Heat Map With 23 Cellular Networks for Simulation From Figures 9/10.

UNIVERSITY OF TORONTO

Environmental Factor 3: Shadowing Values

Observations:

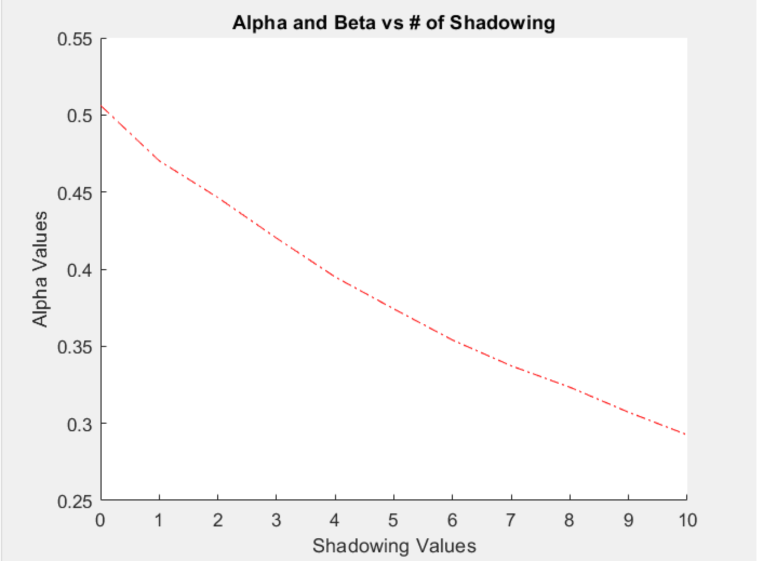


Fig 13. Relationship Between Alpha and Shadowing Values

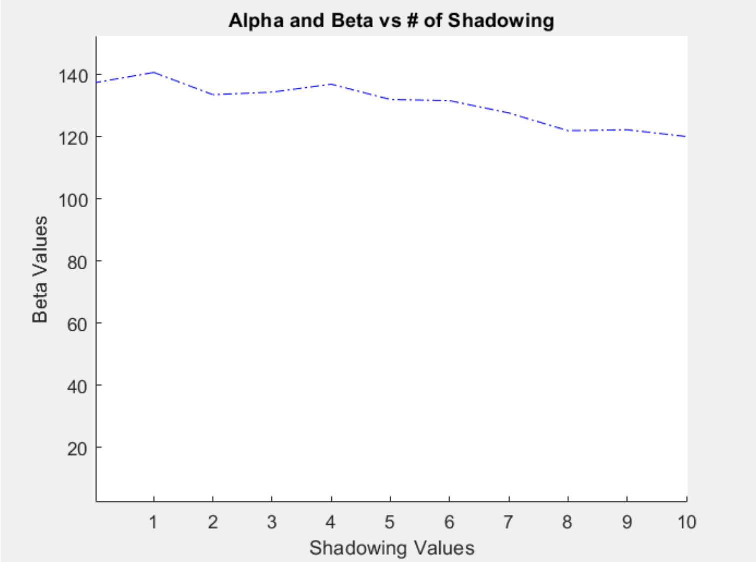


Fig 14. Relationship Between Beta and Shadowing Values

Alpha vs Shadowing Values

As It Can Be Observed in Figure 13, the Relationship Between Alpha and Shadowing Values is Approximately – ve **Linear**, Unlike the Relationship Between Alpha and Number of Access Points Which is +ve Linear.

Beta vs Shadowing Values

As It Can Be Observed in Figure 14, the Relationship Between Beta and Shadowing Values is Approximately - ve **Linear**. By Calculation, the Rate at Which It Decreases is **Greater** Than the Rate at Which Alpha vs Shadowing Values Graph Decreases.

These Relationships **Imply** That the Average Data Rates of the Users in the Map **Decreases** As the Shadowing Value **Increases.** This is an Expected Result Because Higher Shadowing Value Means Greater Power Loss of the Signals.

UNIVERSITY OF TORONTO

Environmental Factor 3: Shadowing Values

Method Used:

1. In the Python-Based Wireless Coverage Simulation Program Code, Remove the Gaussian Normal Distribution and Set the Intended Constant Value to the Shadowing Variable.
2. Run the Simulation Program and Obtain the Data Rates for All Users in the Map. Stores These Data Rates in an Array and Transfer to MATLAB.
3. Make a Probability Density Graph Using the Data Array and Store the Best-Fit Gamma Distribution Parameter Values in the New “Alpha” and “Beta” Arrays.
4. Repeat Steps 2-3 With Different Shadowing Values. Make Sure That the Environment of the Map (eg. # of Users, # of Access Points, etc) is **Constant** for All Simulations.
5. Plot the Alpha and Beta vs Shadowing Values Graph and Analyze the Results.

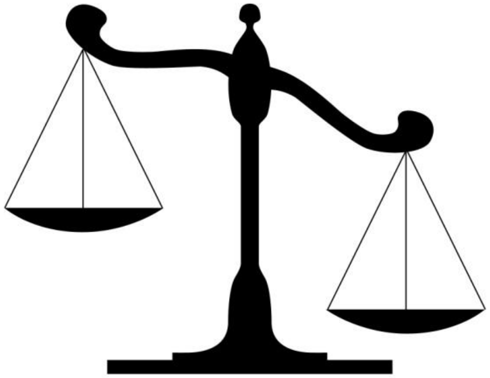
UNIVERSITY OF TORONTO

Future Plans

1. ONLY **Single-Antenna** Cellular Networks Were Accounted For. Future Plan is to Also Consider **Multiple-Antenna** Cellular Networks.

2. ONLY Signals at **Microwave** Frequencies Were Accounted For. Future Plan is to Also Consider the Signals at **mmWave** Frequencies and Check If the Similar Patterns From Microwave Frequencies are Also Shown At mmWave Frequencies.

3. This Study Mainly Focused On **Observing** Any New Interesting Relationships Between the Gamma Distribution Parameters and the Environmental Factors. Future Plan is to Derive **Mathematical Expressions** to Explain and Further Predict These Relationships.



UNIVERSITY OF TORONTO

Appendix: What is Gamma Distribution?

Gamma Distribution is One of the Types of Continuous Probability Distributions With 2 Independent Parameters [1]. Alpha, Also Known As “Shape Parameter”, and Beta, Also Known As “Scale Parameter”, are These Independent Parameters That Determines the Form of the Probability Density Functions, Which are Expressed As:

$$f(x;\alpha,\beta) = \frac{\frac{1^{\alpha}}{\beta}\, x^{\alpha-1}}{\zeta(\alpha)}\, e^{\frac{-x}{\beta}},$$

[1]. Rewritten Using Microsoft Word to Express the Equations in Terms of Scale Parameter Instead of Rate Parameter.

Where $\mu = \alpha\beta$ and $\sigma^2 = \alpha\beta^2$

Gamma Distribution Can Very Well Model the Distribution of Inter-Spike Intervals in Neuroscience [2] and the Cellular Data Rates at Microwave Frequencies As Discussed in This Document.

Here are Some Examples of Gamma Distribution:

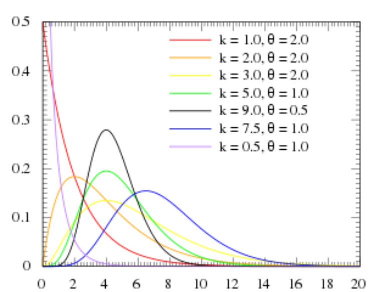


Fig 12. k is the Shape Parameter and Theta is the Scale Parameter

UNIVERSITY OF TORONTO

References

[1] "Gamma Distribution: Definition, PDF, Finding in Excel," *Statistics How To*. [Online]. Available: http://www.statisticshowto.com/gamma-distribution/.

[2] N. Friedman, L. Cai and X. S. Xie (2006) "Linking stochastic dynamics to population distribution: An analytical framework of gene expression", *Phys. Rev. Lett.* 97, 168302.

UNIVERSITY OF TORONTO

Thank You.

UNIVERSITY OF TORONTO