Experiment to Determine the 4G LTE Speed on Mobile Phones

Team: Shuqiong Chen, Yang Liu, Jiawei Xue
Instructor: Chihoon Lee

Experimental Design, 2018Fall
schen78@stevens.edu, (551)-233-3876
yliu249@stevens.edu, (201)-985-4435
jxue4@stevens.edu, (201)-985-4433

ARTICLE INFO

Keywords:
DOE
Factorial Design
Pareto Principle
4G LTE
JMP
Speedtest

ABSTRACT

When it comes to 4G LTE, every person using smartphones may have some intuitive knowledge that how fast your smartphone downloads data often depends on where you are, when you use it, even how you are using it. In this experiment, instead of instinct, fractional factorial design was used to figure out which factors matter and which ones don't, as well as help finding optimal settings for highest download speeds. The Pareto principle had been applied to the analysis process for helping screen out a few important factors before we got the regression equation.

1. Motivation

With the rapid development of telecom technology, people are able to access the high-speed internet network. Cell phones become an indispensable part of our daily life since we use it to access a network (Figure 1). 4G LTE is the most common standard for high-speed wireless communication for mobile devices and data terminals. However, sometimes we may find that the network speed varies a lot in different conditions. Low internet connection happens for various reasons, even when you use the latest version iPhone and pay for a high-speed connection. We would like to figure out the potential factors which may affect the 4G LTE network speed through an experiment. Finding out these factors will help us make a reasonable expectation about Internet speed under certain conditions. This way, we would be allowed to access Internet connection more efficiently. Better yet, we can provide some practical suggestions to smartphone users who may have connection concerns.

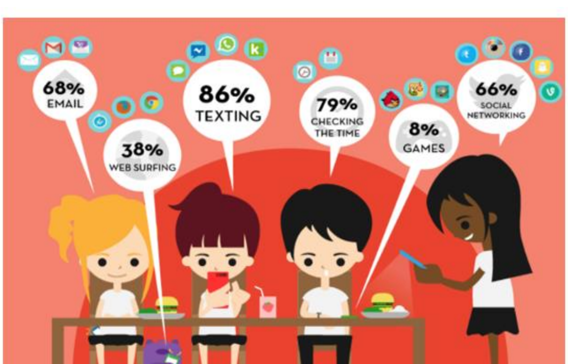


Figure 1: Cellphone in our life

2. Problem description

With this motivation, we had a clear goal-conducting an experiment to test connection speed and identify significant factors. Here came some questions. What is Internet speed? How is Internet speed measured? What is a good Internet speed? Certainly, we did not intend to discuss the Internet itself in detail. We still needed to customize these definitions in our design. The reason to clarify that was because good definitions of these questions would ensure our design perform well and enabled our conclusion to apply to daily practice.

Internet speed refers to the speed which data or content travels from the World Wide Web to your home computer, tablet, or smartphone. The speed of this data is measured in megabits per second ("Mbps"). Internet speeds can be separated into download speed and upload speed depending on the different transmission directions—from the Internet to a terminal or from a terminal to the Internet. Typically, a smartphone user cares most about the download speed. So, in this experiment, we aimed to test the download speed.

Then, it was time to define a good Internet speed. It became a bit of complicated and arbitrary. Different smartphone users may have different definitions of a good Internet speed since that definition will depend on their reason to use the Internet. Some users may use the Internet for recreational activities such as streaming high-definition videos or downloading high-resolution photos. Some users may prefer to chat with friends, use social networking, or casually surf the Internet. Apparently, the former users will have a higher requirement than the latter users. Hence, we chose to use Mbps to measure speed and learned what factors would contribute to the highest download speed instead of a good download speed.

Figure 2: Speedtest

So far, we were clear about test objective and the measurement. Then, how could we get the data? Obviously, the reliable tools were in demand, so we have so many options. After sorting several test tools, we decided to use Speedtest (Figure 2) as the measuring tool because of its friendly user interface and popularity.

3. Experiment design

(a) Responsible variable (Response)

4G LTE network speed- download speed would be the response variable in our experiment.

(b) Independent variables (Factors)

We brainstormed all the possible and common factors that could be studied as independent variables. You may wonder why we stressed the "common" factors. There may exist various factors that have an impact on download speed. Our goal was to study factors which the majority could encounter in daily life and control on some level. Therefore, we selected four factors from various ideas as our study factors (Figure 3) including location, time, phone cover, and battery. We set two levels for each factor.

For Location, crowded place (Crowd space) would be low level (-). In contrast, a place with a few people (Wide space) would be set as high level (+).

For Time, we adopted most carriers' definition for peak hours (between 7 and 11 pm) and set this period as low level (-). The other period would be high level (+).

For Phone Cover, taking consideration of physical protection, better grip, unique outlook, etc., most people prefer to use a phone cover. It's rare to see people use a phone without a cover. So, we set the most common phenomenon, a phone with a cover, as low level (+) and a phone without a cover as low level (-).

For Battery, as we all know, when Low Power Mode is on, your iPhone will last longer before you need to charge it, but some features might take longer to update or complete. Also, some tasks might not work until you turn off Low Power Mode, or until you charge your phone to 80% or higher. Naturally, Low mode would be low level (-) and General Power Mode would be high level (+).

Factor	(-)	(+)
Location	Crowd Space	Wide Space
Time	Busy Time(7:00 am – 23:00 pm)	Idle Time(23:00 pm – 7:00 am)
Phone Cover	With Cover	No Cover
Battery Setting	Low Power Mode	General Power Mode

Figure 3: Factors

(c) Block

We believe that different versions of phones are not homogeneous, even though they are provided by the same company. That is, there possibly are version effects. Then, it would be better to observe each version phone under all four factors. In this design, we set the version of iPhone as a block to avoid the additional noise from version effects. We should note that in design matrix block1 meant iPhone 7 and block 2 meant iPhone 8 plus.

(d) Matrix design

In this design, as we discussed previously, there were four factors (each with two levels) and one block. We viewed this block as one factor so as to design a matrix more efficiently. Considering the degrees of freedom, we intended to replicate this design once to reduce some a combination of bias and chance. In this case, if we used two-level full factorial design, there would be $\mathbf{2^{4+1} * 2 = 64}$ runs in total and tend to be a redundancy. A good method to exploit this redundancy is to use fractional factorial design (FFD). Hence, a fractional factorial design was applied to our study. The summarized key inputs were shown in Figure 4.

The number of factor	4
The number of level of each factor	2
The number of block	1
The number of replication	1
Total runs	$2^{(4+1)-2} * 2 = 16\ runs$

Figure 4: Key inputs

After clarifying these key inputs, we used JMP, one of the popular statistical software, to help us generate the randomized order. Then, it's time to collect data.

4. Data collection

According to the design matrix created by JMP, we collected the data in two places. The crowed place in our design was ShopRite located in 9th street, Hoboken City and the place with a few people was a park in Union city. The reason why we chose ShopRite, a market, as crowded space, rather than another crowded place like Stevens Library or shopping mall. The collected data are shown in Figure 5.

	Pattern	Block	Location	Time	Phone Cover	Battery	4G LTE Download Speed-1st Run	4G LTE Download Speed-2nd Run
1	++--	1	Wide_Space	Idle_Time	With_Cover	Low_Power_Mode	47.1	46.7
2	-+-+	1	Crowd_Space	Idle_Time	With_Cover	General_Mode	42.6	67.7
3	+-+-	1	Wide_Space	Busy_Time	No_Cover	Low_Power_Mode	55.1	54.4
4	--++	1	Crowd_Space	Busy_Time	No_Cover	General_Mode	21.7	25.2
5	-++-	2	Crowd_Space	Idle_Time	No_Cover	Low_Power_Mode	48.6	53
6	+--+	2	Wide_Space	Busy_Time	With_Cover	General_Mode	76.4	51.9
7	++++	2	Wide_Space	Idle_Time	No_Cover	General_Mode	79.3	78.6
8	----	2	Crowd_Space	Busy_Time	With_Cover	Low_Power_Mode	47.2	44.4

Figure 5: Collected data

5. **Data analysis**

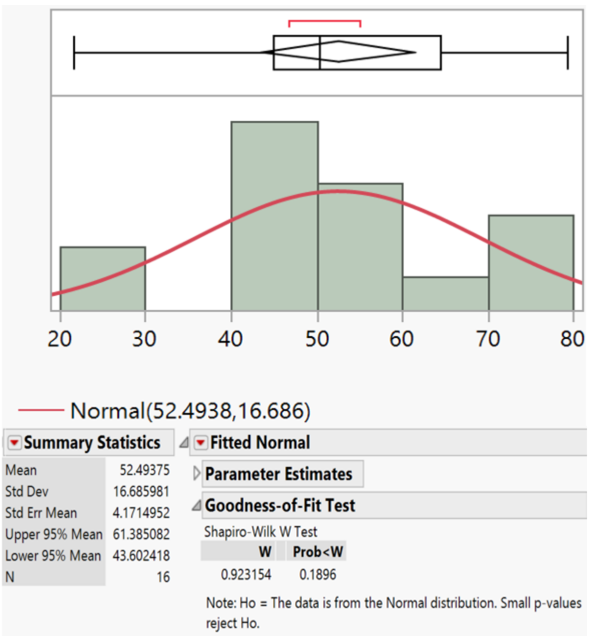


Figure 6: Normal distribution

First of all, we checked whether the data we collected were normally distributed. One thing we needed to clarify was that in our design, we set 95% as the confidence level. Even though, it seemed that some left skewness existed in Figure 6. However, as we could see, in this Goodness-of-Fit test, the p-value was greater than 0.05. Therefore, we could get a conclusion that our data is normally distributed.

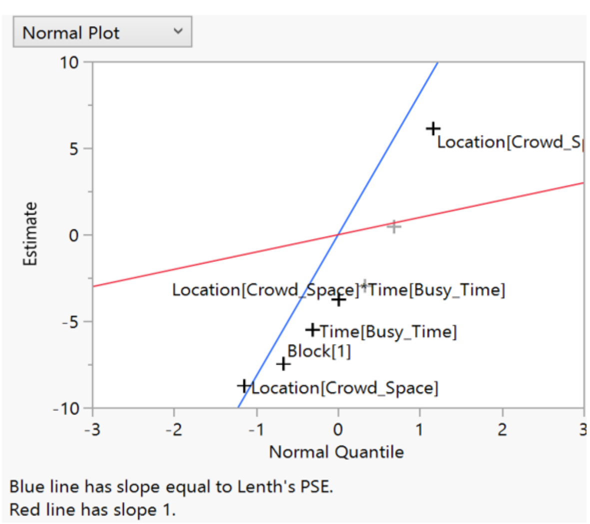


Figure 7: Normal plot

Second, using normal plot helped us find out[3] which factors were active factors or significant factors. In this normal plot (Figure 7), the vertical coordinate represented the value of the estimate and the horizontal coordinate represented its normal quantile. Normally, we could consider points that follow a line with a slope of σ, the blue line, as inactive effects. As we could see, this plot suggested that Location, Time and Block are active factors or significant factors. Other two factors, Phone cover and Battery mode, seem not important. But there might exist some interaction between these four factors. So, in the next step, we would dig out their interaction.

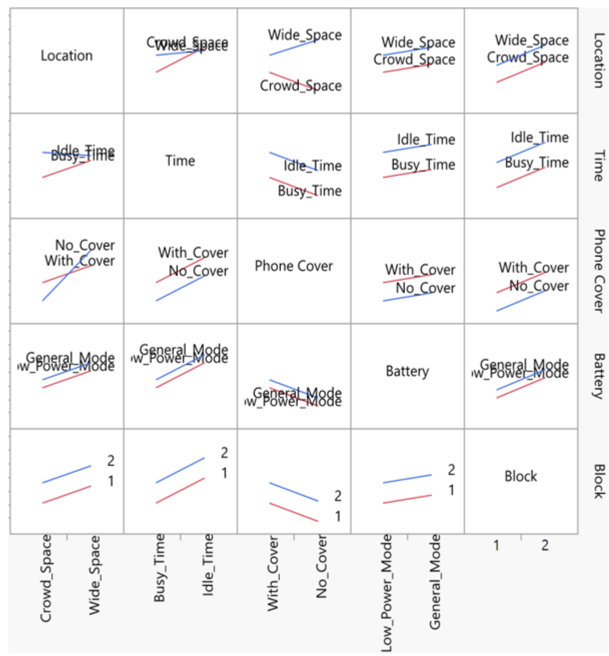
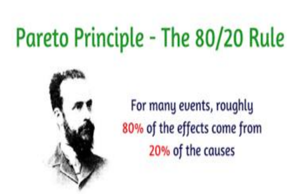


Figure 8: Interaction profiler

In this interaction profiles plot (Figure 8), non-parallel lines give evidence of possible interaction. Looking it carefully, we found that there may exists some interaction, between Location and Time, between Location and Phone Cover. However, we should further check their p-values before getting a conclusion that they are important effects and should be put into regression model.



Then, we used the Pareto plot (Figure 9) to detect effects which we should put into our regression model. The Pareto plot presented absolute values of the estimates in decreasing order. The grey curve summed the absolute values. As we could see, there were four effects above the dotted red line, which accounted for the 80% of the whole effects. They were Location, Block, the interaction between Location and Phone Cover, and Time.

Pareto Plot of Estimates

Term	Estimate
Location[Crowd_Space]	-8.693750
Block[1]	-7.431250
Location[Crowd_Space]*Phone Cover[With_Cover]	6.168750
Time[Busy_Time]	-5.456250
Location[Crowd_Space]*Time[Busy_Time]	-3.718750
Battery[Low_Power_Mode]	-2.931250
Phone Cover[With_Cover]	0.506250

Figure 9: Pareto plot

Remember, our confidence level is 95%. So, we needed to further check their p-values before summarizing a regression equation.

Scaled Estimates

Nominal factors expanded to all levels

Term	Scaled Estimate	Std Error	t Ratio	Prob>|t|	Lower 95%	Upper 95%
Intercept	52.49375	2.228061	23.56	<.0001*	47.355831	57.631669
Location[Crowd_Space]	-8.69375	2.228061	-3.90	0.0045*	-13.83167	-3.555831
Location[Wide_Space]	8.69375	2.228061	3.90	0.0045*	3.5558311	13.831669
Time[Busy_Time]	-5.45625	2.228061	-2.45	0.0400*	-10.59417	-0.318331
Time[Idle_Time]	5.45625	2.228061	2.45	0.0400*	0.3183311	10.594169
Phone Cover[With_Cover]	0.50625	2.228061	0.23	0.8260	-4.631669	5.6441689
Phone Cover[No_Cover]	-0.50625	2.228061	-0.23	0.8260	-5.644169	4.6316689
Battery[Low_Power_Mode]	-2.93125	2.228061	-1.32	0.2248	-8.069169	2.2066689
Battery[General_Mode]	2.93125	2.228061	1.32	0.2248	-2.206669	8.0691689
Block[1]	-7.43125	2.228061	-3.34	0.0103*	-12.56917	-2.293331
Block[2]	7.43125	2.228061	3.34	0.0103*	2.2933311	12.569169
Location[Crowd_Space]*Time[Busy_Time]	-3.71875	2.228061	-1.67	0.1337	-8.856669	1.4191689
Location[Crowd_Space]*Time[Idle_Time]	3.71875	2.228061	1.67	0.1337	-1.419169	8.8566689
Location[Wide_Space]*Time[Busy_Time]	3.71875	2.228061	1.67	0.1337	-1.419169	8.8566689
Location[Wide_Space]*Time[Idle_Time]	-3.71875	2.228061	-1.67	0.1337	-8.856669	1.4191689
Location[Crowd_Space]*Phone Cover[With_Cover]	6.16875	2.228061	2.77	0.0243*	1.0308311	11.306669
Location[Crowd_Space]*Phone Cover[No_Cover]	-6.16875	2.228061	-2.77	0.0243*	-11.30667	-1.030831
Location[Wide_Space]*Phone Cover[With_Cover]	-6.16875	2.228061	-2.77	0.0243*	-11.30667	-1.030831
Location[Wide_Space]*Phone Cover[No_Cover]	6.16875	2.228061	2.77	0.0243*	1.0308311	11.306669
Location[Crowd_Space]*Battery[Low_Power_Mode]	0	0	0.00	1.0000	0	0
Location[Crowd_Space]*Battery[General_Mode]	0	0	0.00	1.0000	0	0
Location[Wide_Space]*Battery[Low_Power_Mode]	0	0	0.00	1.0000	0	0
Location[Wide_Space]*Battery[General_Mode]	0	0	0.00	1.0000	0	0
Time[Busy_Time]*Phone Cover[With_Cover]	0	0	0.00	1.0000	0	0
Time[Busy_Time]*Phone Cover[No_Cover]	0	0	0.00	1.0000	0	0
Time[Idle_Time]*Phone Cover[With_Cover]	0	0	0.00	1.0000	0	0
Time[Idle_Time]*Phone Cover[No_Cover]	0	0	0.00	1.0000	0	0
Time[Busy_Time]*Battery[Low_Power_Mode]	0	0	0.00	1.0000	0	0
Time[Busy_Time]*Battery[General_Mode]	0	0	0.00	1.0000	0	0
Time[Idle_Time]*Battery[Low_Power_Mode]	0	0	0.00	1.0000	0	0

Figure 10: Scaled estimates plot

Through the plot of scaled estimates, we verified that there were five effects with p-value less than 0.05, including the four effects we already detected in pareto plot, plus the intercept. Based on these outputs, we got this regression equation as follows:

Estimated Speed=

8.69375*[Location: Wide Space]

+5.45625*[Time: Idle Time]

+7.43125*Block[2]

+6.16875*[Location: Wide Space]*[Phone Cover: No Cover]

+52.49375

Highest speed: 80.24375

Summary of Fit

RSquare	0.847851
RSquare Adj	0.71472
Root Mean Square Error	8.912246
Mean of Response	52.49375
Observations (or Sum Wgts)	16

Figure 11: RSquare

From this equation, we could expect the highest download speed reach about 80 megabits per second. This is the pretty high speed which we might satisfy with. But did this equation model fit our data well? In Figure 11, the adjusted $\boldsymbol{R^2}$ reached 0.71472. That is, nearly 71.4% of total variance had been explained by this regression model. This regression equation was not perfect but still worked.

6. Conclusion and recommendation

Our conclusion is that if you were at the place with a few people during the idle time, surfing the internet via your latest iPhone without phone cover, then you are close to the best Internet connection.

This perfect situation, however, does not always exist for every smartphone user. Then, we recommend some ways to live a better life with the Internet connection based on our study results.

First, the place matters most. If you plan to use Internet for some recreational activities like watching high-definition videos at a bus station, supermarkets and etc. This is not a good plan. You would better download some high-definition videos in advance.

Second, it is an advisable strategy that you focus on work or study in business hours. Because this period from 7 am- 11 pm will have a bad influence on the Internet connection. The self-fulfillment that you find in your work or study will give you more happiness than in surfing the Internet.

7. Future work

As regard this experiment, from defining the problem to designing the experiment matrix, four aspects at least can be improved.

First, as far as the design matrix, we can use full factorial design to achieve the higher accuracy and also perform the One Factor At A Time (OFAT). Then, it will get a more reasonable conclusion by comparing full factorial design, OFAT with FFD.

Second, in this experiment, we focused on this problem how do we reach the highest download speed. However, the upload speed may also matter to some smartphone users. Hence, some extension of our study problem can be done in future.

Third, there may exist more factors that we did not realize and did not involve in our design.

8. Acknowledgement

We would like to express our deepest appreciation to our professor (Chihoon Lee) who not only deliver us great classes, provide us interesting reading materials and useful online learning resources, help us conceptualize a real project, but also give full support for our project work and guide us all along. Without his encouragement, support and guidance, we could not gain so much from this course.

References

Gordon, H. B., Johannes, L., & Arthur, J. S. (2006). Experimental design on the front lines of marketing: Testing new ideas to increase direct mail sales. Intern. J. of Research in Marketing 23 (2006) 309–319.

The reason to use phone cover

https://t2conline.com/importance-of-a-mobile-phone-cover/

Low power mode

https://support.apple.com/en-us/HT205234