Understanding the impact of the Weather on Human Mobility via LTE Access Traces in Seoul Districts

Zofia Marciniak
KAIST
zofia.marciniak@kaist.ac.kr

Sumin Han
KAIST
suminhan@kaist.ac.kr

Dongman Lee
KAIST
dlee@kaist.ac.kr

Abstract

In the urban area, there are many activity choices such as shopping or staying home. One of the main factors which affects people's choices is the weather. In this study, the influence of weather conditions on human activity in the urban area is explored. We trace the change in the intensity of the mobile LTE access usage under different weather conditions. We find that pressure and rain are the most impactful conditions, while most affected groups are people in their 20's and older. People in their 30's seem to be the least affected and follow their usual activity pattern. Our result can provide insights for spot recommendation, traffic prediction, and sales prediction.

Keyword

LTE access trace, weather, activity

1. Introduction

In the urban area, there are many activity choices such as shopping or staying home. One of the main factors which affects people's choices is the weather. For example, people may go outside for shopping during the sunny days or stay indoors during the rain.

The correlation between the weather and human activity has been studied in several research. These studies understand and confirm certain correlations between weather and activity types [1], mobility of older adults [2] and retail sales [3]. However, they focus on a set of individuals or places and do not explore the weather impact on the city areas in the bigger image.

In this study, we use LTE access trace data to find the impact of the weather on popularity of different types of area. We measure intensity of human activity under different weather conditions using the weather. We find that women, people in their 20's and olders are most likely to be affected by the weather conditions, especially pressure and rain, while people in their 30's are the least affected by all conditions. We believe our findings can provide insights for spot recommendation, traffic prediction, or sales prediction.

2. Previous work

Brum-Bastos et al. [1] leverage GPS trajectories collected from participants and meteorological data to show travel mode and amount of time spent at different places under weather conditions.

Clarke et al. [2] conduct telephone surveys to find the correlation between the weather and the older adults' activity. They find that older adults' decisions for where to travel are affected by different weather changes.

Badorf et al. [3] makes an analysis based on daily sales data from retail shops and weather data from all across the country. They calculate the impact of weather on sale depending on store location and theme to show that weather changes should be included in sales prediction models.

3. Dataset

In this section, we describe how we construct our LTE access trace data and weather data.

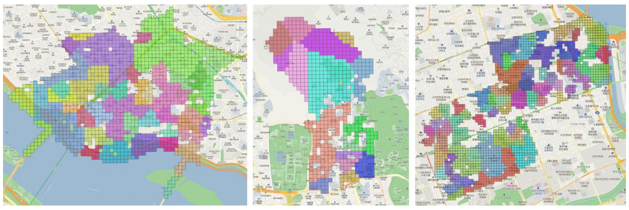

Figure 1. LTE data in the commercial areas in Seoul (Hongik, Anguk, Samsung subway station area). The color represents the similar LTE access trace trend pattern after seasonal decomposition, and grouped by Pearson correlation over 0.95.

3.1 LTE access trace data

The LTE access trace data consists of the numbers of anonymized people in different age groups, who access the LTE cell towers in an hourly timestamp. The dataset covers 4,577 LTE cell towers around the commercial area as described in Figure 1, from 1st March 2018 to 28th February 2019.

3.2 **Weather data**

We crawl the weather records from Seoul City weather station via a public website[1]. The dataset consists of temperature, humidity and pressure measurements every 3 hours (00:00, 3:00, 6:00, 9:00, 12:00, 15:00, 18:00, 21:00) and precipitation measurement every 12 hours (9:00, 21:00).

4. Methods

4.1 Clustering LTE cells

We cluster the adjacent LTE cells into the region groups which show a similar time series pattern. We first extract seasonal patterns from STL time series decomposition for each LTE access trace. Then, we calculate dissimilarity of each adjacent two LTE cells by Pearson Correlation. We group each pair of two cells if the Pearson Correlation is over 0.95. In result, the total of 133 clusters are found in our LTE cells as Figure 1.

Then we label the types of each cluster manually using street view images into commercial, residential, and transportation. Each category is evaluated based on the general outlook of the area. For residential, we look for the presence of apartment buildings or houses. For commercial, we look for restaurants or clothes stores. For transportation, we look for a big road with bus stops or subway stations.

Figure 2. Examples of residential (visible living buildings, apartaments), commercial, transportation area (big road type).

4.2 Grouping the weather data values

We consider rain, temperature, humidity, and pressure from the weather dataset. We use those four entries to cluster the weather clusters using KMeans Clustering for K = 3. The clusters show features as in Table 1. Cluster-0 is characterized as hot weather, as shows the high temperature. Cluster-1 is characterized as the cold weather, as it shows the high pressure and the low temperature. Cluster-2 is characterized as the rainy weather, as it shows the high precipitation and the high humidity.

[1] http://rp5.ru/

Table 1. Shows for each cluster its minimum, maximum and average for every weather entry type.

Cluster	Temperature (in °C)	Precipitation (in mm)	Humidity (in %)	Pressure (in mmHg)
0 (hot)	min 4.4625 max 33.6875 avg **20.897**	min 0 max 11.1 avg 0.261	min 33.125 max 69.875 avg 57.911	min 737.72 max 764.49 avg 751.95
1 (cold)	min -14.8125 max 24.125 avg **3.1140**	min 0 max 6.0 avg 0.0775	min 16.875 max 70.375 avg 47.363	min 746.38 max 770.09 avg **759.77**
2 (rainy)	min -5.7875 max 30.525 avg 17.95	min 0 max 78.0 avg **6.1326**	min 67.000 max 97.375 avg **78.308**	min 735.09 max 765.56 avg 750.98

4.3 Join analysis

After both area types of LTE cell towers and weather clusters are set, we join these datasets to analyze how much the area is exploited on different weather for each type of area: commercial, residential, and transportation. We take the average of the hourly LTE access count of everyday.

The LTE data entries are grouped into weather categories based on the weather cluster number. From the clustered data entries, we take mean values of each age group. We scale the data by changing the unit into percentage. Therefore, for each age group, we can see the percentage change in activity by weather condition.

5. Result

First of all, we can see that the cluster-hot shows the lowest active people using LTE from the results in Figure 3.

For the commercial area, women are much more active than men (Table 2). Also, women's activity drastically decreases from weather cluster-hot to cluster-rainy, more than men. Meaning, women are more likely to be affected by rainy weather in commercial areas. People in their 20's and 25's are the ones that decrease their activity the most for weather cluster-cold and cluster-rainy (Table 3). Age groups from 35 to 45 do not show much change in activity between clusters (but it decreases with each cluster). For people in their 50, they show the most activity in cluster-cold, but it is very close to cluster-hot and cluster-cold. It shows that for commercial area activities, older people might prefer cold weather rather than hot one.

For the residential area, both men and women show similar activity rate, with slight dominance of women activity in cluster-hot and cluster-cold, and slight dominance of men in cluster-rainy (Table 4). For the age groups activity (Table 5), the activity rate in cluster-cold and cluster-rainy is almost identical.

Similarly to the commercial area, people in their 20-30's are much more active during the cluster-hot.

For the transportation area, dominance of activity of men is significant (Table 6). In this case, the activity of men in transportation areas is the lowest during cold weather. This leads to a conclusion, that men are less likely to use transportation area during cold weather. People aged 30-45 are more likely to use transportation area during rainy weather than cold weather (Table 7). The biggest gap is in activity of people in their 25's between hot weather and cold weather, where hot weather dominates. It follows with people aged 30-35, who are less, but also more active during hot weather type, The trend decreases for people aged 20, 40, 45 and 50.

6. Discussion

In our result, it shows several correlations between age and the weather, with emphasis on differences during high temperature and strong correlations especially for young people. Our data was collected over too short time to get a reasonable result with a prediction model, as weather conditions are repetitive in yearly time intervals. However, in further analysis with a bigger dataset, it would be possible to create a prediction model that predicts people's activity during certain weather conditions.

7. Conclusion

We find there is an impact on the activity and area where people spend their time during certain weather types. This could be applied in developing the advertisement systems and targeted advertisements. With these observations, the advertisement system could be built based on the weather forecast and applied in a transportation means. The results can help understand the human interest in going out and choosing different kinds of activities. By making the analysis, I wish to achieve better understanding of human interest and improve spot recommendation and advertisement based on weather conditions.

Acknowledgment

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.2019-0-01126, Self-learning based Autonomic IoT Edge Computing). We appreciate Korea Telelcom(KT) for providing valuable LTE access data.

Reference

1. Brum-Bastos, V.S., Long, J.A., Demšar, U., 2018. Weather effects on human mobility: a study using multi-channel sequence analysis. In: Computers, Environment and Urban Systems. 71. Pergamon, pp. 131–152.
2. Clarke, P. J., Yan, T., Keusch, F., & Gallagher, N. A. (2015). The Impact of Weather on Mobility and Participation in Older U.S. Adults. American journal of public health, 105(7), 1489–1494. https://doi.org/10.2105/AJPH.2015.302582
3. Badorf, F., & Hoberg, K. (2020). The impact of daily weather on retail sales: An empirical study in brick-and-mortar stores. Journal of Retailing and Consumer Services, 52. https://doi.org/10.1016/j.jretconser.2019.10192

Appendix

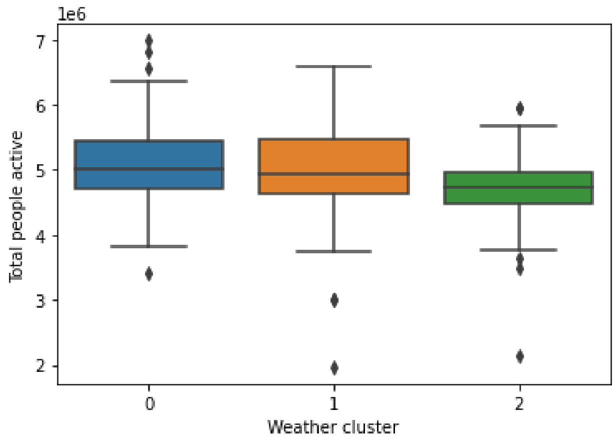

Figure 3. Total activity within different weather clusters

Gender	Cluster-hot	Cluster-cold	Cluster-rainy
Men	2010.0	1934.7	1826.4
Women	2065.5	2033.8	1842.6

Table 2. Gender-grouped average activity in commercial cluster (in people)

Age group	Custer-hot	Cluster-cold	Cluster-rainy
20	1119.6	1081.7	992.8
25	1026.6	991.1	900.5
30	590.8	573.1	528.2
35	444.6	440.7	409.3
40	317.8	309.7	297.0
45	322.5	317.3	301.7
50	253.5	254.9	239.5

Table 3. Age-grouped average activity in commercial area (in people)

Gender	Cluster-hot	Cluster-cold	Cluster-rainy
Men	917.4	874.7	875.5
Women	923.8	895.3	871.3

Table 4. Gender-grouped average activity in residential area (in people)

Age group	Custer-hot	Cluster-cold	Cluster-rainy
20	308.5	289.5	289.5
25	417.6	397.1	390.7
30	319.7	300.2	298.8
35	271.6	260.7	258.1
40	196.4	194.9	191.5
45	187.5	183.9	182.0
50	140.0	143.6	136.2

Table 5. Age-grouped average activity in residential area (in people)

Gender	Cluster-hot	Cluster-cold	Cluster-rainy
Men	972.3	867.2	890.5
Women	752.8	686.3	673.8

Table 6. Gender-grouped average activity in transportation area (in people)

Age group	Custer-hot	Cluster-cold	Cluster-rainy
20	281.9	261.4	252.9
25	400.5	357.0	351.9
30	287.9	251.5	257.0
35	244.1	216.7	222.8
40	180.7	160.5	169.9
45	181.8	167.1	171.5
50	148.1	139.3	138.4

Table 7. Age-grouped average activity in transportation area (in people)