

Modeling and Visualizing Regular Tweeting Pattern of Purdue Campus

Yuqian Huang (Geomatics, Lyles School of Civil Engineering)

Introduction

Objectives

With the strength of social media, many crowd-sourced sensing and collaboration projects can benefit from this kind of **Volunteered Geographic Information (VGI)**. And this project is trying to find out the potential **spatial and temporal pattern** of the twitter user and provide the probabilities of the clusters.

Data

The data for this project is the Geo-tagged Tweets around Purdue Campus from Jan 1st 2014 to Sep 30th 2015. There are totally 114,937 tweets contributed by 5,445 users.

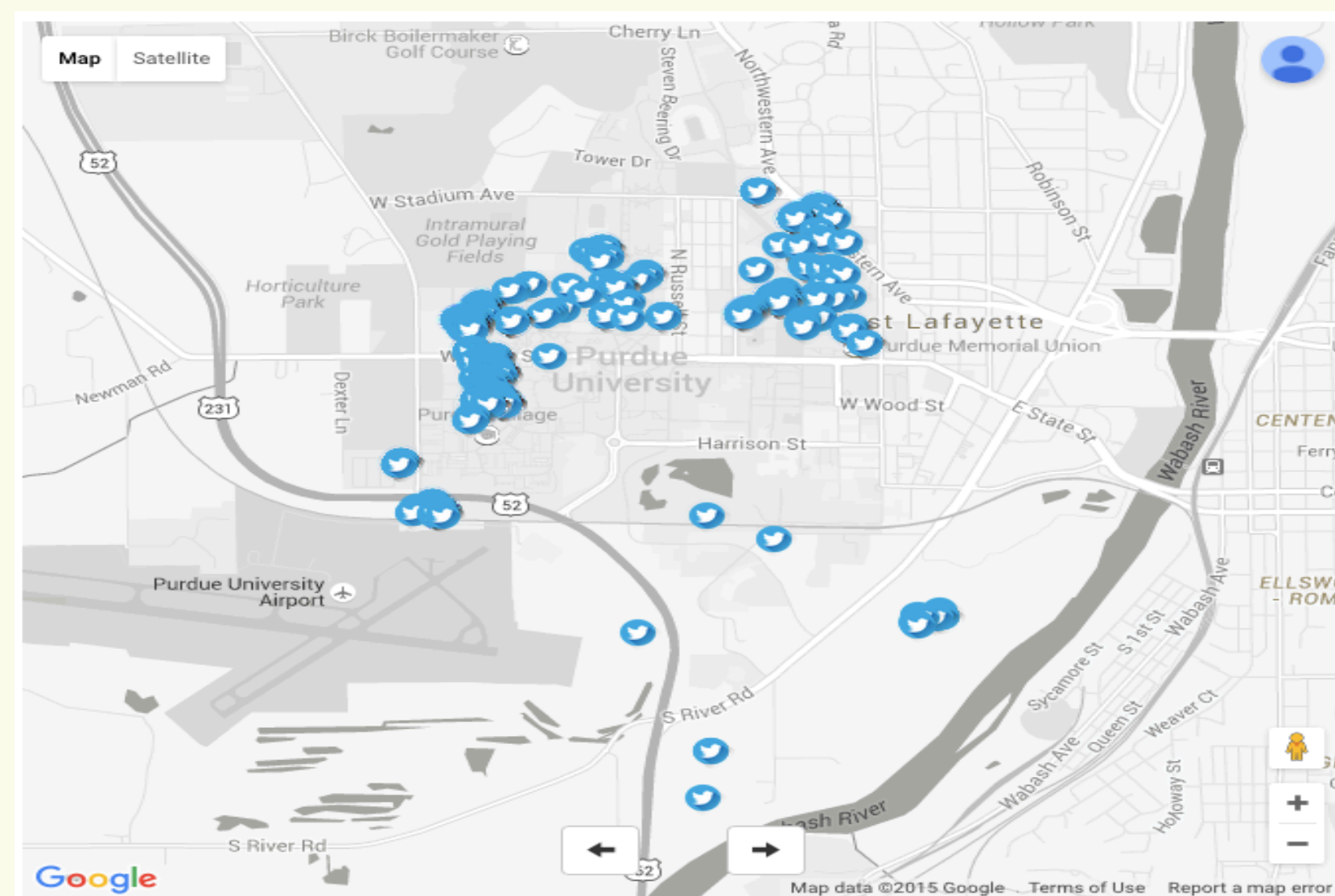


Fig. 1 Sample of Tweets from the Most Active Users in 2014

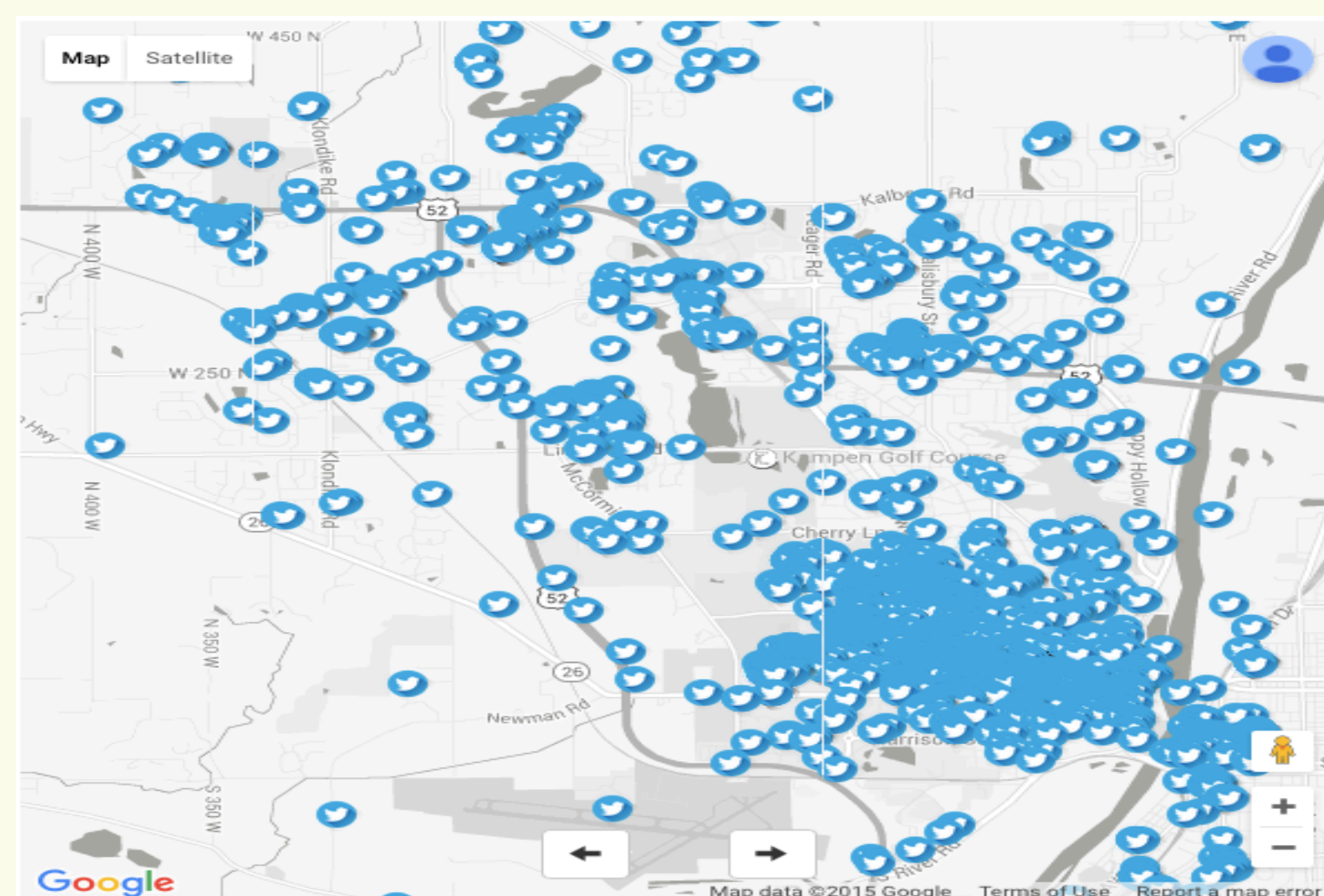


Fig. 2 Sample of Tweets from the Most Active Users in 2015

Methodology

- Clustering: Density-Based Spatial Clustering of Applications with Noise (DBSCAN)
- Keyword Detection: Descending order of word frequency

Processing Procedures

1. Group user's tweets depends on the hour of the tweets and apply DBSCAN to detect potential clusters.

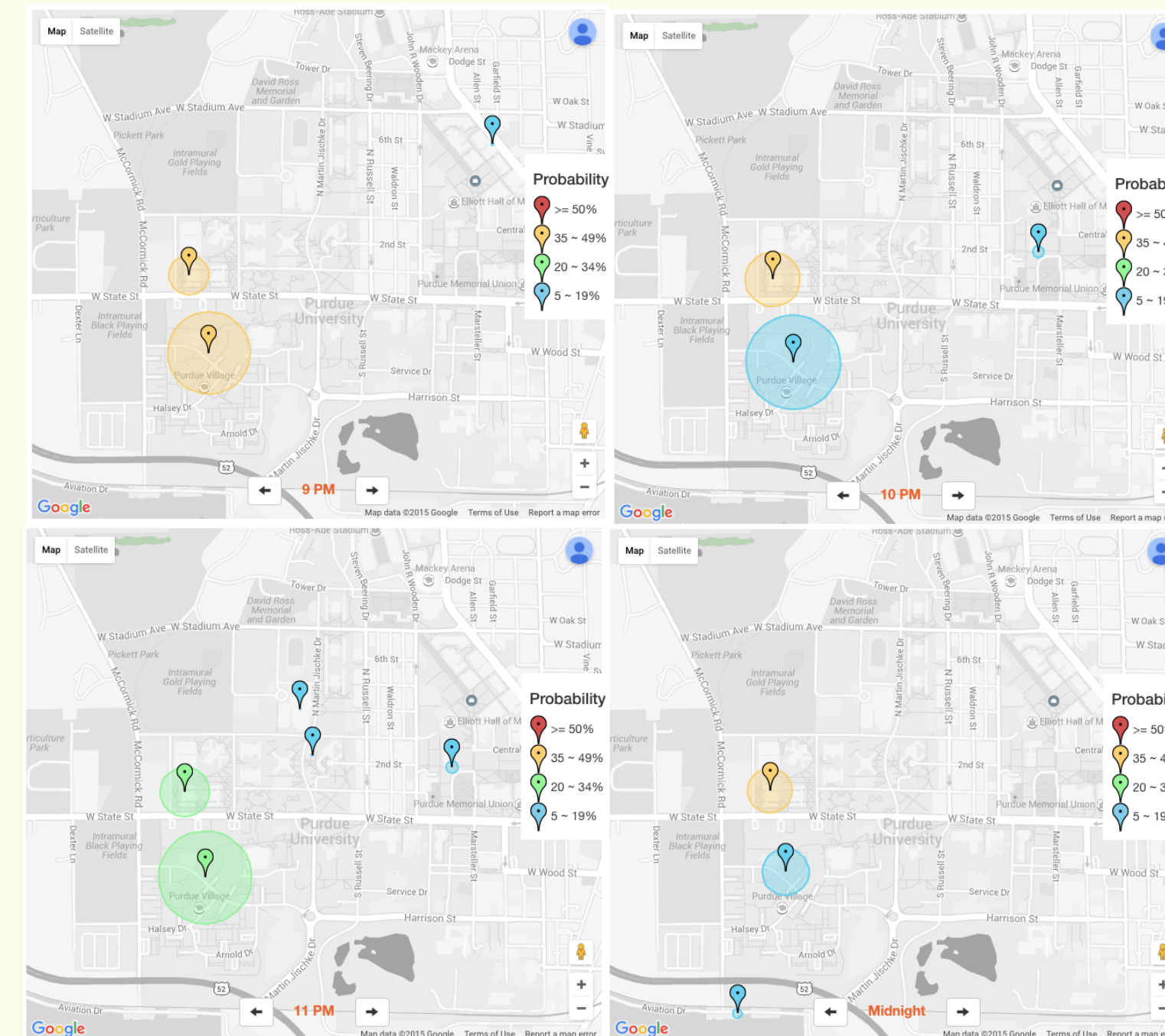


Fig. 3 Patterns from a User in Different Period(9pm, 10pm, 11pm, midnight)

2. Calculate spatial and temporal probability, radius, number of tweets and keywords of every cluster.

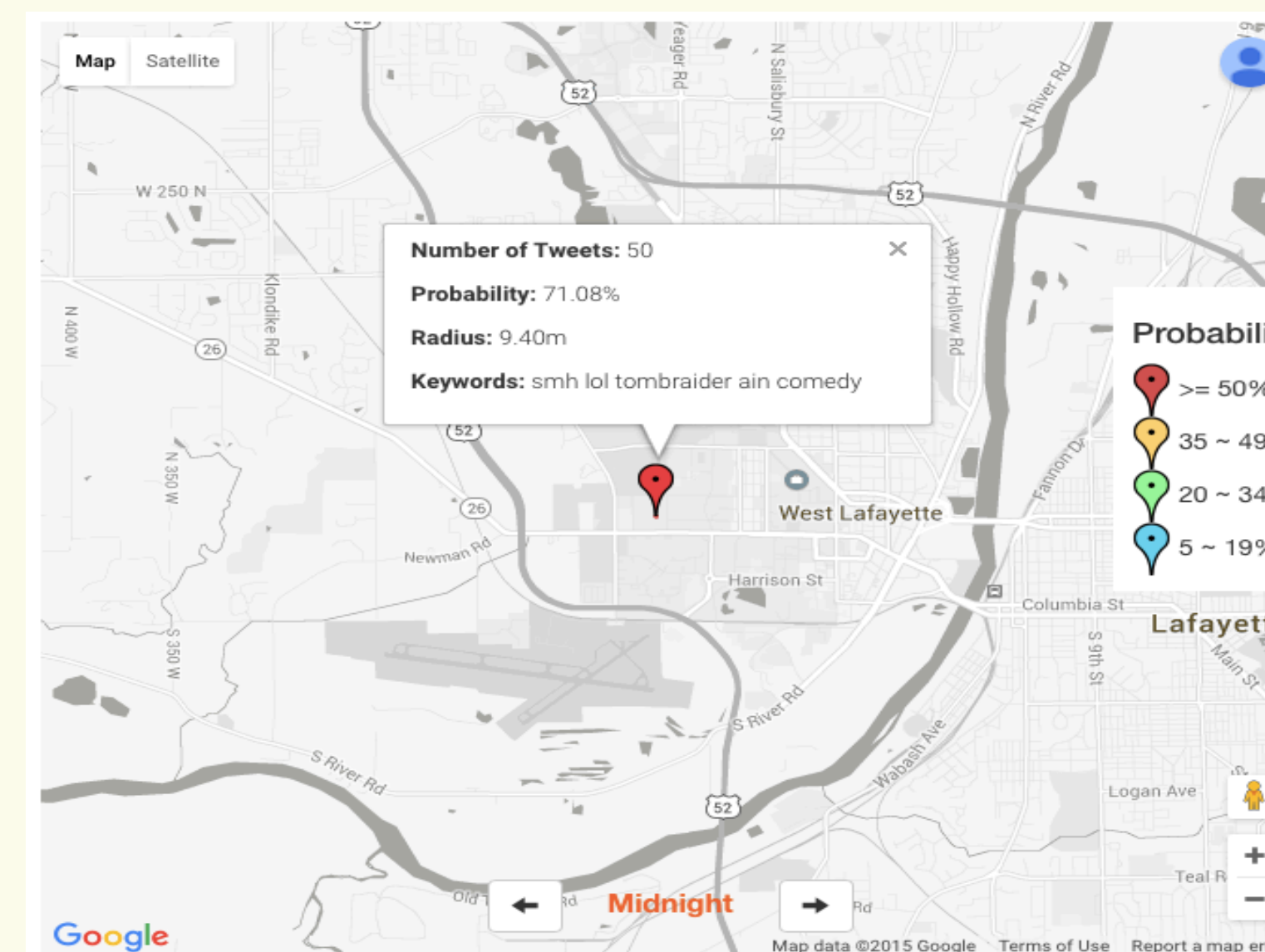


Fig. 4 Information Window of a Cluster

3. Summarize the probabilities of different types of clusters together into a tweeting frequency bar chart.

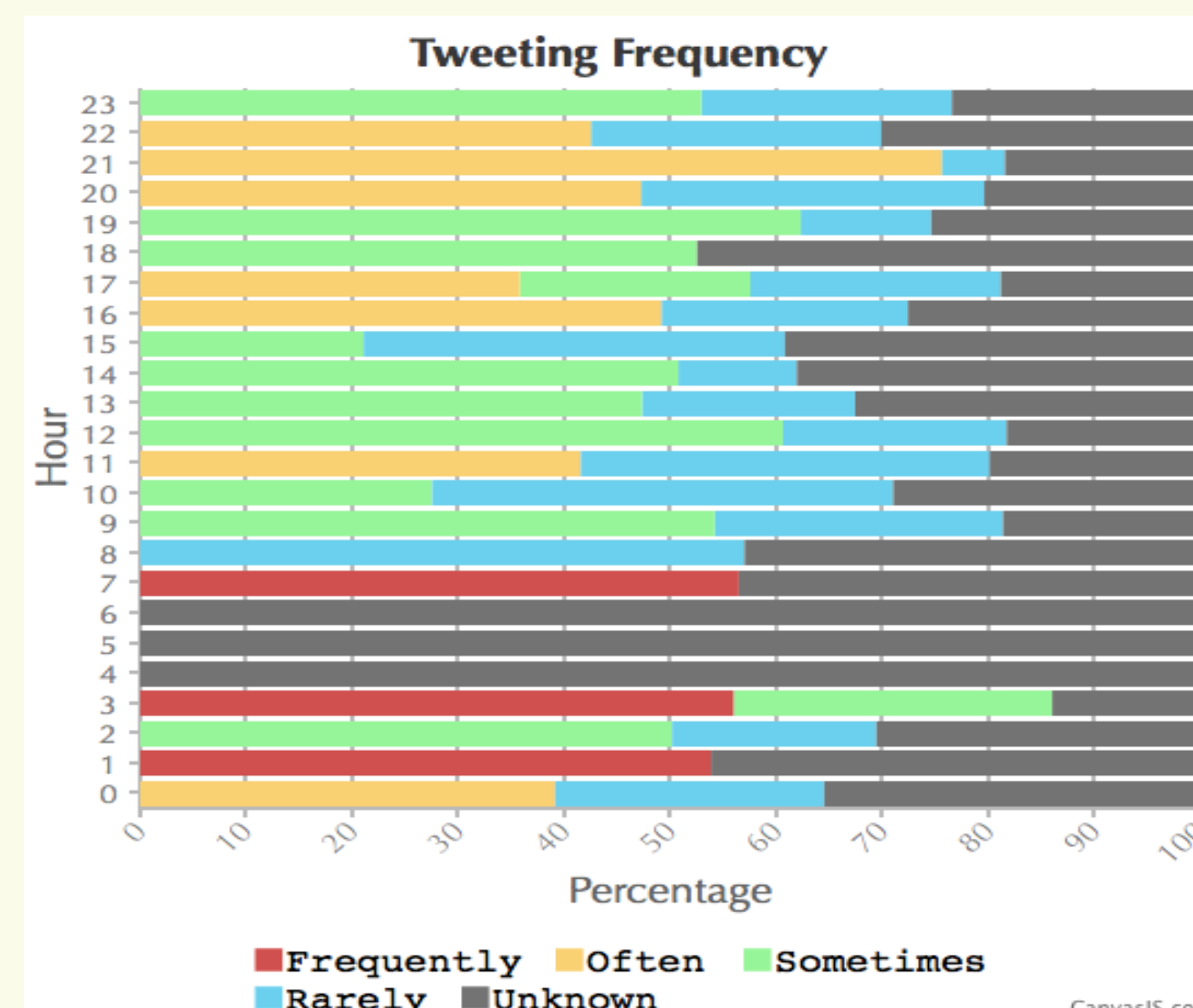


Fig. 5 Tweeting Frequency Bar Chart

Cases Analysis

Changes of Tweeting Pattern

These two bar tweeting frequency bar charts are from the same user but in different years. As what we can see in the bar charts, this user tended to tweet in specific locations in 2014, while the user tweeting pattern changes quite a lot in 2015. Because there are only blue bars on the chart, which indicates that the user has a really high diversified tweeting pattern, and it is unlikely to predict where the user is in 2015 with only these clues.

Hey! This guy has moved away!

These two maps show a user's patterns from the same period in 2014 and 2015. It is clear that this user has changed the usual tweeting location to another one. And because it is in the midnight (12 am), most students usually stay in their dormitories. Hence, according to these clues, a reasonable guess from these two figures is that this user has moved from Owen Residence Hall to Hilltop.

Tiny Groups

Beside doing clustering on specific users, monthly data are also discussed in this project. These two figures are both from Dec 2014, but they look different. This is because of the change of the definition of "Rarely" from "5 ~ 19%" to "0.5 ~ 19%", and the huge difference comes after the change. This indicates that there are actually a lot of tiny groups, where there are only some tweets around specific locations. This high diversity is somehow reasonable when combining all the peoples around the campus, because people's patterns are so different from each others'!

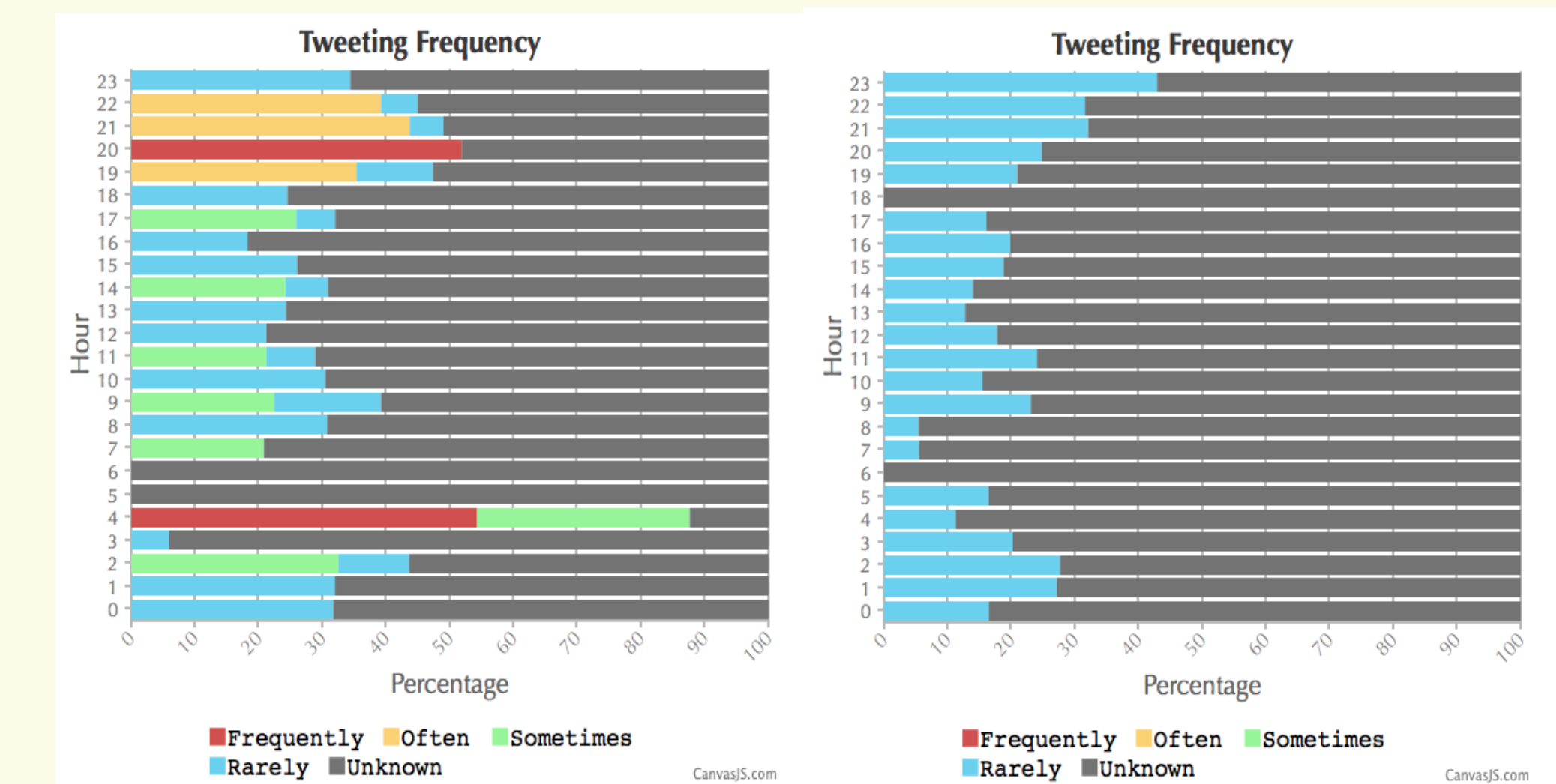


Fig. 6 Bar Chart from a User in 2014 and 2015

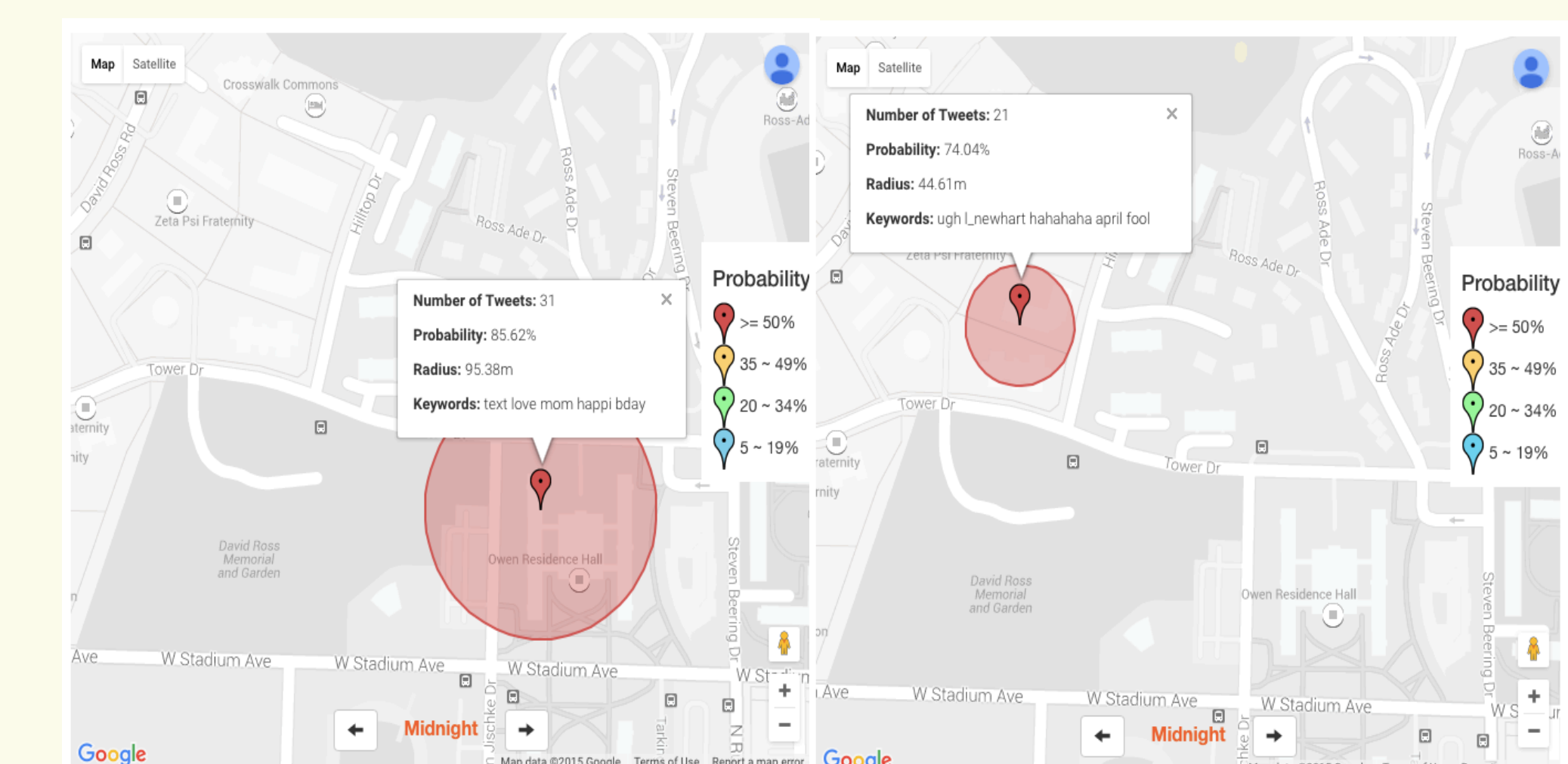


Fig. 7 Cluster Center of Pattern from a User in 2014 and 2015

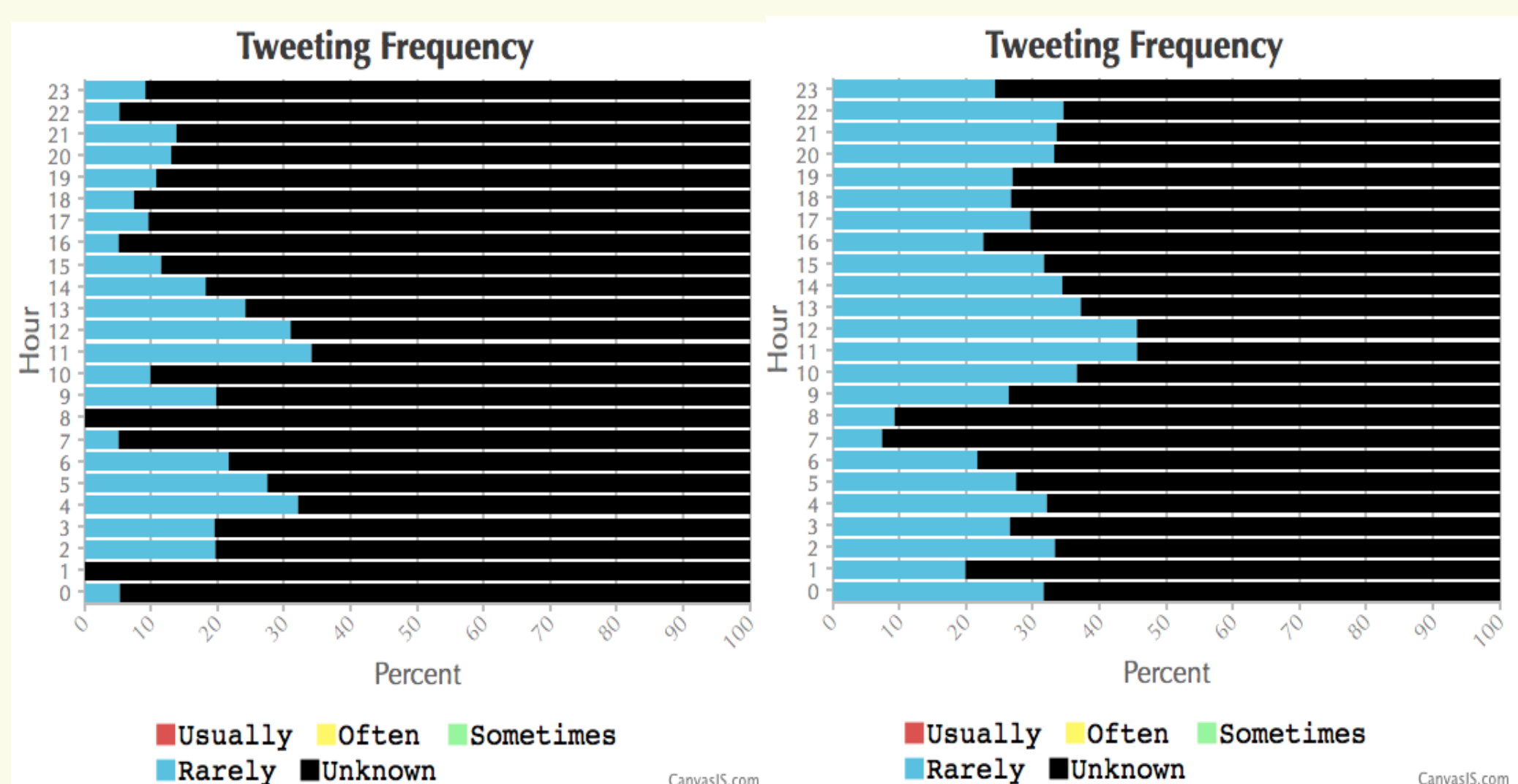


Fig. 8 Cluster Pattern from Dec 2014 with different range of "Rarely" (Left: 5 ~ 19%, Right: 0.5~19%)

Conclusion

This project, successfully detect potential patterns of the most active users, not only provides a general intuition of spatial and temporal patterns, but it also provides measurement of probability for every clusters. But when facing month data, it is difficult to see through with the tiny groups problems and find out the majority. To figure out this question and improve the analysis result, testifying this method on other campus datasets would be helpful. If you are also interested in this topic, please go to the website of this project and hope you enjoy it: <http://purduetweets.azurewebsites.net/>

Reference

Qunying Huang & David W. S. Wong (2015): Modeling and Visualizing Regular Human Mobility Patterns with Uncertainty: An Example Using Twitter Data, Annals of the Association of American Geographers, DOI: 10.1080/00045608.2015.1081120