Movement and Network Visulization all in R Studio!

VAST2021-MC2 Using R and Shiny

Miao Lu[1] Zhang Xinru[2] Zhong Ping[3]
Singapore Management University MITB ISSS608 G2-T6

Issue and Problem

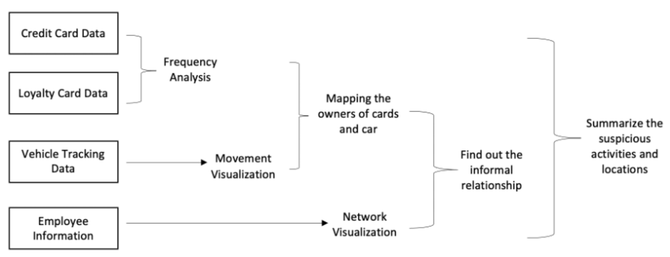
In January 2014, during a celebration of GAStech,a natural gas production site in Kronos, several employees go missing. An organization known as the Protectors of Kronos (POK) is suspected in the disappearance, but things may not be what they seem. We have two separate datasets about transaction history of credit card, loyalty card. But we don't know who each card belongs to. Besides, we also have a vehicle GPS history dataset and car assignment table. But the GPS data only contains coordinate data, and the volume is too huge to analyze.

Motivation

Recognizing mapping relationships among credit card, loyalty card, car ID and employee can be very time-consuming and complicated. Besides, recognize the informal relationships and explain it efficiently is not easy. With the help of R, Shiny app, and other use open-source packages, we developed a simple application to analyze data and visualize the mapping relationships between cards and employee as well as the informal relationships we found.

Approach

At first, we will perform the Frequency Analysis for each store to find the popular ones. Then we visualize the vehicle movement data to analysis the daily route of each car. Combine these two approach, we can map the ownership of car and cards that is visualized in Parallel Plot, and dig out the informal relationship that is visualized in Network Graph.At last, we summarized the suspicious locations from all the analysis and visualization.



Key Output and Result

Location Frequency Analysis

Among those 2 weeks, Katerina's Café, Hippokampos, Guy's Gyros, Brew've Been Served are the more popularity locations. Katerina Café is most popular in 8 pm dinner time, Hippokampos is most popular in 1pm lunch time, and Guy's Gyros is as popular at 1pm lunch time as it is at 8pm dinner time, while Brew've Been Served is most popular in 7-8am.

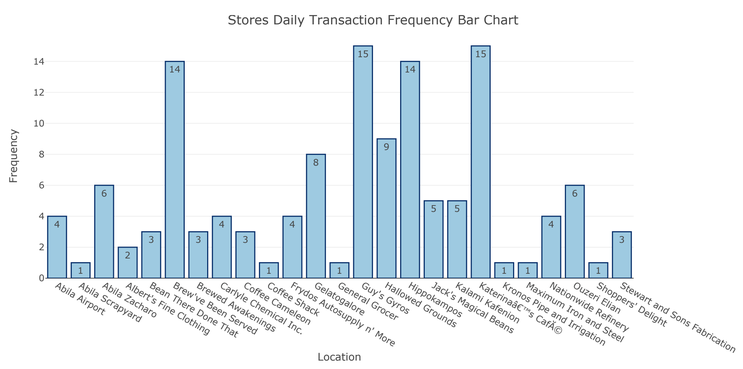


Figure 1: Stores Daily Transaction Frequency Bar Chart on 01/09/2014

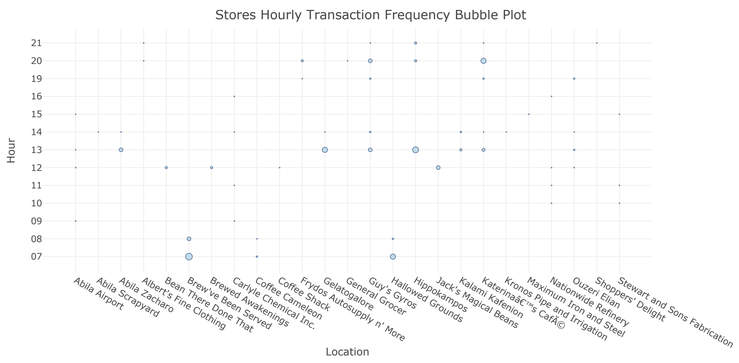


Figure 2: Stores Hourly Transaction Frequency Bubble Plot on 01/09/2014

Movement Visualization

The movement map is in an overwhelming position in the project to help us visualize the daily route of each car, map the routes and transaction activities to find the ownership of cars and cards, dig out the informal relationship and summarize the suspicious locations involved in the kidnapping.

Figure 3: Car Movement Map of Car2 in 01/06/2014

Card Ownership Analysis

The parallel plot is used to visualize the ownership among Car, Credit Card and Loyalty Card, involving 2 kinds of ownership that are the normal ownership and abnormal sharing cards relationship shown in two plots separatly.

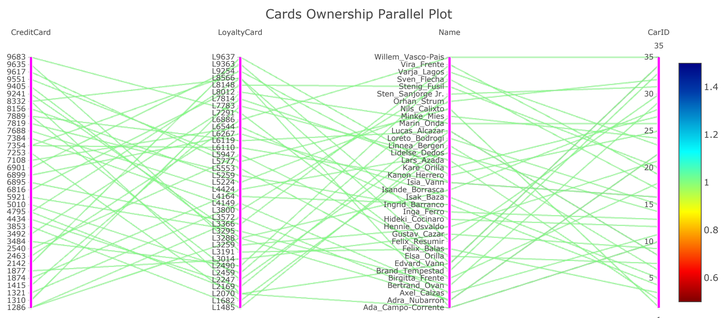


Figure 4: Cards Ownership Parallel Plot

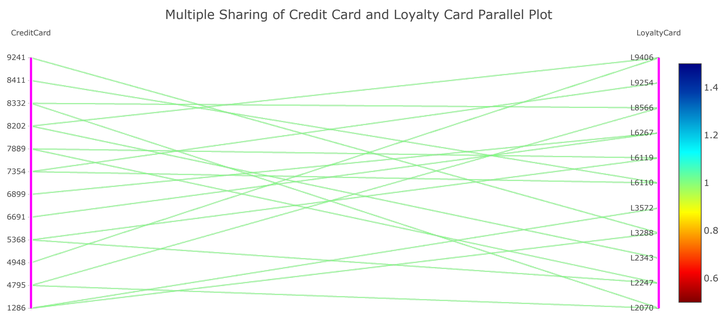


Figure 5: Multiple Sharing of Credit Card and Loyalty Card Parallel Plot

Network Analysis

In total, there are 29 employees have 9 kinds of informal relationships out of all 35 employs, including living together, visiting hotel privately, playing golf together, department gathering, suspicious/midnight visit, dining together, sharing cards and meeting at wasteland.

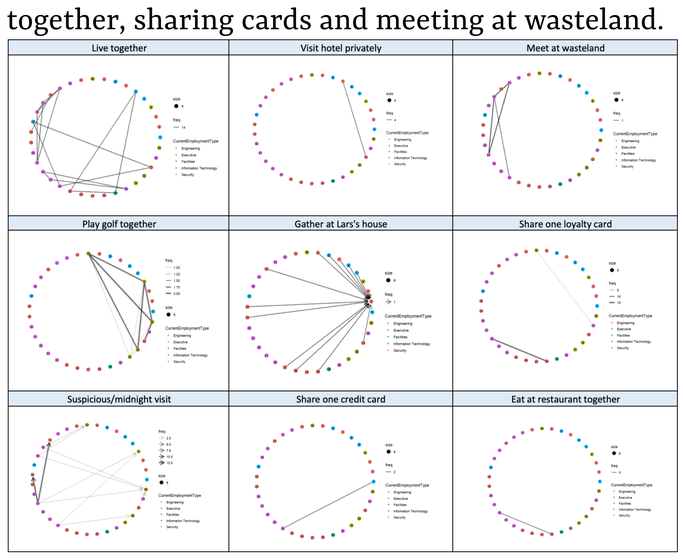


Figure 6:Informal Relationship Network Graph

Suspicious Locations Analysis

11 suspicious locations were summarized at last, including house of owners of Car 2, 14, 18, and executive team, Chostus hotel, site controller's working places and some stores with abnormal transaction frequency. Suspicious activities happened there includes midnight department gathering, security team's monitor to executive team, private gathering and the surge of transactions of some locations at the night before the disappearance.

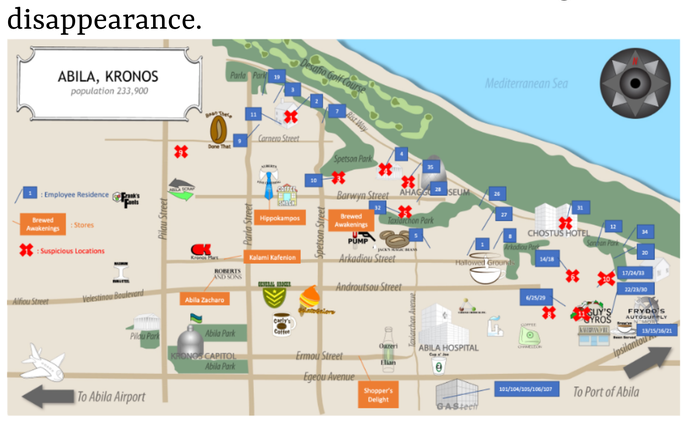


Figure 7: Suspicious Location Map

Limitation and Future Work

1. Some employee's GPS may get broken. Car 9's GPS cannot collect all data, and Car 28's GPS points were upper than the actual position. It is hard for us to tell it is broken accidentally or intentionally that they want to cover their real routes, which means that we cannot evaluate the suspicion of them.

2. The vehicle data is not continuous and lacking of interactivity among different cars. In other words, there is no GPS data if the car stopped. Ideally, if we get the GPS data in very single minutes, we can select a specific time to check whether some employees are stay together.

In the future, if these limitations can be overcome by interviewing the suspicions and implementing more data wrangling techniques, we can draw a more accurate guess and improve the fitness of our visualization application.

References

1.What is Tracking Analyst? —ArcMap | Documentation. (2020). Arcgis. https://desktop.arcgis.com/en/arcmap/latest/extensions/tracking-analyst/000-what-is-tracking-analyst.htm

2.Hajjar, A. J. (2021, August 5). Graph Analytics: Types, Tools, and Top 10 Use cases. AIMultiple. https://research.aimultiple.com/graph-analytics/

3.Geovisualization. Geovisualization - an overview | ScienceDirect Topics. (n.d.). https://www.sciencedirect.com/topics/earth-and-planetary-sciences/geovisualization.