Regular and Unusual Data Visualization IIIT-H

P Yashaswi* IIIT-H
Yarrabelly Navya† IIIT-H
Veera Raghavendra Chikka‡ IIIT-H

Abstract

VAST 2014 Mini Challenge 2 was a rolling challenge where the contestants were asked to submit , the daily routines and unusual patterns among the employees of GASTech employees which would in turn fetch information for investigation of GASTech dissapearences .Participants were given access to geospatial data related to vehicles and the personal/ business credit/debit card transactions for the local GAStech employees spanning two weeks prior to the kidnapping. In this summary paper ,we show how we used applications of Geotools , QGIS and D3.js for the visualizations necessary to address the mini-challenge.

1 Introduction

In this year's VAST mini-challenge 2, participants were asked to analyze geospacial data and credit card data of the empoyees and extrapolate daily routines and unusual patterns among the employees which would furnish clues to the investigation of GASTech disappearences. In our approach we used ,
i)An adaptation of Geotools toolkit, an open source Java library for visualizing geospatial data thereby visualizing the paths in which the cars are moving.
ii)QGIS ,to locate and label the necessary placemarks that were mentioned in loyalty/credit card transactions .
iii)D3.js for building the data visualization frameworks that helped to depict the unusual patterns in data.
In the next sections, we describe in specific how these tools were used in the resolution of the VAST Challenge.

2 Daily Routines

2.1 Time/Date Specific Analysis

Working hours of the employees would contribute to the major part of the daily routines for any company .To determine the working hours , geospatial data was analysed at different hours of the day and density of cars was found. Interface for such a visualization was developed using application of geotools as shown in Fig1 .Working hours of the company can be assumed to be hours in which there was a higher density of cars near the office (shown by red color in the Fig 1).

The interface also gives us an idea of density of cars on various days of the week. The locations of higher density for weekdays can be considered as the places which are related to the company . This was drawn from the fact that these places were frequently visited by the employees during working hours. Thus, this information can contribute to daily routines of the employees.

2.2 Employee Specific Analysis

An interface smilar to that in Fig1 which helps in analysing data specific to a particular employee type is shown in Fig 2. Primarily the gps data was used to extract information specific to a particular employee type by observing recurring patterns .In addition to this , the loyalty/credit card data helps in extracting information specific to the particular employee and thus helps in finding employee specific daily routines.

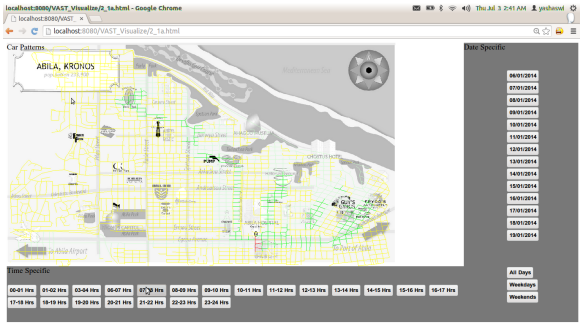

Figure 1: Interface showing car geospacial patterns wrt time/date(Geotools).

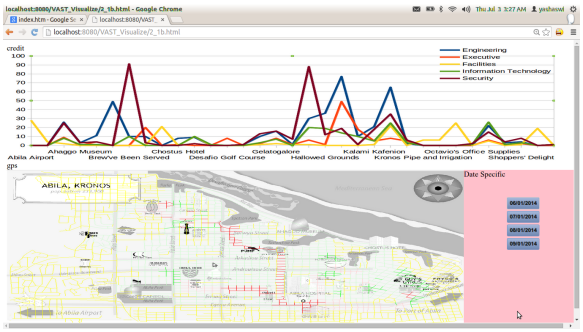

Figure 2: Interface showing Employee type specific car geospacial patterns and locations for transactions(Geotools) .

3 Unsual Patterns

3.1 Conflicting Data

When we tried to merge gps locations and the locations in credit/loyalty card transactions , we found a few instances where the same person was found to be present at two different locations at the same time (i.e. location in loyalty/credit card transaction was different from that in gps) . Such a situation can only occur when credit card ,car or both belonging to an employee were not being used by the employee but some other entity . Such instances can be considered as unusual patterns as they may be helpful to find the reason behind the missing employees.For visualizing this data we used qgis tool .One such instance in shown in Fig 3.

3.2 Peculiar instances

We found standard deviations of amounts considering transactions of every location separately .The locations having high standard deviations indicate that there were some unusual transactions (abnormally high or low amounts) at that location.Such recursive

*e-mail: p.yashaswi@research.iiit.ac.in
†e-mail:yarrabelly.navya@research.iiit.ac.in
‡e-mail:raghavendra.ch@research.iiit.ac.in

IEEE Symposium on Visual Analytics Science and Technology 2014
November 9-14, Paris, France
978-1-4799-6227-3/14/$31.00 ©2014 IEEE

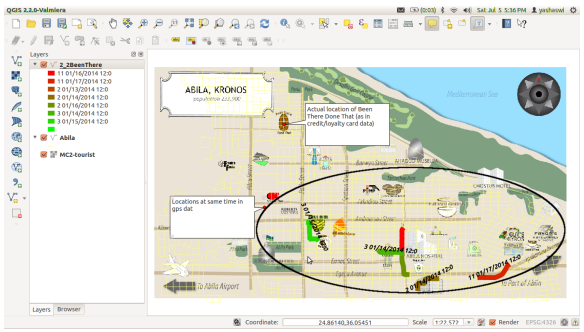

Figure 3: Conflicting Locations(QGIS).

and heirachical visual analysis was done using D3.js as shown in Fig 4.

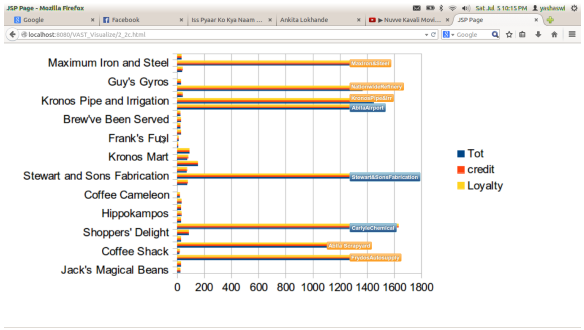

Figure 4: Graph showing standard deviation in amounts accross various locations(d3.js).

Loyalty cards are generally used by employees to get profits over high transaction amounts.So, some transactions where the loyalty card amounts were unsually less compared to credit card amounts can be considered as unsual patterns and such cases can end up being helpful for further investigation .These kind of instances were also visualized through D3.js as shown in Fig 5.

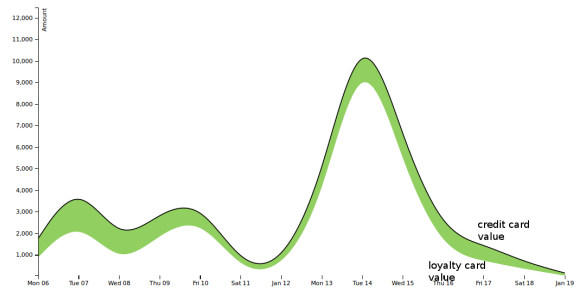

Figure 5: Difference in amounts between loyalty and credit card transactions(d3.js).

4 Conclusion

We bulit an interactive visual interface using geotools for geospatial data . The interface provides a cumulative outlook of daily routines of the employees which is easy to comprehend and also is user-friendly . The unusual patterns were extracted using qgis tool and D3.js framework. This choice was judicious as it offers the correct metrics to analyse data of this nature and the results derived were satisfactory.

References

[1] Geotools toolkit *http://www.geotools.org*.
[2] QGIS tool *https://www.qgis.org/en/site/forusers/download.html*.
[3] D3.js *http://d3js.org*.