Patterns of Life Visual Analytics Suite for analyzing GPS movement patterns

Simon Attfield[1], Peter Passmore[1], Neesha Kodagoda[1], Pragya Paudyal[1], David Neilson[1], George Pagiatakis[1], Brian Joyce[1], Robert Wells[1], William Wong[1], Adrian Wagstaff[2], James Bullock[2], Adam Malin[2], Dougie Holmes[2], Graham Phillips[2], John Marshall[2], Stewart Bertram[3]

Middlesex University, UK[1], MASS Consultants Ltd., UK[2], Intelligence Sans Frontieres, UK[3]

ABSTRACT

We developed the Patterns of Life (PoL) suite for the IEEE VAST 2014 Mini-Challenge 2. The suite comprises PoL Atlas for visualising movement and transactions spatiotemporally, PoL Classifier for classifying locations and identifying trips, and PoL Location Timeline, for visualising where time is spent.

Keywords: Visual Analytics, Information Visualisation, Spatiotemporal Analysis; Movement Data;

Index Terms: H.5.2 [User Interfaces]: Information Visualization;

1 INTRODUCTION

The IEEE VAST 2014 Mini-Challenge 2 (MC 2) [1] sets the task of discovering behaviour patterns of GASTech employees prior to the disappearance of some employees. The data provided includes two weeks of employee vehicle tracking data and credit and loyalty card transaction data. The task is to identify daily routines and unusual events using visual analytics and to deal with uncertainties and conflicts in the data. We describe the Patterns of Life (PoL) visual analytics suite and discuss some examples from the analysis.

2 PATTERNS OF LIFE SUITE

2.1 PoL Altas

We use PoL Altas (see figure 1) to visualise the vehicle tracking GPS data and transaction data. Altas has two views: an interactive map (top) and a timeline (bottom). To the left is a vehicle selector which controls the vehicles selected for display.

GPS data is plotted as points on the map. The timeline is a grid in which rows correspond to vehicles and columns correspond to time 'bins'. Credit and loyalty card data is overlaid on both. The views are coordinated [2] such that interaction in one (e.g. filtering, panning, zooming) cross filters what is seen in the other. Given a fixed number of time-bins (horizontally), zooming the timeline changes the bin time interval (and the resolution of the view).

Data are visually mapped between the map and the timeline using colour coding. There are three modes: *Geospatial* - points coloured according to x,y location on the map and bins coloured by average map location (mean map x and mean map y) within the bin; *Proximity to location* - events within a user defined spatial vicinity of a selected location are indicated; *Proximity to Subject* - events within user defined spatial and temporal vicinities of a selected subject are indicated. In figure 1 we show the interface in geospatial mode displaying some repeating truck movements in the data.

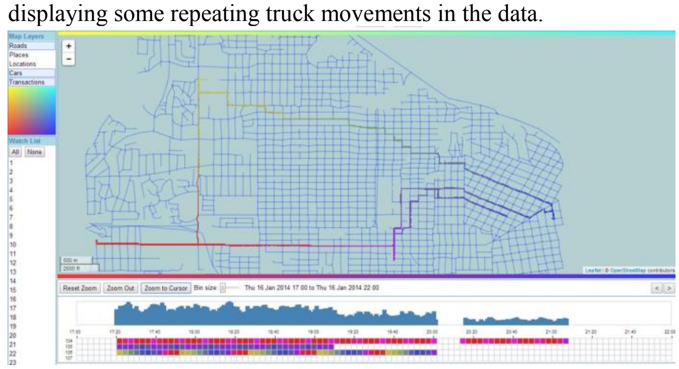

Figure 1: Showing truck movements on 16th, viewed using PoL Atlas. Repeating colour patterns on the time line show repeating journeys.

2.2 PoL Classifier

For further analysis, the GPS data was segmented into trips using PoL Classifier (a threshold of 1 minute to mark stops). With trips segmented, the Classifier tool can be used to visually plot successive individual trips to help the user infer and interactively label end locations (e.g. 'home') (see figure 2).

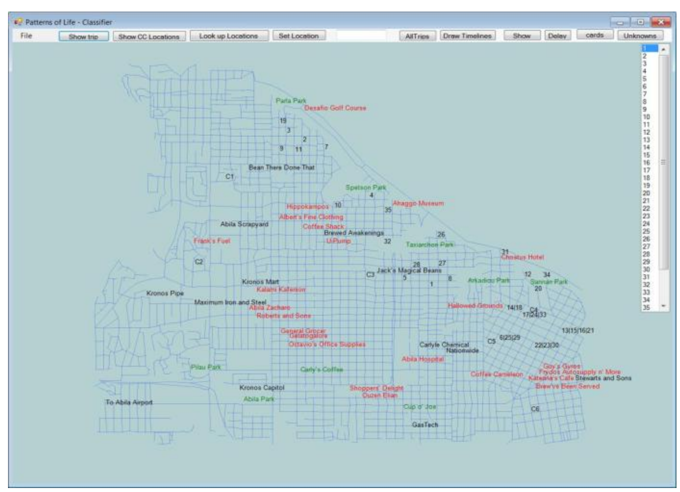

Figure 2: With inclusion of 6 new labels (C1-6) all trips and locations were classified. Locations derived from the tourist map were classified B4 (green), VAST supplied data B2 (red), and inferred data is C3 (black).

{s.attfield}{p.passmore}{n.kodagoda}{p.paudal}{d.neilson}{g.pagiatakis}{b.joyce}{w.wong}@mdx.ac.uk; {robertwells89@gmail.com}; {awagstaff}{jbulloch}{amalin}{dholmes}{gphillips}{jmarshall}@mass.co.uk; {berts123@hotmail.co.uk}

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

2.3 POL Location Timeline

With locations and trips classified, we use PoL Location Timeline to view time spent at named locations (see figure 3). Credit and loyalty card data is again overlaid. Vehicles are represented as horizontal bars and time read from left to right. Colour indicates a given location asTech premises) or location type (e.g. home, coffee shop).

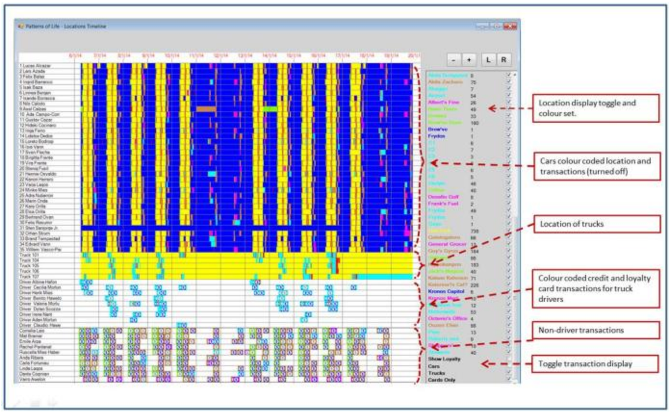

Figure 3: Using PoL Location Timeline to see where time was spent by GasTech employees over the two-week period of the data: yellow = *GAStech*, blue = *home*, light green = *coffee bar*, brown = *eatery*, magenta = *shop,* blue = *other locations*, red = *travel*.

3 Example of Use

We illustrate normal and abnormal behaviors using the PoL Suite.

3.1 Day of a life of an GASTech employee (normal)

Figure 3 shows that GAStech employees tend to start weekdays with a trip to a coffee bar, arriving at work between 7:30 and 8:30. They typically go out for lunch between 12:00 to 13:00, leave work for home between 17:00 and 18:00 and usually go out in the evening to shop or eat. They are usually home before 22:00. Employees mostly spent weekends at home, although some eat lunch out and perhaps shop. Most eat out in the evenings.

3.2 Night visits to executives' houses (abnormal)

Figure 4 shows security guards making early hours visits to executive's homes on four occasions. We noted this behaviour as unusual. Each executive is visited once.

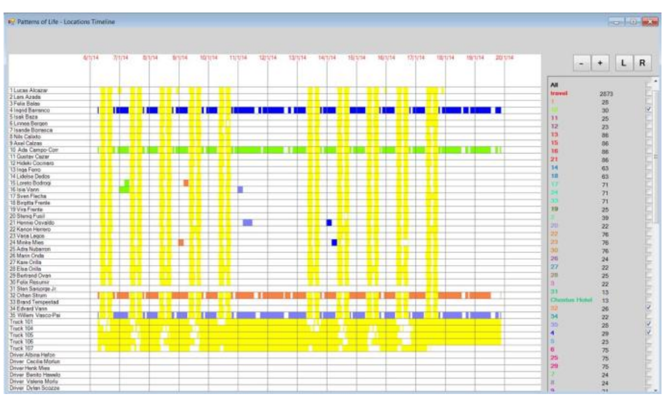

Figure 4: Using PoL Location Timeline to visualise night visits to executives homes by security guards.

Trucks behaving strangely (abnormal)

Figure 1 shows another example of what we classed as unusual activity visualised. Using PoL Altas, we see trucks repeatedly driving in circles or back and forth continuously, repeating routes without stopping. The routes are shown on the map whilst the repeating patterns are clearly visible on the timeline. Three trucks began this behaviour simultaneously on 16th Jan. One returned to GasTech (purple) before going out again and repeating its cycle.

4 Uncertainty in Data

Latitude and longitudes of locations were derived by: cross referencing credit card and GPS data (median coordinates at transaction time); inferring from GPS and possibly map data or inferring by the user or by credit card use; or inferring from map data (collocation of the tourist map with the Abila street map). Unknown locations visited by trucks were resolved by credit card use. We linked trucks to drivers based on location and credit card use. Certainty levels about locations were recorded using a confidence scale (table 1)

Table 1. Confidence scale.

Source evaluation		Intelligence evaluation	
A	Always reliable	1	Assumed true as provided by VAST
B	Mostly reliable	2	Assumed true but may contain errors
C	Sometimes reliable	3	Inferred by analysis team so may contain errors
D	Unreliable	4	Cannot be judged
E	Untested source	5	Suspected to be erroneous

5 Summary

The PoL suite was developed to assist in detecting patterns, both normal and abnormal, using geo-temporal data. PoL Atlas visualises geo-temporal data points and shows how patterns unfold over time. PoL classifier allows the user to segment the data into 'trips'. PoL Location Timeline displays the periods time spent at classified locations.

References

[1] K. Cook, G. Grinstein, and M. Whiting, "VAST Challenge 2014: Challenge Descriptions," VAST2014, 2014. [Online]. Available: http://vacommunity.org/VAST+Challenge+2014. [Accessed: 15-Aug-2014].

[2] J. C. Roberts, "State of the Art: Coordinated & Multiple Views in Exploratory Visualization," Fifth Int. Conf. Coord. Mult. Views Explor. Vis. (CMV 2007), no. Cmv, pp. 61–71, Jul. 2007. .