A Three Step Process to Design Visualisations for GeoTemporal Analysis

(VAST 2014 Mini Challenge 2)

Alvin Chua*, Ryo Sakai†, Jan Aerts‡, and Andrew Vande Moere§

Datavis Lab, University of Leuven, Belgium

1 Introduction

Given vehicle tracking data, loyalty and credit card logs of employees from a fictitious company, *GAStech*, participants of VAST 2014 mini challenge 2 were tasked to extract the common daily routine of employees and identify any suspicious activities that may be present in the data. In this paper, we reflect on our analysis procedure focusing on each step of the process that contributed to problem solving. Accordingly, we describe the features incorporated into our software at each stage of the process and justify the design decisions that were made. Inspired by DiBiase's approach to visual analysis [1], our procedure consists of three stages (Fig. 1). With off-the-shelf software, such as *R* and *QGIS*, we conducted exploratory data analysis [2] to generate a diverse range of insights. The insights were evaluated based on their relevance to the given task. They were used to formulate hypotheses and data task abstraction [3] resulting in a set of complementary tools comprising of an origin-destination map, a timeline and a flow diagram that we developed in *processing*. These tools were not designed to function as an integrated software package but were treated as rapid prototypes that would afford more flexibility in the design and development cycle [4].

2 Exploratory Analysis

Our goal at this stage was to extract summary statistics and familiarize ourselves with the data so as to identify inconsistencies such as missing or conflicting information. To facilitate this process, we scripted in *R* to quickly visualise with *ggplot2*. The focus at this stage was to minimise the time spent on software development. Subsequently, parallel coordinates were used as a brute force technique to thoroughly check for any patterns that may have been overlooked. In addition, *QGIS* was used for interactive spatial analysis. The process resulted in the following tasks to be addressed by our purpose built visualisation:

- Aggregate movement information to reveal routine and spontaneous commutes by employment group.
- Reveal where employees gather and provide reference to the nature of meeting.
- Merge credit card and loyalty card transactions and reveal the employee's location on the map at the time of transaction.

3 Designing Tools for Confirmatory Analysis

Our design comprises of three components; The OD map is an

*E-mail: alvin.chua@asro.kuleuven.be

†E-mail: ryo.sakai@esat.kuleuven.be

‡E-mail: jan.aerts@esat.kuleuven.be

§E-mail: andrew.vandemoere@asro.kuleuven.be

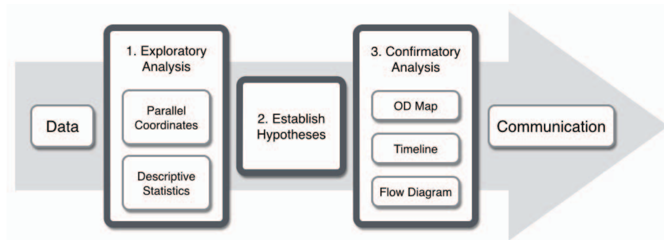

Fig. 1. The overview of our analysis procedure.

abstract representation of the employee's movement as a graph. Conversely, the timeline offers a temporal representation of the employee's movement as well as reference to their credit card transactions. The flow diagram optimises the information shown on the timeline by emphasizing moments in time when the employees are spatially collocated. All three components were linked to allow spatial temporal filtering.

OD Map expresses the daily commutes by employees as a weighted directed graph. The edges are weighted by the frequency of travel between two locations. The user may interactively filter the graph through an edge weight histogram to display only a subset (i.e. routine/spontaneous) of commutes made by the employees. The visual encoding we adopted was informed by the heuristics described by van Wijk [5].

Timeline represents all of the employee's activities in chronological order over the course of a day. Horizontal line segments depict the duration of time when an employee is stationary while vertical lines are used to mark the point in time where credit transactions were made. The timeline provides an interface that integrates information related to credit card transaction to an employee's location on the map. The user identifies possible gatherings by searching for line segments that overlap or align. Collocation on the map can then validate possible gatherings. In such instances, information from credit card transactions is useful for understanding the nature of the activity undertaken by the employees.

Flow Diagram is an optimised representation of the timeline specifically developed to emphasise the spatial collocation of employees. Splines are used to express the movement of employees across the passage of time across the horizontal axis. Each spline represents an employee colour coded by employment group. Splines are bundled when employees are spatially collocated and the length of each bundle reflects the duration of each "gather". Rectangles are drawn over these bundles to improve visual saliency. The use of bundling in the Flow Diagram arose as a response to the need for vertical sorting of the collocated employees in Timeline. While Timeline assists in identifying alignments and overlaps on the timeline, it results in an arbitrary ordering of employees and the vertical axis. Furthermore, sorting the employees along the vertical axis provides reference to only one event. The Flow Diagram was able to overcome these limitations with spline bundling.

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

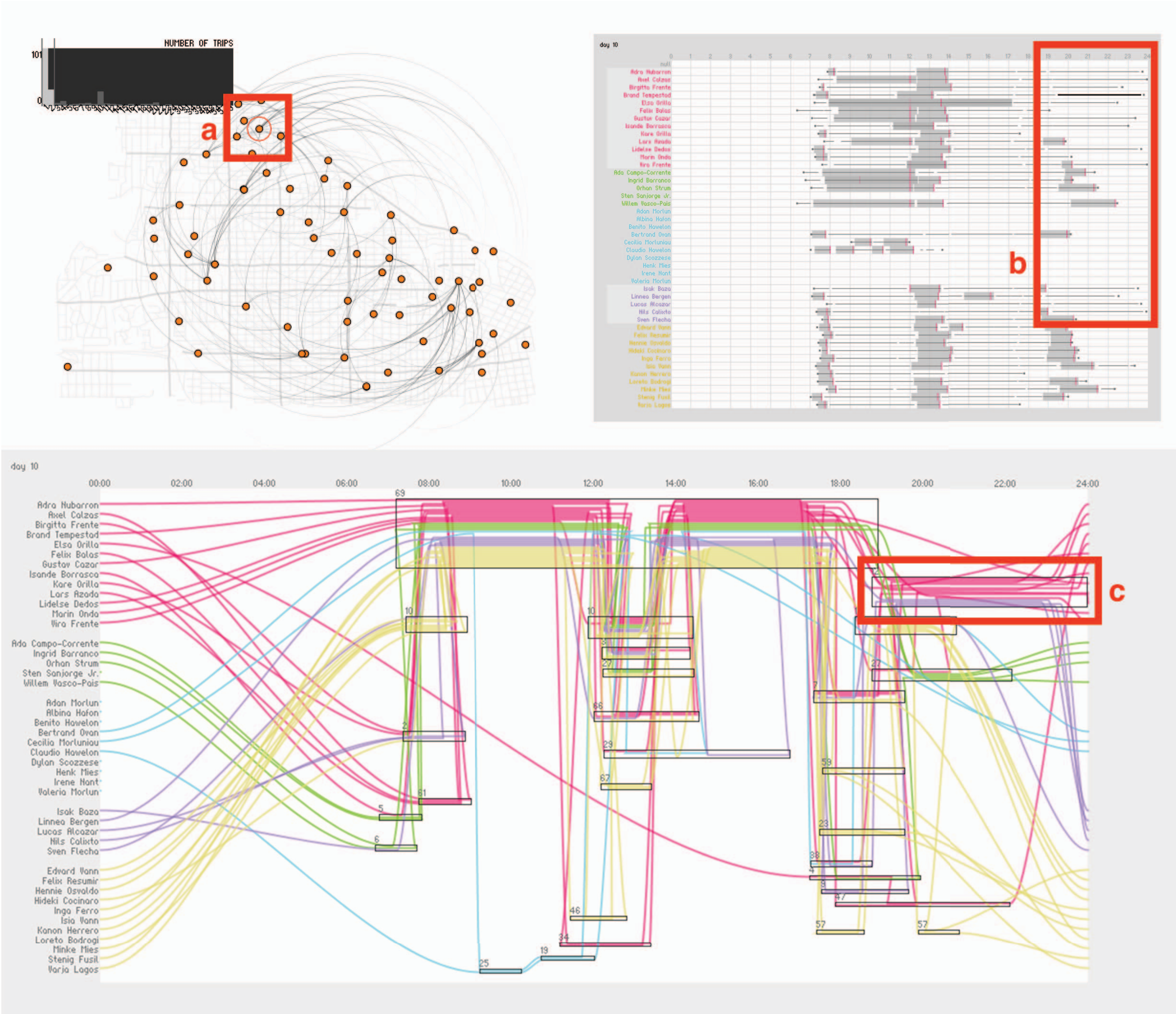

Fig. 2. An overview of the purpose built visualisation. It consists of three linked components: (a) An OD map, (b) a timeline and (c) a flow diagram. The various components each offer a different representation of the data. In this instance, the regions marked by the red rectangles reveal a large gathering of employees.

4 Use Case

We describe a use case involving our bespoke tools that led to the discovery of a large gathering of employees. The user may trace the movement employees over the OD map by brushing through line segments on the timeline. The location of an employee is then reflected on the map. The user may make comparisons between the commuting behaviour of different employment groups through the OD map. Fig. 2b shows the collocation of 14 employees from engineering and IT on the timeline and Fig. 2c shows the same collocation on the Flow Diagram. Cross-referencing the same pattern on the OD map in Fig. 2a verifies this insight.

5 Conclusion

We created a set of three complementary tools based and successfully used them to explore the VAST Mini Challenge 2 Dataset. For detailed descriptions of our findings for the challenge, see http://goo.gl/THj4C2. The design of these tools was guided by the insights generated through our analytical procedure. By organising our workflow into three distinct phases, we obtained a diverse range of insights, which in turn enabled us to approach the problem in a systematic manner. As a result, we were able to design tools with very specific features while maintaining flexibility in development.

References

[1] D. DiBiase, “Visualization in the Earth Sciences,” *Earth And Mineral Sciences*, vol. 59, no. 2, pp. 13–18, 1990.

[2] J. T. Behrens, “Principles and procedures of exploratory data analysis.” *Psychological Methods*, vol. 2, no. 2, p. 131, 1997.

[3] T. Munzner, “A nested model for visualization design and validation.” *IEEE Transactions on Visualization and Computer Graphics*, vol. 15, no. 6, pp. 921–928, 2009.

[4] J. Dykes, A. M. MacEachren, and M. J. Kraak. "Facilitating interaction for geovisualization." Exploring geovisualization, pp. 265-291, 2005.

[5] D. Holten and J. J. van Wijk, “A user study on visualizing directed edges in graphs,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2299–2308, 2009.