Short Paper

LI Yuhong
Singapore Management University
yuhong.li.2020@mitb.smu.edu.sg

XU Dan
Singapore Management University
da.xu.2020@mitb.smu.edu.sg

YU Yaxin
Singapore Management University
yaxin.yu.2020@mitb.smu.edu.sg

ABSTRACT

Interactive applications developed using Shiny for the R language can satisfy move disciplines interested in movement which are visualization and visual analytics (Ware C. 2013). Shiny allows customizing the user interface of the application, providing an elegant environment for displaying user input controls and analog output-the latter is updated at the same time as the input changes. The flexible nature of R language makes it possible to freely choose and observe the movement trajectories of different objects, thus promoting the combination of Shiny and R in model visualization.

1.   1. INTRODUCTION

Movement data in space and time provide critical information about the behaviour of objects studied, i.e., vehicles, humans, and animals. Insights from spatial and temporal movements drive many human and ecological studies. These studies enable in-depth analysis and understanding of complex processes like an animal and human migration, forest fires and natural disasters as well as mobility and transportation.

](Picture1.jpg){width=50%}

For most people, it is really hard to explore geospatial data without using a map. With this application, we demonstrate how to create a free and open-source solution using R and R Shiny, that allows users to interact with the geospatial data, and quickly find what they want to see using the interactive Shiny app.

2.   2. MOTIVATION OF THE APPLICATION

Recently, a number of studies in computer science focused on identification of human movement behaviour patterns from GPS trajectories. These studies look at identification of significant places (i.e., locations which play an important role in the activities of a user) (Ye et al., 2009), classification of human behaviour in these places (Huang et al., 2013) and analysis of spatial interactions between significant places identified from trajectories (Sila-Nowicka et al, 2014). In addition to this, new tracking technologies for observation of human movement have been deployed as alternatives to GPS technology.

A common design approach for visual analytics systems for highly-dimensional complex data is to use linked views, that is a set of interactively connected visualisations, each of which provides a different perspective on the data (Roberts, 2007). This methodology takes a set of data displays (each showing a selection of the given dimensions in some particular way) and then allows the interaction in one view (e.g., selection, zoom in, zoom out, etc.) to simultaneously modify displays in all views. This is relevant to visual exploration of spatio-temporal data, since it enables generating a unique spatial, temporal or spatio-temporal perspective on a pattern appearing in a linked attribute-only view (Andrienko, 2010).

Trajectory data, defined here as a discrete time series of measured locations, are collected at detailed temporal resolutions and on particular temporal schedules. Global Positioning System (GPS) devices are now in everything from watches to cars and, of course, smartphones. This means that GPS datasets have the ability to track a large proportion of the world's population. Depending on sampling frequency and schedules, such data volume can be very large (long, densely sampled trajectories).

Additional complexity is introduced with simultaneous collection of related information either directly from other sensors or derived from environmental data (Cagnacci et al., 2010). Although there are privacy concerns, when appropriately anonymized and aggregated, GPS datasets have the potential to help tackle the issues raised in human movement events.

3.   3. REVIEW AND CRITIC ON PAST WORKS

Identifying a representative route from a set of trajectories is a common topic in GIScience and there are many approaches outside move. Brudson (2007) uses principal curves to identify the most probable route from a set of GPS pedestrian trajectories. Similar methods are widely used in navigation and even developed for reconstruction of representative 3-D trajectories: reconstruct 3-D bicycle tracks from GPS trajectories – a method that could be of interest for movement of animals freely moving in 3-D (birds, sea mammals).

In move, Etienne et al. developed a method to identify the main naval route from a set of vessel trajectories, sampled at equal times. Pelekis et al. take an alternative approach and consider the uncertainty in trajectory measurements by constructing a fuzzy vector representation of each trajectory. They use this representation to construct a so-called centroid trajectory as the representative path based on density of trajectory points at each moment in time.

Many move studies mentioned to employ the linked views paradigm. Others include not only spatio-temporal displays, but also attribute visualisations. Tominski et al. (2012) introduce a trajectory wall. Here, car trajectories are represented as the 2-D plane of the geographic space. Ribbons are stacked over their geographic path and their segments coloured according to an attribute (speed). This view is interactively linked to a time lens, where clusters of temporally similar trajectories are shown using edge-bundling.

The system allows identification of temporal patterns in car movement and in particular anomalies in regular flow, such as traffic jams. Andrienko et al. present linked views for a comprehensive visual exploration of any type of movement trajectories. These systems include a number of attribute visualisations, density maps, temporal visualisations and a number of other displays that allow incorporation of contextual information.

4. 4. DESIGN FRAMEWORK

Our application used the free and open-source R language which offers a thriving programming environment for statistical and graphical analysis. We load all the required libraries to the app. The application interface is usually created using fluidPage which ensures that the page is laid out dynamically. Our considerations for creation are listed following: ● Explanatory Data Analysis for Background ● Using standard R packages on the CRAN for supportability. ● Employing Shiny App to webify the codes for simplicity. ● Offering options and visualizations for interactivity.

4.1 4.1 Data Preparation

The dataset used in this project was same as the Vast Challenge 2021 Mini Challenge 2. The dataset provided by GASTech company involves 4 attributes: timestamp, car-id, car-type, and gate-name. The raw data was provided in csv format. First step for data preparation was to conduct in R using base R and the dplyr package to narrow the scope of the data by removing unused data and joining data into correct format. Some variables were grouped logically in the base of common definitions. The further preparation was made according to the real requirements of the different analysis methods applied.

4.2 4.2 Background Analysis

Our application allows interactive bivariate frequency analysis to be performed under the Background Analysis tab.

Popular Locations In Fig. 2, the bar chart shows the number of transactions by employees of the GASTech company in different locations to reflect the popularity of different consumption locations. This bar chart also allows users to select the type of card to view the transaction status of different types of cards in different locations.

Figure 2: Transaction frequency for different locations](Picture2.png)

It is noted that Katerina's Caf has the largest traffic volume of both credit cards and loyalty cards. About one-third of the locations have more than 50 transactions over the investigated 2 weeks.

Hotpoints of News Articles

The Hotpoints of News Articles sector gives the interactive visualization to view the top n most used words within each news group. The word cloud diagram can visually display the words that appear more frequently in different news. The user can select the news name he is interested in through the drop-down box on the left side of Figure 3 to observe. At the same time, our application also adds Minimum Frequency and Maximum Number of Words these two horizontal setting boxes. For example, the most frequent words in News Desk include Kronos, government, gastech, etc.

Figure 3: Wordcloud of most used words](Picture3.png)

Figure 4: Bar chart of top 20 most used words in each news](Picture4.png)

The bar chart in Figure 4 shows the top 20 most frequently occurring words in News Desk. If the user selects a different news group, Figure 3 and Figure 4 will change accordingly at the same time.

4.3 4.3 Consumption Analysis
Figure 5

Figure 5](Picture5.png)

4.4 4.4 Path Analysis
The GPS trajectories of the geographic information associated with the GASTech data is defined as locations as shown in Figure 6.

Figure 6: GPS trajectories for each employee](Picture6.png)

The geographic information associated with the GASTech data is defined as zones. There are in total 34 locations in the investigated region and every location can be either an origin location or a destination location. In other words, there are 34×34 possible combinations of the origin-destination location pairs. Users can explore the GASTech employee's movement data between any origin and destination location their daily moving. This tool also offers an intuitive way to select the day of the week (Figure 7) directly from the list where users simultaneously get the information of location, time, and intermediate stops. Users can choose what data they want to see by selecting Car ID of employee from the drop-down list pane. They can zoom in and out or pan the map to find the location and employee's movement of interest.

Figure 7: Day of the week selection](Picture7.png){width=50%}

5. 5. DISCUSSION
Our analysis of the temporal sequences of network motifs in individual movement networks revealed successions of spa-

tial patterns corresponding to changes in behavioural modes that can be attributed to specific spatio-temporal events of each employee trajectory. Future applications of our method to multi-layered movement and social network analysis yield considerable promises for extending the study of complex movement patterns at the population level.

6. 6. FUTURE WORK

The application was developed using Shiny on R, supported with a range of geo-spatial packages to provide users with a suite of techniques to derive insights from the data. Given the emphasis on visual analytics in this project, the various components of the application were developed with due consideration of the taxonomy of interactive visualizations. We suggest that our further development of the app draw the idea of adding additional user interaction features within each view in the app. The users could be given the option to highlight certain sets of interest in the frequency bar chart which can speed up the interpretation of findings. As another example, the bar chart could incorporate an option to compare the current view with the finding for a different news group (i.e., adding a reference line) to make comparisons easier for users. Also, future work can also address the two trajectories with a potential encounter computed based on the Brownian bridge movement model to explore their connection or similar trajectories.

References

[1]   Fenner, M. 2012. One-click science marketing. Nature Materials. 11, 4 (Mar. 2012), 261–263.

[2]   Meier, R. 2012. Professinal Android 4 Application Development. John Wiley & Sons, Inc.