VizTrails: An Information Visualization Tool for Exploring Geographic Movement Trajectories

Martin Becker
University of Würzburg
becker@informatik.uni-wuerzburg.de

Philipp Singer
GESIS - Leibniz Institute for the Social Sciences
Cologne, Germany
philipp.singer@gesis.org

Florian Lemmerich
GESIS - Leibniz Institute for the Social Sciences
Cologne, Germany
florian.lemmerich@gesis.org

Andreas Hotho
University of Würzburg and L3S Research Center
hotho@informatik.uni-wuerzburg.de

Denis Helic
Graz University of Technology
dhelic@tugraz.at

Markus Strohmaier
GESIS - Leibniz Institute for the Social Sciences and University of Koblenz-Landau
markus.strohmaier@gesis.org

ABSTRACT

Understanding the way people move through urban areas represents an important problem that has implications for a range of societal challenges such as city planning, public transportation, or crime analysis. In this paper, we present an interactive visualization tool called VizTrails for exploring and understanding such human movement. It features visualizations that show aggregated statistics of trails for geographic areas that correspond to grid cells on a map, e.g., on the number of users passing through or on cells commonly visited next. Amongst other features, VizTrails allows to overlay the map with the results of SPARQL queries in order to relate the observed trajectory statistics with its geo-spatial context, e.g., considering a city's points of interest. The systems functionality is demonstrated using trajectory examples extracted from the social photo sharing platform Flickr. Overall, VizTrails facilitates deeper insights into geo-spatial trajectory data by enabling interactive exploration of aggregated statistics and providing geo-spatial context.

Categories and Subject Descriptors: H.5.3 [**Information Interfaces and Presentation**]: Group and Organization Interfaces—*Web-based interaction*

Keywords: Human Trails; Flickr; Hypotheses; HypTrails

1. INTRODUCTION

Understanding the way people move through urban areas represents an important problem that has implications for a range of societal challenges such as city planning, public transport or crime. Recent research has studied human movement trajectories in cities through a variety of data sources including mobile phone data, GPS and WiFi tracking, location-based social media platforms, online photo sharing sites, and others (cf. related work section of [1]).

Recently, we have extended this line of research by studying the underlying processes of a set of trails derived from human movement exemplified in urban photo trails [1]. In that direction, the HypTrails approach [2] allows to formulate and compare different hypotheses about the production of these trails. In order to better understand how the original photo trails materialize, and also to gain further insights on how our own hypotheses explain these trails, we have implemented a visualization tool called VizTrails[1]. In this article, we present VizTrails including an overview of its architecture (Section 2) and visualization aspects (Section 3).

VizTrails shows aggregated information for grid cells on a map featuring interactive visualization of statistics, such as the number of users passing through cells, the in- and out-degree from and to other cells or the cells commonly visited next. Amongst other tools, VizTrails enables overlaying the map with content from arbitrary SPARQL queries for relating the observed trajectory statistics with geo-spatial context. VizTrails is designed for minimizing the required pre-processing steps. Overall, VizTrails facilitates deeper insights into geo-spatial trajectory data by enabling interactive exploration of aggregated statistics and providing geo-spatial context.

2. ARCHITECTURE

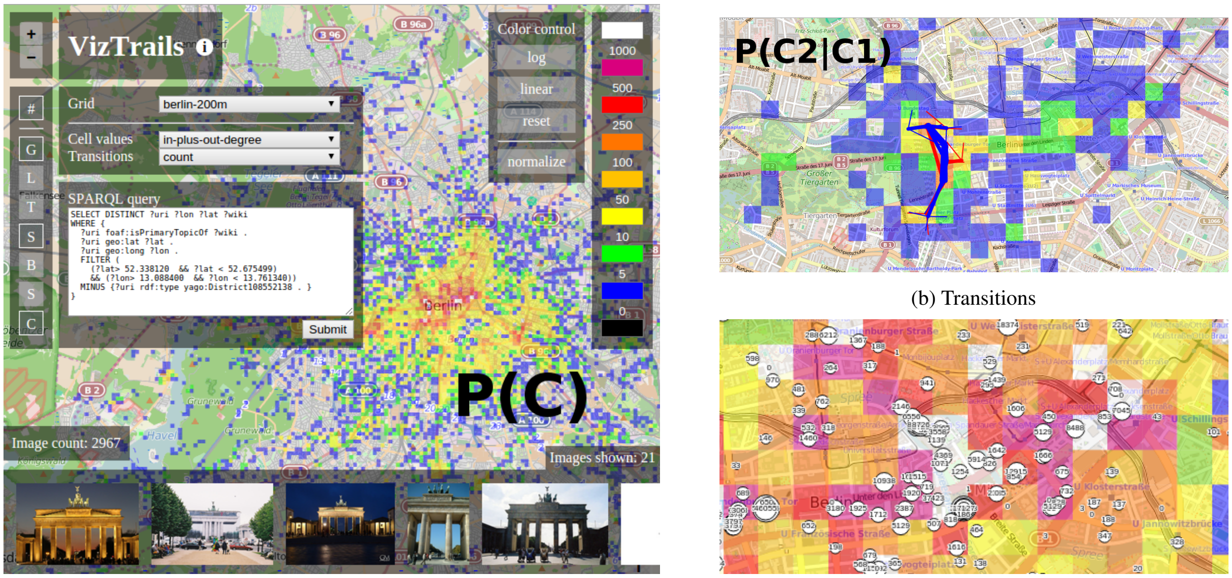
VizTrails is a web application based visualization system. It consists of two independent layers: the REST-layer for serving data and the UI-layer for visualizing the provided data.

The REST-layer is connected to a database and provides endpoints for accessing data points, user trajectories, grid cells, cell transitions, and more. It is built to be modular, i.e., the underlying database is easily exchangeable. Thus, it can not only serve data from relational databases like MySQL or PostgreSQL, but can also directly access data from distributed NoSQL databases like HBase or Cassandra. This is especially useful when large amounts of trajectory data are processed via parallel computation frameworks like Hadoop or Spark which directly write to such distributed data storage systems.

The UI-layer is browser-based. It pulls the data from the endpoints provided by the REST-layer and visualizes it via HTML, JavaScript and corresponding frameworks like jQuery or OpenLayers. As a primary goal of VizTrails, the UI-layer enables data exploration in real-time. Since the listing of available grids and transitions is directly coupled with the REST-layer, new grid and transition types are immediately available in the user interface. This allows for a smooth workflow from generating and analyzing data towards visualizing it.

[1] `http://dmir.org/viztrails`

HT '15, September 1–4, 2015, Guzelyurt, Northern Cyprus.
ACM 978-1-4503-3395-5/15/09.
http://dx.doi.org/10.1145/2700171.2791021



(a) Grid View

(b) Transitions

(c) POIs

Figure 1: *Visualization components (map tiles from OpenStreetMap).* In (a) we show the general grid view visualizing different values for individual grid cells providing a general overview of some global statistics. In this case, photo counts (e.g., $P(C)$) in Berlin are depicted. The second subfigure (b) demonstrates how transitions from or to a cell are visualized when clicking on that particular cell (e.g., $P(C2|C1)$). This allows to explore how people move from or to different places. Third, (c) shows how entities from DBPedia and their respective view counts on Wikipedia are visualized providing trajectories with spatio-semantic context. These different visualization modes aid in exploring data about human movement trajectories in an intuitive and explorative way.

3. VISUALIZATIONS

We visualize geo-spatial trajectory data by discretizing an area defined by a bounding box into grid cells as depicted in Figure 1a. Trajectories are then projected onto this grid. This allows us to visualize aggregated statistics on the set of all trails that contain a location within this grid cell. These include single cell statistics, cell transitions, and the respective geo-spatial context. In the following, we describe these visualizations in the same order.

Cell frequencies. For an overview of the general spatial distribution of the recorded datapoints, we color each grid cell according to the number of data points in that cell. The color as well as the value intervals associated with each color can be freely chosen. In addition to the number of data points in each cell, this visualization can be used to visualize any other scalar valued statistics depending on the values the grid provides (in our case we also provide in- and out-degree for each cell). A dialog allows to choose from a number of different grids and associated values and updates as new grids are available in the database. Upon choosing a grid the map automatically pans and zooms to the appropriate extent.

Markov chain transitions. Now, in order to explore trajectories, the UI allows to visualize first-order Markov chain transitions. When clicking on a cell, cell colors change from a coloring based on overall statistics, to colors associated with the count of transitions starting at a point within the clicked cell. We also show lines for the most probable trails from (red) or to (blue) that cell. Thus, for example in the Flickr case, it can easily be judged where people will go from the current cell in order to take their next picture. Figure 1b shows the transitions from the “Brandenburg Gate” in Berlin. Here people mostly move towards three destinations, namely the “Reichstag” building, the “Potsdamer Platz” and the “Museum Island” (marked in orange on the screenshot). Note, that this feature not only allows to visualize actual trajectories, but can also be used to contrast them with hypotheses about transitions as applied in [1, 2].

Spatio-semantic context. In previous work [1], we have found that the processes resulting in human trajectories are strongly connected with geo-spatial features such as points of interest and their corresponding popularity in the social and semantic web. In order to be able to directly correlate trajectories with such features, we provide the possibility to query and visualize geo-spatial entities from DBPedia via SPARQL. In addition, these entities can be weighted by the view counts of the respective Wikipedia articles[2] (if available), as shown in the example screenshot show in Figure 1c.

Flickr. Although VizTrails can visualize arbitrary geo-spatial trails, our demonstration example features urban *photo* trails from the Flickr platform. As an additional feature for this dataset, we can also search for particular photo ids or show public photos that have been taken within a bounding box drawn on the map, cf. Figure 1a.

4. CONCLUSION

In this paper, we have introduced an interactive visualization tool called VizTrails that allows exploring human movement and corresponding trails. To this end, we have used a grid-based approach to visualize a number of metrics as well as mutual transitions between cells. VizTrails also allows to set these trails into geo-spatial context using semantic web data via SPARQL queries. Thus, it enables interactive exploration and also facilitates deeper insights into spatial trajectory data.

Acknowledgements. This work was partially funded by the DFG in the research project "PoSTs II".

References

[1] M. Becker, P. Singer, F. Lemmerich, A. Hotho, D. Helic, and M. Strohmaier. Photowalking the city: Comparing hypotheses about urban photo trails on flickr. 2015. under review `http://dmir.org/pub/2015/photowalking.pdf`.

[2] P. Singer, D. Helic, A. Hotho, and M. Strohmaier. Hyptrails: A bayesian approach for comparing hypotheses about human trails on the web. In *International Conference on World Wide Web*, 2015.

[2]extracted from `http://dumps.wikimedia.org/other`