An Interactive Dashboard for Traveler Mobility Analysis

Lukas Vorwerk
Department of Informatics
Technical University of Munich
Garching, Germany
lukas.vorwerk@tum.de

Linus W. Dietz
Department of Informatics
Technical University of Munich
Garching, Germany
linus.dietz@tum.de

ABSTRACT

Showcasing research in data mining is a challenging topic. Developed models of everyday phenomena are often only understandable by domain experts which might lead to low general adoption. We demonstrate an interactive dashboard to visualize the complex domain of international travel over time. Given that spatio-temporal phenomena such as mobility can not effectively visualized using consumer software, we developed a web-based system, where users can explore how the behavior of international and domestic travelers changes over time and easily create their own analyses.

KEYWORDS

Tourist mobility, Data mining, Visualization, Dashboards

1 INTRODUCTION AND RELATED WORK

Human mobility analysis is a widely researched topic with many facets such as next location prediction [10], discovering recurring activity patterns [7], determining different traveler types [2], or deriving touristic travel regions [3]. An important problem of all research efforts concerning human mobility is the visualization of the results. To establish an understanding of the resulting models, visualizations are oftentimes very selectively included into the manuscripts as plots. This, however, only shows a part of the findings and it is hard for practitioners in the tourism industry to utilize the academic findings to improve their services. To improve upon this, we have developed an interactive traveler mobility dashboard that visualizes patterns of global mobility. The implementation as a web application offers users the flexibility to create their very own analyses.

The underlying data stems from a self-collected data set from Twitter. Due to that most content of Twitter is available for the general public, it has been frequently used as a source for researching human mobility [5, 6]. We employed our earlier approaches to mine trips from geotagged tweets [1, 3] and aggregate the metrics into an interactive dashboard available at http://mobility-dashboard.cm.in.tum.de.

Dashboards provide users with a customizable overview about the state of a system without requiring them to work with the raw data and using programming techniques. Thus, they are an ideal tool for the general public to explore complex topics such as stock markets, misinformation on social media[1] or global pandemics[2].

In this position paper, we present our initial prototype for a traveler mobility dashboard based on mobility data derived from social media. We present the motivations and design considerations to build a dashboard that can visualize analyses concerning trends in global traveling behavior and point out relevant use cases and applications.

2 TWITTER MOBILITY DATA SET

Mobility data about individuals is valuable, since it can be used for many purposes and it is quite complicated to collect it on a larger scale. Most data sets are limited to national boundaries, since traditional methods of collection, such as mobile phone communication records [4] or national statistics are tied to the respective administrative regions. With the increasing adoption of online location-based social networks (LBSNs), mobility traces of humans have become more widely available and this data has the advantage that it is typically not impeded by traveling to foreign countries. Twitter is especially useful for researchers, since the service is inherently public and it provides official APIs to query user timelines.

Whenever a user tweets with the geolocation option enabled, her approximate location is attached to the message. The downside of LBSN mobility data is that it gives an incomplete picture of the user's mobility — the location is only known when she decides to tweet. This makes it necessary to assess the quality of the data with respect to the use case, eventually, discarding trips of insufficient data quality.

Table 1: Scope of the underlying Twitter mobility data set (as of January 27, 2021)

Feature	Value
Tweets with geolocation	16,551,608
Users	302,747
International Travelers	128,539
Domestic Travelers	174,208
Trips	1,788,432
International Trips	395,963
Domestic Trips	1,392,469
Countries visited	241
Observation period	2015 – 2021

Since 2018, we collected timelines of Twitter users and using the Python `tripmining` library[3], we segment geolocated tweets of a user's timeline into periods of being at home and on travel. If a user tweets at a relatively constant pace, this data can be quite reliable, which is also one of the of the library's features to assess using various metrics.

Table 1 summarizes the statistics of our Twitter trips data set used for the dashboard. As the data collection is ongoing, these

[1] http://csmr.umich.edu/projects/iffy-quotient/ [8]
[2] https://coronavirus.jhu.edu/map.html
[3] https://github.com/LinusDietz/tripmining

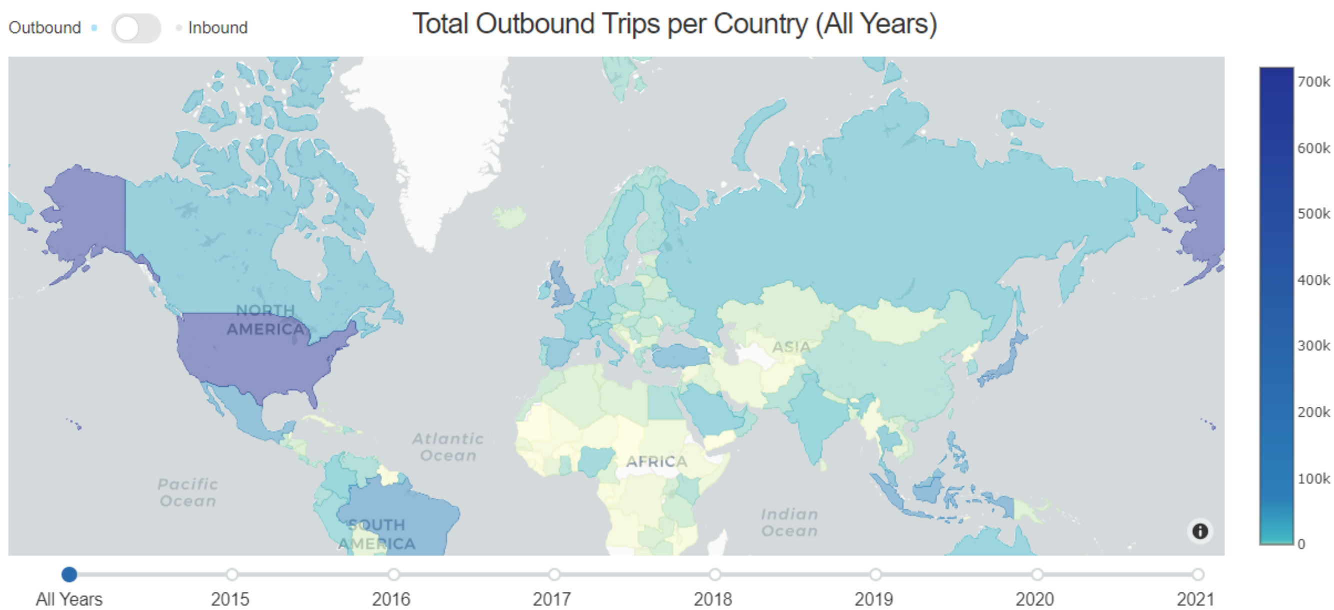

Figure 1: Screenshot of the Mobility Dashboard. See http://mobility-dashboard.cm.in.tum.de

numbers are constantly growing. While only parts of the data set are currently publicly available[4], we are able to provide mobility data to interested researchers on a trip-level.

3 MOBILITY DASHBOARD

The dashboard consists of three interactive visualizations that invite users to engage with the data and create mobility analyses on their own. The central feature is a choropleth world map that reflects the number of inbound or outbound trips on a country-level basis. The exact number is shown when hovering over the respective area on the map. By using a toggle switch, the user can decide whether to display inbound or outbound trips. Upon clicking on the area of a specific country, users are presented with a detailed analysis of this country. They can now determine where the residents of this country prefer to travel to and from which other countries travelers visit this country in which quantities. Figure 1 shows the number of outbound trips in our database undertaken since 2015.

Besides the map, a list containing the top ten countries for the selected scenario is displayed. An additional visualization in the form of a line chart displays the mean and median trip duration per year or month. To provide insights into the data the visualizations are based on, the dashboard also includes the statistics shown in Table 1 and a chart that indicates the number of trips per year or month (see Figure 2). Furthermore, a time slider offers the possibility to uncover temporal trends over the last decade. All of the described dashboard elements are interconnected, i.e., if the user selects a specific country or year, all visualizations are updated accordingly.

The dashboard is developed using Dash[5], a "Python framework for building web analytic applications". The usage of Dash enables further interactions with the visualizations such as zooming in on specific chart areas. To the best of our knowledge, this platform is the first offering an interactive analysis of global travel trends.

[4]https://github.com/LinusDietz/JITT2020-Mining-Trips-Replication [3]
[5]https://dash.plotly.com/introduction

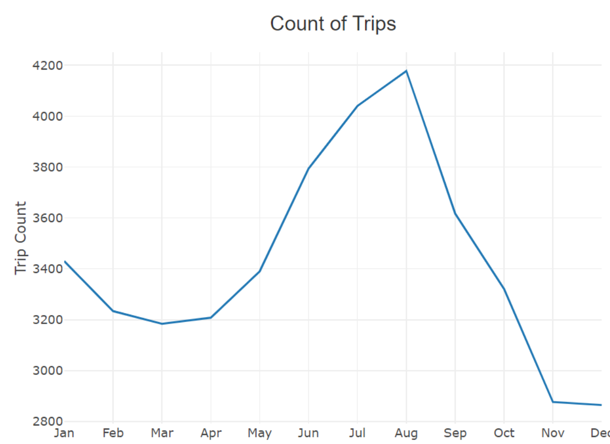

Figure 2: Inbound Trips to the United Kingdom in 2019.

4 LIMITATIONS AND USE CASES

Given the intrinsic properties of the data, some analyses work better than others using the dashboard. What works quite well is to get an indication of the popularity of a destination country as well as the popular travel seasons aggregated on a monthly basis. Certainly, the raw numbers have little meaning, since they only account for the mobility of Twitter users. The relative popularity of countries in the Western World, however, should be meaningful. On the contrary, the numbers for countries where Twitter is not much adopted, e.g., due to state censorship, are certainly not representative. Furthermore, the overall amount of data might be enormous (cf. Table 1), but when subdividing it into 7 years (currently about 73 months), 241 countries, and inbound and outbound trips, the amount of data points can become quite low, especially for small countries. Finally, it should be noted that the data collection is slower than its creation, and it takes time until enough recent trips are collected. For this reason, the most recent data is the most incomplete, since trips that are currently occurring can not yet be included in the data set.

Therefore, the target audience of the dashboard is rather the general public than researchers. The number of possible analyses is intentionally limited at the moment, since we believe that increasing complexity would hinder layman to interact with the dashboard to produce their own simple, but useful analyses.

5 CONCLUSIONS AND FUTURE WORK

This paper described the interactive visualization of international traveler mobility from Twitter. The underlying data has been collected over the last year and shows the high value of LBSN data for such analyses. With the presented dashboard, we provide a tool that can be used to perform own analyses on the complex domain of international travel. Currently, the dashboard allows to explore the temporal variability of the number of inbound and outbound trips of Twitter users.

In the future, we plan to improve the granularity of the analysis into smaller regions, such as federal states or touristic regions [9]. Furthermore, it should be possible to set the timeframe to custom dates, to be able to observe seasonal trends e.g., introduced by school breaks or holidays associated with traveling, such as Thanksgiving in America. Finally, we plan to develop further analyses, e.g., to reveal which countries have often been traveled to together.

REFERENCES

[1] Linus W. Dietz, Daniel Herzog, and Wolfgang Wörndl. 2018. Deriving Tourist Mobility Patterns from Check-in Data. In *WSDM Workshop on Learning from User Interactions*. Los Angeles, CA, USA.

[2] Linus W. Dietz, Rinita Roy, and Wolfgang Wörndl. 2018. Characterisation of Traveller Types Using Check-in Data from Location-Based Social Networks. In *Information and Communication Technologies in Tourism*, Juho Pesonen and Julia Neidhardt (Eds.). Springer, Cham, 15–26.

[3] Linus W. Dietz, Avradip Sen, Rinita Roy, and Wolfgang Wörndl. 2020. Mining Trips from Location-Based Social Networks for Clustering Travelers and Destinations. *Information Technology & Tourism* 22, 1 (March 2020), 131–166. https://doi.org/10.1007/s40558-020-00170-6

[4] Marta C. González, César A. Hidalgo, and Albert-László Barabási. 2008. Understanding individual human mobility patterns. *Nature* 453, 7196 (June 2008), 779–782. https://doi.org/10.1038/nature06958

[5] Bartosz Hawelka, Izabela Sitko, Euro Beinat, Stanislav Sobolevsky, Pavlos Kazakopoulos, and Carlo Ratti. 2014. Geo-located Twitter as proxy for global mobility patterns. *Cartography and Geographic Information Science* 41, 3 (Feb. 2014), 260–271. https://doi.org/10.1080/15230406.2014.890072

[6] Raja Jurdak, Kun Zhao, Jiajun Liu, Maurice AbouJaoude, Mark Cameron, and David Newth. 2015. Understanding Human Mobility from Twitter. *PLOS ONE* 10, 7 (July 2015). https://doi.org/10.1371/journal.pone.0131469

[7] Anastasios Noulas, Salvatore Scellato, Cecilia Mascolo, and Massimiliano Pontil. 2011. An empirical study of geographic user activity patterns in Foursquare. In *Fifth International Conference on Weblogs and Social Media (ICWSM '11)*. AAAI, Palo Alto, CA, USA, 570–573.

[8] Paul Resnick, Aviv Ovadya, and Garlin Gilchrist. 2019. *Iffy quotient: A platform health metric for misinformation.* Technical Report. School of Information Center for Social Media Responsability University of Michigan.

[9] Avradip Sen and Linus W. Dietz. 2019. Identifying Travel Regions Using Location-Based Social Network Check-in Data. *Frontiers in Big Data* 2, 12 (June 2019). https://doi.org/10.3389/fdata.2019.00012

[10] Weimin Zheng, Xiaoting Huang, and Yuan Li. 2017. Understanding the tourist mobility using GPS: Where is the next place? *Tourism Management* 59 (April 2017), 267–280. https://doi.org/10.1016/j.tourman.2016.08.009