BeRoute

Mapping Project

Zhu Lin, Valeriia Shurupina
2019-2-1

Content

1. Introduction

"BeRoute" is a web-based map application aims to help beer drinkers to find out suitable places for drinking beer and provide them a route to the recommended location in Munich. There are three main parts in this project: emotion scores layer, beer locations and beer routes. Emotion score layer reflects the atmosphere in the space. It helps users to decide which districts they want to go. Beer locations are divided into different categories depending on their types. It helps users decide which kind of place they want to stay for drinking beer. After two-time filtering, users will get suitable locations for them and obtain specific beer routes to these places.

To achieve the above functions, we analyzed multiple social media data, calculated emotion score and did spatial data interpolation. Especially for creating the emotion map, we focus on analyzing Flickr tags. With the development of internet community, more and more users are intended to share their experiences on the social media such as Facebook and Flickr. Flickr as one of the most popular photo social media contains a great number of photo data, these photos contain significant information about location, time, tags, description and weather. We believed result based on Flickr has a certain representativeness and could play a function on analyzing geographical emotion.

The following chapters will describe our project and our work method in details. Chapter 2 provides an overview concept about the website; chapter 3 goes to introduce our used data and the data crawling methods; chapter 4 introduces the work flow for creating the emotion map; chapter 5 is about the route machine; chapter 6 and 7 discuss our present works and our future plan; the sources for the project will be found on chapter 8.

2. Map Description

Designing our mapping project in web environment we used three core technologies of the World Wide Web: HTML, CSS and JavaScript.

As a base of our web page there is a map with different additional tools. Among them there are a navigation bar on the top, zoom control panel, base map and layer switcher, and window with route information along right side. On the left side we implemented a sidebar with filters and a short description about objects (Fig. 1).
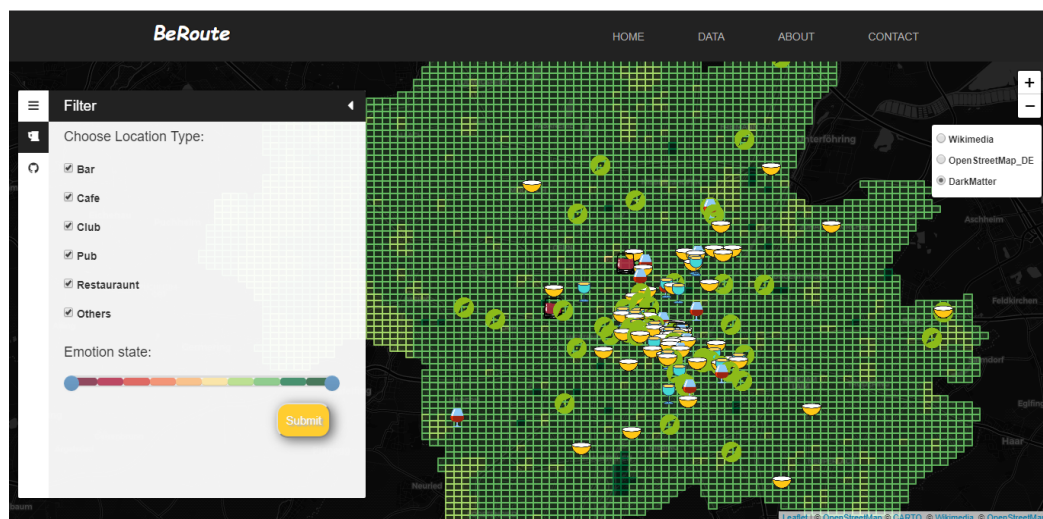

Figure 1. Design of BeRoute web page

We also give to our users free choice among three maps taken on Leaflet: "Wikimedia", "OpenStreetMap_DE" and "DarkMatter". It can help with examination of the details on different backgrounds.

Speaking of the map content, there are a set of places, scores of emotions, and routes. To visualize all these we chose point map symbols for the places categorizing them (Fig. 2).

Figure 2. Categories of places: bar, cafe, club, pub, restaurant, others

To show and make the choice of emotions interactive the range selector was created and divided into 10 equal parts with values from 0 to 10. By choosing the values on the range selector and clicking a submit button all the polygons are filtered and left only ones with chosen values. According to above described filtering process all the checked in checkboxes places (that are inside the polygons) are also filtered and shown on the map (Fig. 2).

Another interactive feature that was implemented is a route planner. Once user clicks on the point symbols a routing mechanism creates the shortest route.

3. Data and data processing

3.1 Flickr Data

Flickr is an image hosting service and video hosting service. Photos and videos can be accessed from Flickr without the need to register an account, but an account must be made to upload content to the site. Registering an account also allows users to create a profile page containing photos and videos that the user has uploaded and also grants the ability to add another Flickr user as a contact [1].

Looking for the crowded and popular places among tourists in Munich, we grabbed and parsed Flickr data containing id, photo id, title, tag, coordinates, location, taken time and their url (Table 1). As our main tools we used Python programming language and Flickr API. Thus, we found 29.610 photos that were taken in Munich area from 01.10.2010 to 30.09.2018. Meanwhile, to grab data within the required area we settled a square boundary with the following coordinates: x(top right) = 11.7229, y(top right) = 48.2482, x(bottomleft) = 11.3608, y(bottom left) = 48.0616 (WGS 84).

Table 1. Example of Flickr data

id	photo_id	title	tag	coordinates	location	taken_time	url
1	4326 4608 170	Blau ride	{blau,bike,bridge,bikeride ,woods,forest}	(48.076129,11.390118)	Gauting	9/22/2018 5:41:07 PM	https://www. m/photos/and o/432646081

3.2 Untappd data

Untappd is an application founded on 2010 for Beer social service. It aims to create a social network by sharing the location where you are drinking beers. Users could check in when they drink beers, post pictures with beers and rank the beer. Till now, it has accumulated about 8 million users [2] and stored large amounts of data

related with beers. Based on these data, they also provided a series of ranking lists on beer venue, beer brand and beer brewery in different cities for users. Considering the huge number of users, we believe these rank lists are reliable at certain extent.

For our project, we are interested in the popularity rank for beer venues in Munich. It provides more than 500 places where beer drinks always visit in Munich and could be the reference for location recommendation in BeRoute[2]. Table 2 shows one record example of Untapped data.

Table 2. Untapped data example

Name	District	Tag	Website
Tegernseer Tal Bräuhaus	Munich	Brewery, Gastropub	/venue/802206

3.3 Google Map data

Scrapped untappd data only provide information about the location name and location tags. There are some important attributes like coordinate, location address, website and address are missing. To make up these data, we decided to import Google Map data.

Google Map is a well-known application launched in 2005 [3] and possess a large amount of Geo-information data such as satellite image, street map, location address and so on. Google provides free API key to developers for acquiring certain range of data. For this project, we used Google places API to get the detailed information about locations. Table 3 shows one record example of Google place data.

Table 3. Google Place data example

name	Coordinates	Address	type
The Keg Bar	[11.5855095, 48.1581038]	Trautenwolfstraße 1,80802München, Germany	Bar

4. Emotion Layer

In our opinion, an ideal spot for beer drinker is the place whose surroundings' atmosphere meet with their emotion. Many location recommendation systems tend to introduce most popular places to users. The results not always meet user's need.

We hope to guide users to identify their emotion, select their location preference and provide more reasonable recommendations.

As Volunteered geographic information, Flickr data is a great data source for extracting geo-related information. Photos from Flickr not only contain georeferenced information but also be tagged with various words relating with photo contents. In 2016, Eva and Dirk used Flickr data to analyze Flickr tag data [4].

Inspired by their works, we referenced the research method and created a Munich emotion distribution layer by analyzing Flickr photos' tags. This layer aims to segment Munich based on the emotion score and helps users to pick out districts which match their current mood for drinking beer.

To create the emotion data layer, there are three main questions need to be

answered:
1. How to extract emotion and quantize them?
2. How to transfer the irregular photo points with emotion score into continuous emotion surface?
3. How to present the emotion variable? They are related with 4.1. semantic analysis 4.2. Emotion data interpolation 4.3. Color for emotion.

4.1 Semantic analysis

In the field of Natural Language Processing, many scientists are researching on digging out intrinsic meaning behind texts. For example, many of them created different word lists for different languages to quantize the emotion behind different words. Their works provide us great supports on creating our emotion map. In our project, we focus on analyzing German and English words from photo tags by using these word lists.

For analyzing German words, we referenced "German Affective Norms List" created by Maximilian Koper & Sabine Schulte im Walde from institute for Natural Language Processing in university of Stuttgart. It contains 350 000 German lemmatized words and each word is rated on four psycholinguistic attributes: Abstractness, Arousal, imageability and valence [5] Each attribute of one word obtains one score range from 0 to 10. For analyzing English Words, we used the word list created by Warriner, A.B., Kuperman, V., & Brysbaert, M. in 2013. They collected 13,915 English words and calculated their valence, arousal and dominance value [6].

To capture the sense of emotion for each tag words, arousal and valence are two attributes we need. The score of valence and arousal define the pleasantness of a word and intensity of emotion provoked by a stimulus separately. As Figure 3 shows, each word existing in the above word lists would have a value location ([valence, Arousal]) in the following coordinate system. The Euclidian distance between value location and principle point is the final emotion score for a word.
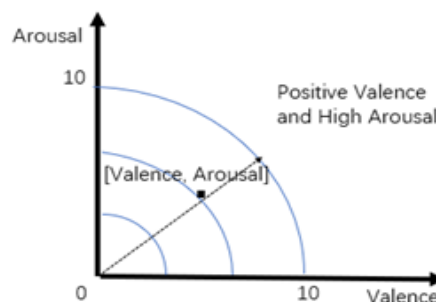

Figure 3

To calculate the emotion score for each photo tag, firstly, we scanned photo tags and find out words which are also existing in word lists. Then, calculating the emotion score for matched words and get mean value among them to define the final emotion value for each tag. The procedure as Figure 4 shows.
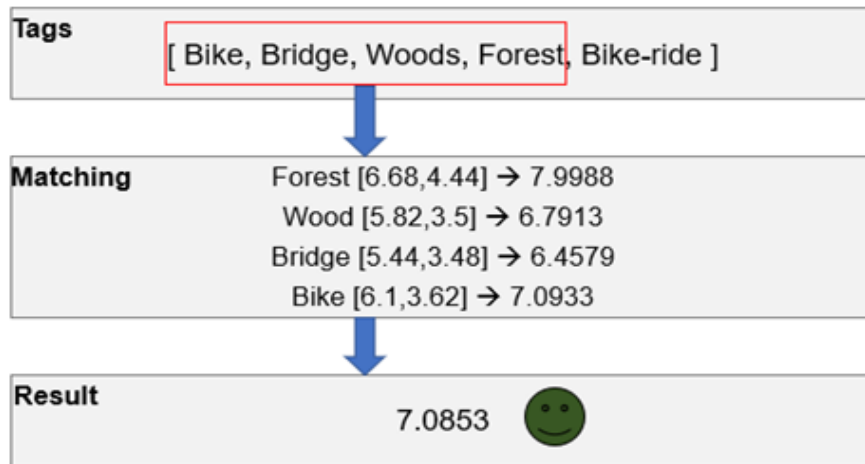
Figure 4. Calculating procedure

Till now, we finished the work on quantizing the emotion, each single photo point obtains unique emotion value. From 29611 Flick data, we deleted data with 0 value.

4.2 Emotion data Interpolation

Based above steps, we obtained an irregular distributed emotion dataset. How to know values of the emotions in every point of the city? Spatial interpolation helps us to transfer the scattered data into continuous surface.

Firstly, we used Delaunay Triangulation and Voronoi Polygons to proportionally divide and distribute a points coverage into regions. Secondly, according to the calculated network, we did an interpolation using IDW method. Before the interpolation we used spherical function to simulate variogram curve and define the output surface stored in 100*100 regular grid (Fig. 5).
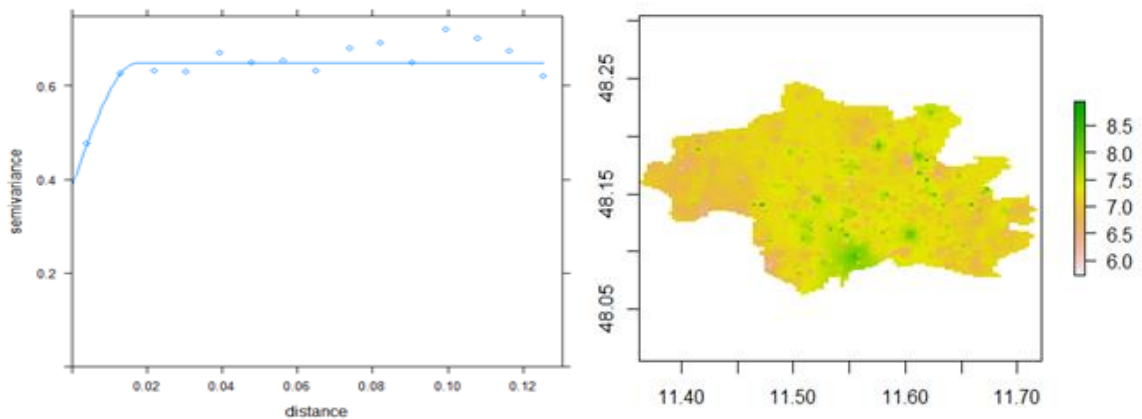

Figure 5. Variogram curve and surface based on emotion scores

4.3 Color for Emotions

To present our dataset with positive and negative emotion scores it is important to find acceptable color scheme. Choosing it we had a consultation with other cartographers where almost everyone suggested to take diverging color scheme from red to green (Fig. 6).
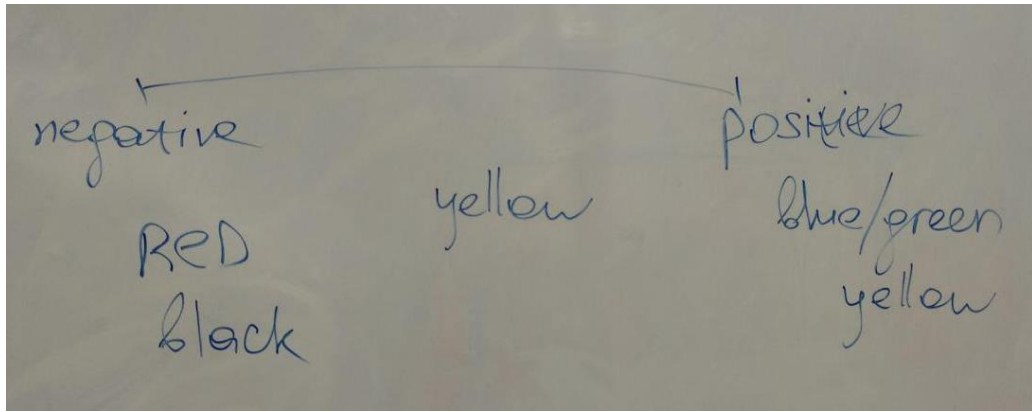
6

Figure 6. Suggested color schemes during session with cartographers

Another important thing we considered is an emotion score distribution. Using ArcGIS we could see the result (Fig. 7). Here it is obvious that the most number of scores are between 6 and 8 values.
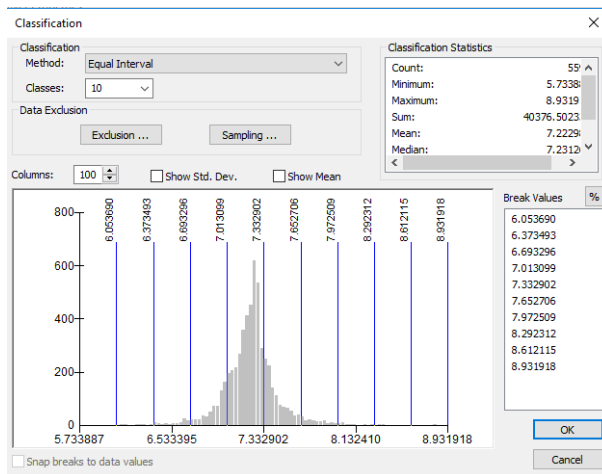

Figure 7. Emotion scores distribution

Finally, having technical issues with making a range selector, we used diverging color scheme with 10 classes taken on Color brewer [7] (Fig. 8). It goes from saturated green to yellow and orange, and then to red colors, e.g. from the mood with relatively high emotion state to respectively low.
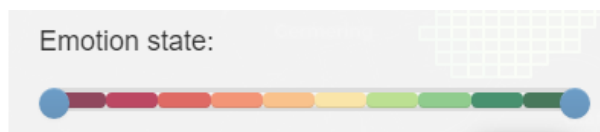

Figure 8. Color scheme for emotion values

5. Beer Route

This part aims to provide a function on leading users from their current position to the selected-location and relies on Leaflet Route Machine with OSRM route engine. Nowadays, there are many commercial companies or research institutes such as TomTom, ESRI, Mapbox provided route engine and opened API for developers. For this project, we considered to provide the shortest walking route for users. This is the most basic requirement and achieved by most route engine. Thus, we chosen OSRM in Leaflet Route machine, the free and easiest one, to meet our need.

6. Discussion

We successfully gathered all the necessary data from Flickr, Untapped, and Google maps and analyzed it according Word Lists. As the tools for the first part Python and R programming languages were used. To present our data and make it interactive we created and designed a web page using HTML, JavaScript, and CSS.

As a result, we received our final service allowing to filter emotion values, choose types of places where user would like to go to, and create routes from the current location to the chosen place.

On the one hand, we almost achieved our plan for the mapping project, On the other hand, there are still some issues deserved to be discussed. First, the limitation from our research views. We tended to depict the emotion differences in the space using Flickr photo tag but neglected to analyze how emotion changes with the time in the certain district. Second, during the semantic analysis procedure, we required the word in tags need to be entirely matched with word in the word lists. This would lead some combining word which also deserved to be scored will lose the chance. Third, there is great space for improving the route. Route could be customized not only based on the distance but also road situation, users' habit. Provide more flexible choices could build more friendly interactive application.

The value of this project is proposing an idea on transferring abstract sense like human's emotion to valuable data and serves our life.

7. Outlook

We think BeRoute is a potential project, in the future, we would like to improve it in three parts:

1. Enriching the information on the map for example adding route situation and traffic light data;

2. Creating our own Munich base map and create the real beer route;

3. Adding time variable and analyzing the how emotion changes with time. Users could have the choice to create their beer route during their pleasure time.

8. Data sources and Code

All of the data sources and codes for our project are stored in our public GitHub repository: https://github.com/BeRouteCarto

Reference

[1] https://en.wikipedia.org/wiki/Flickr

[2] https://en.wikipedia.org/wiki/Untappd

[3] https://en.wikipedia.org/wiki/Google_Maps

[4] Ahmouda A, Hochmair H H. Using Volunteered Geographic Information to measure name changes of artificial geographical features as a result of political changes: A Libya case study[J]. GeoJournal, 2018, 83(2): 237-255.

[5] Maximilian Koper, Sabine Schulte im Walde. Automatically Generated Norms of Abistractness, Arousal, Imageability and Valence for 350,000 German Lemmas. In: Proceeding of the 10th Conference on Language Resources and Evaluation(LREC). Portorož, Slovenia, May 2016.

[6] http://crr.ugent.be/archives/1003

[7] http://colorbrewer2.org