PlaceProfile: Employing Visual and Cluster Analysis to Profile Regions based on Points of Interest

Rafael Mariano Christófano, Wilson Estécio Marcílio Júnior[a] and Danilo Medeiros Eler[b]
São Paulo State University (UNESP), Presidente Prudente/São Paulo, Brazil

Keywords: Area Profiling, Smart Cities, Smart Mobility, POIs, Clustering, Visualization, Google Maps.

Abstract: Understanding how commercial and social activities and points of interest are located in a city is essential to plan efficient cities in smart mobility. Over the years, the growth of data sources from distinct online social networks has enabled new perspectives to applications that provide mechanisms to aid in comprehension of how people displaces between different regions within a city. To support enterprises and governments better understand and compare distinct regions of a city, this work proposes a web application called PlaceProfile to perform visual profiling of city areas based on iconographic visualization and to label areas based on clustering algorithms. The visualization results are overlayered on Google Maps to enrich the map layout and aid analyst in understanding region profiling at a glance. Besides, PlaceProfile coordinates a radar chart with areas selected by the user to enable detailed inspection of the frequency of categories of points of interest (POIs). This linked views approach also supports clustering algorithms' explainability by providing inspections of the attributes used to compute similarities. We employed the proposed approach in a case study in the São Paulo city, Brazil.

1 INTRODUCTION

Since the work of (Ravenstein, 1885), researchers have been focused on understanding displacement patterns to identify how people need to move in a region. With the growth of big cities, the increase in population, society's evolution, and technology innovation, cities have become more diverse and complex than ever. The world is increasingly interconnected. Accordingly, the displacement of people to carry out their daily activities has become a major challenge. Therefore, creating solutions to improve mobility so that people can move from one point to another in an agile and safe way has been a challenge for local governments to manage (D'Andrea et al., 2018). Thus, city planning is closely related to human mobility in an urban territory. As a result, this planning directly influences the population's access to services such as hospitals, schools, parks, and events.

Urban mobility is related to the movement of people and goods in a city, with the objective of developing economic and social activities in the urban areas, urban agglomerations and metropolitan regions (Silva, 2014). In the past, to understand mobility and activities in a city, researchers collected survey data on small samples and low frequencies. Currently, information about activities in a region can be extracted through data from collaborative social networks, in which users enter data on routes, trips, purchases, points of interest, as well as the data acquisition from Internet of Things (IoT) devices. With this information, government officials, authorities and entrepreneurs can understand the organization of regions in a city as well as plan its development. For example, it is possible to label and compare different regions of a city according to the activities employed in an area under analysis.

In the literature, several works present approaches to analyse the urban mobility (Batty, 2009), (Claramunt et al., 2000), (Demissie et al., 2013), (Jiang et al., 2012), and to label regions (Andrienko et al., 2013), (Song and Miller, 2012), (D'Andrea et al., 2018), (Jiang et al., 2012). Usually, those works use machine learning techniques (i.e., classifier and clustering) to deal with data acquired from geolocated data of IoT devices, sensors, points of interest, online posts, traffic information, and other data sources. To improve the user analysis, information visualization technique are employed to enhance map layouts and

[a] https://orcid.org/0000-0002-8580-2779
[b] https://orcid.org/0000-0002-9493-145X

Christófano, R., Marcílio Júnior, W. and Eler, D.
PlaceProfile: Employing Visual and Cluster Analysis to Profile Regions based on Points of Interest.
DOI: 10.5220/0010453405060514
In *Proceedings of the 23rd International Conference on Enterprise Information Systems (ICEIS 2021) - Volume 1*, pages 506-514
ISBN: 978-989-758-509-8

present visualizations on the city maps. While categorizing city regions based on clustering techniques can provide information about city patterns, such an approach might hide important information due to aggregation.

This paper presents a web-based application called PlaceProfile to aid in profiling and characterizing city areas based on visual analysis of points of interest (POIs) of regions in a city. Given a set of POIs and their categories, the proposed approach uses an iconographic visualization to profile areas based on the main categories of POIs, and employs clustering techniques to label areas of a region based on the frequency of the POIs present in each area. The iconographic approach extends clustering analysis power by showing how POIs in different city areas are related and showing the information details that could be used to different clustering results. PlaceProfile also presents a linked views strategy to coordinate the map clustered areas with a radar chart visualization. To augment the interpretability power of clustering results, this coordination mechanism also provides a way to explain the clustering results to aid analysts in a detailed analysis by showing predominant POI activities for each area. We provide a use case to demonstrate how PlaceProfile can be employed to analyze areas of São Paulo, Brazil.

Summarily, the main contribuitions of PlaceProfile are:

- A iconographic visualization approach to show at a glance the main activities of areas in a city;
- A coordination approach between the areas of interest and a radar chart to aid in explainability of clustering and profiling results.

The remaining of this manuscript is organized as follows: in Section 2, we review related works on urban mobility focusing on data mining visualization approaches; in Section 3, we present our tool by detailing all of the pipeline involving in data preparation and visualization; a case study is presented in Section 4; we conclude our work in Section 5.

2 RELATED WORKS

The first applications of visualization systems for the analysis of urban mobility were based on the resources of Geographic Information Systems (GIS) and traditional visualization methods (bar and line charts) with limited interaction capabilities (Claramunt et al., 2000). The advent of new technologies for research and development of visual representations (such as D3.js[1] and Google Maps API[2]) accelerated the number of applications for knowledge discovery considering data from Smart Cities (Sobral et al., 2019).

Usually, human displacement is analysed from various perspectives, such as vehicle traffic, people movement dynamics, incidents, activities in regions of a city, and daily patterns of human activities. In the domain of applications for vehicle traffic analysis, (Andrienko et al., 2013) proposed an application of street diagrams with time-space to analyse urban traffic congestion in the city of Helsinki, Finland. Using a flower visual metaphor, the authors used a rose chart in which the segments of the circle represent the hours of the day, and such segments represent the number of traffic jams and the segment size represents the time duration of the traffic jams.

Heat maps are usually employed to analyse vehicle traffic, which are superimposed on a geographic map of the region to be analysed, (Song and Miller, 2012), (Liu et al., 2013), (Pu et al., 2013). The use of such a technique presents the user with the real perception of the place with the highest incidence of vehicle congestion.

In the domain of applications that analyse the dynamics of displacement of people, that is, what makes a person or group of individuals move from point A to point B, has concentrated on mandatory urban points. For example, in the work of (Sagl et al., 2012), researchers sought to better understand the typical space-time patterns of collective human mobility on the operational scale of a city and its periphery, so that the work was able to reveal similarities and differences in functional configuration of cities in terms of mobility. In the work of (Demissie et al., 2013), the researchers analysed data obtained by mobile telephony, specifically, data known as *call details record* (CDRs), to understand the process of downloading a mobile call on the move, that is, when an active connection is switched from one transmission tower to another. The objective was to emphasize the necessary points in the coverage of the tower signal, to detect the points of cellular congestion and human mobility patterns.

For applications that analyse traffic incidents, (Albino et al., 2015) and (Pack et al., 2009) performed analyses on incident data acquired by departments responsible for public administration. In (Pack et al., 2009), the researchers developed an application that offers an analysis of the data sets of transport incidents. The tool offers the user an intuitive set of features that includes data filtering, geospatial visualiza-

[1]https://d3js.org/

[2]https://cloud.google.com/maps-platform/

tions, statistical classification functions and multidimensional data exploration features.

In the domain of applications with the objective of labelling the activities in regions of a city, (D'Andrea et al., 2018) used data collected from various sources to extract significant characteristics for identifying activities in areas of a city. Moreover, other works have tried to understand patterns from data to answer some questions, such as 'What is the activity that the user is performing based on her geographical position? and 'What is the purpose of her movement?' (Xiong et al., 2014), (Hung and Peng, 2011), (Paul et al., 2013). Although these researchers raise a few important questions, these surveys are limited to identifying activities in specific locations and the data produced in these works were not made available so that other surveys could include data on mobility in order to understand why people move.

And finally, applications that seek to identify daily patterns of human activities in order to improve transport logistics for school, work, leisure, among other. (Jiang et al., 2012), for example, developed a survey on the routes of residents in the Chicago metropolitan area, to analyse the data collected, the researchers used the Principal Component Analysis (PCA) method and the clustering algorithm K-means, in this work the researchers concluded that the techniques were effective in analysing the correlation between the variables and in identifying similar groups respectively. In their most recent work (Jiang et al., 2017), the researchers used data produced by mobile device, known as Call Detail Records (CDR) to examine the mobility patterns of anonymous individuals in the metropolitan region of Singapore.

3 PLACEPROFILE: UNDERSTANDING PATTERNS BASED ON POINTS OF INTEREST

Many complex data containing information about places of interest in a city, cost of living, traffic, and preferences are produced daily. Most data sources on the web, in general, provide information about any city or geographic region, while other data sources are specific to a particular city (D'Andrea et al., 2018). Understanding how this enormous amount of data can be used and the analytical process can help decision-makers.

This paper proposes a web application tool called PlaceProfile, which uses information on points of interest to profile and label areas of a city. PlaceProfile enables visual metaphors and uses clustering algorithms to support users to identify the main activities of different regions of a city or metropolitan region. Figure 1 shows the PlaceProfile's architecture, which consists of four functional components: Preparation, Data Collect, Data Mining Analysis, and Visualization.

The **Preparation** components (see Figure 1 (a)) consists of defining parameters that will be used for the Data Collect step. Firstly, the user defines a region for analysis using the zooming feature – a feature inherited from Google Maps API. Then, cells of the same size (in meters) are used to impose micro-regions on top of the user-defined region. Notice that such a process will result in a grid on top of the region being analysed. Thus, cells (areas) with a smaller size can be used if one wants a more fine-grained analysis.

The data about the user-defined region is collected in the **Data Collect** step (see Figure 1 (b)) by using Google Place API library. Table 1 shows an example of the collected data for a region, where `Type` corresponds to a classification assigned to a point of interest with 130 possible values, and `User Rating` is the evaluation assigned by users – the user rating goes from 1 (bad) to 5 (good). Usually, the data is extracted in batches, stored in a database, and later processed.

Table 1: Raw data collected from Google Places.

Description	Sample value
ID	ChIJfUjHo(...)
Name	Museu Paulista
Geo. coordinates	(-23.5855993, -46.6097431)
Type	museum, establishment
User Rating	4.6

During the **Data Collect** process, the collected points of interest (POIs) are visualized on the grid (as shown in Figure 2) using red circles to encode the density of POIs in each cell. The red points representing the POIs are positioned in their respective geographical location.

In the **Data Mining Analysis** step (see Figure 1 (c)), all of the 130 types of categories retrieved from Google Places are grouped into eleven macro categories (D'Andrea et al., 2018), as show in Table 2, enabling easier selection or deletion by users for latter analysis. For each cell, the number of macro categories is retrieved to compute features (see Table 3) for clustering algorithms and iconographic visualization. PlaceProfile allows clustering using *k*-Means (MacQueen et al., 1967), *c*-Means (Dunn, 1973), and Agglomerative Clustering (Frigui and Krishnapuram, 1997).

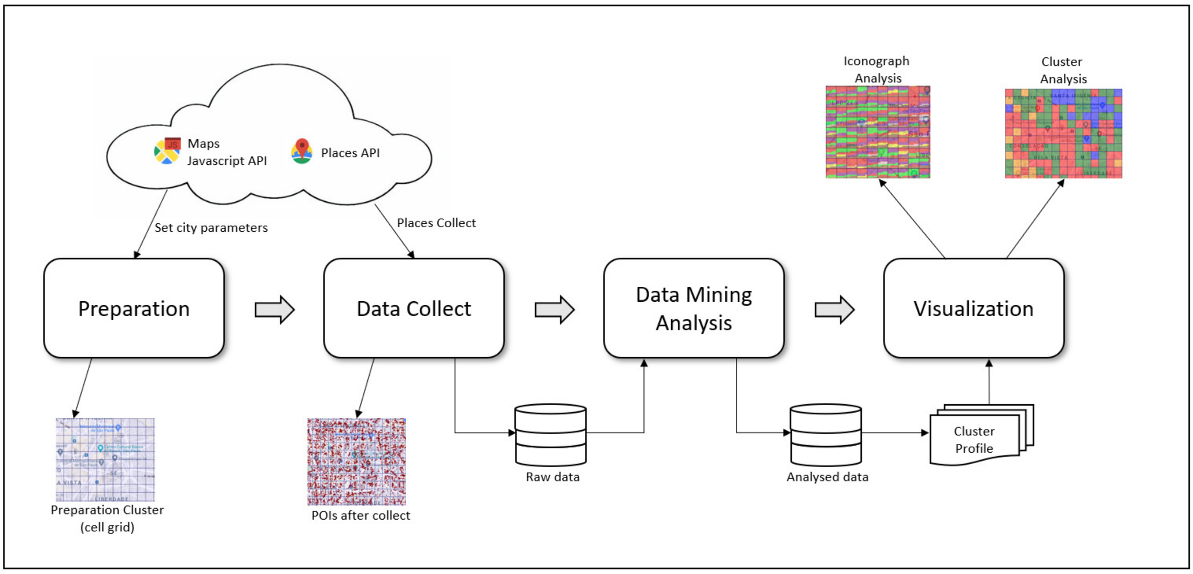

Figure 1: PlaceProfile consists of a web application visual data mining of points of interest. The main components of PlaceProfile are Preparation, Data Collect, Data Mining Analysis, and Visualization.

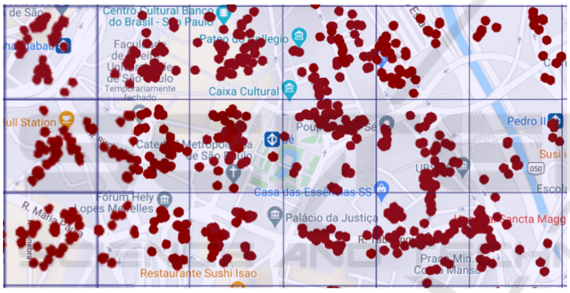

Figure 2: PlaceProfile: The collected data (red dots) are plotted on the grid in their respective cell overlaid on the map.

Importantly, *PlaceProfile* allows the inclusion of other features for later analysis, such as geographic coordinates, a cell identifier, or cell location in the grid. Besides that, users can discard cells with only a few POIs to perform analysis using only informative cells. Finally, the only required parameter is the number of clusters used to analyse the division imposed by the clustering algorithm in the cells, that is, how many k distinct areas on the data will be labeled to create the profile for the user-defined region.

The data resulting from the mining process is stored in a database to facilitate the definition of different visualization strategies. For example, one may use the processed data to create a visualization approach that emphasizes the reasons that citizens change location in a city of a metropolitan region.

In the **Visualization** step (see Figure 1 (d)), three different strategies are used to represent the results of the previous steps: (i) Iconograph Analysis, (ii) Cluster Analysis, and (iii) Radar Chart Analysis.

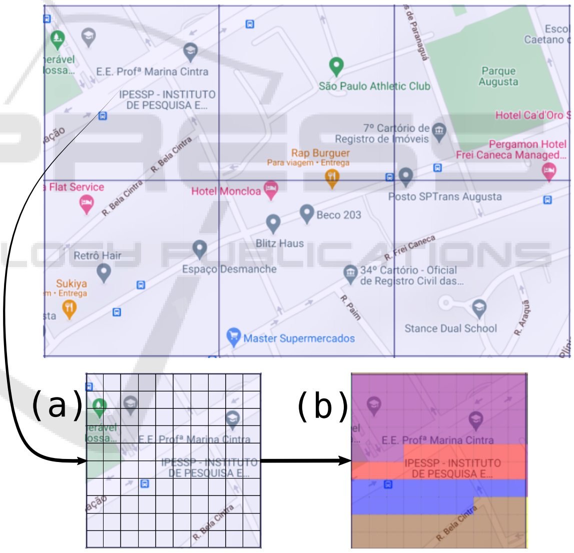

Figure 3: Visualizing the most common macro categories. Given a region divided into cells, we divide the cell space using a grid (a) and color code the grid based on the proportion of macro categories (b). Ordered from top to bottom, the last color (brown) represents the remaining macro categories aggregated.

Figure 3 shows six cells (areas) in the central region of São Paulo before any analysis. To design the Iconograph representation, each cell is divided into a hundred equal parts (see Figure 3), which creates an internal grid in the cell (a). Then, for each cell, the three main macro categories are highlighted by receiving proportional visual space according to their

Table 2: Clustering Google Maps categories into macro categories (Macro Cat.).

Macro Category	Category from Google Maps
Food	bakery, bar, cafe, food, liquor_store, meal_delivery, meal_takeaway, restaurant
Finance	accounting, atm, bank, finance
Admin.	city_hall, courthouse, embassy, police, fire_station, local_government_office
Transport.	airport, bus_station, subway_station, taxi_stand, train_station, transit_station, light_rail_station
Cultural	art_gallery, library, school, university, movie_theater, museum
Entert.	night_club, amusement_park, bowling _lley, campground, zoo, aquarium, stadium, casino
Health	pharmacy, physiotherapist, beauty_salon, dentist, doctor, gym, hair_care, hospital, veterinary_care, health, spa
Services	travel_agency, funeral_home, park, post_office, parking, roofing_contractor, locksmith, general_contractor, lodging, moving_company, car_repair, electrician, car_rental, laundry, gas_station, plumber, painter, real_estate_agency, lawyer, recreational_vehicle_park, insurance_agency, car_wash
Religious	mosque, cemetery, church, hindu_temple, synagogue, place_of_worship
Stores	shoe_store, shopping_mall, pet_store, bicycle_store, book_store, car_dealer, , movie_rental clothing_store, jewelry_store, florist, store, furniture_store, convenience_store, department_store, electronics_store, hardware_store, home_goods_store, storage, grocery_or_supermarket
Misc.	point_of_interest, establishment, country, floor, intersection, locality, natural_feature, geocode, colloquial, area, room, post_box, neighborhood, postal_code, postal_town, political, postal_code_prefix and _suffix, premise, route, street_address, subpremise, street_number

occurrence in the cell (**b**). The most prominent POI macro category appears at the top of the cell, then the second highlighted macro category, the third macro category highlighted comes next, and the fourth represents all the remaining macro categories aggregated.

The result of dividing cells according to their most common macro category is shown in Figure 4 for the six cells presented in Figure 3. Notice that each color represents a macro category and the color occupation in each cell represents the proportion of activities related to that macro category. Finally, the sum of all the remaining macro categories is shown in proportion and represented by the brown color. In this example, the purple color represents the services macro category, red is related to stores, yellow represents food, green represents health, blue represents cultural, and brown represents all other macro categories present in the cell. The analyst can notice the categories that are not predominant in each cell by looking at the legend that describes each macro category employed in the visualization.

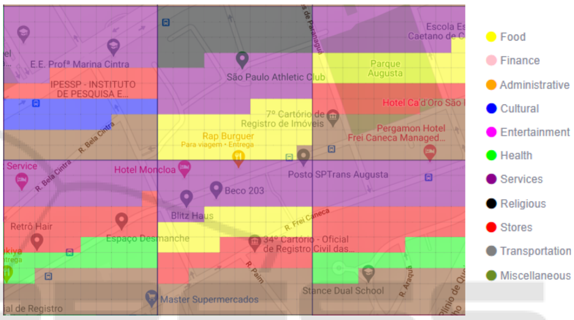

Figure 4: Using colors to encode the most common macro category in a cell. This region is particularly represented by *services* and *stores* macro categories.

In the cluster analysis view, each cell is labeled with a different color for each data group partitioned by the clustering algorithm. Figure 5 shows the result of cluster analysis on the same six cells highlighted in Figure 3. Notice that five out of the six cells belong to the same cluster (cells in red) and one cell has different features (cell in green). The user can understand the pattern adopted by the clustering algorithm to separate groups of data. We must highlight that the colors in the cluster analysis do not relate to the colors used for encoding the macro categories.

To complement the analysis of macro categories and assist in the interpretation of clustering results, PlaceProfile also uses a Radar Chart visualisation to help users to perceive predominant activities for each cell. Interactivity is present in this view, allowing the user to select cells of interest to analyse and make comparisons. Figure 6 shows the differences between the two groups, the red cells stand out for macro categories of services and store activities while in the green cell the highlight is for stores, services, and health.

Table 3: The table illustrates a sample of the summary of the macro categories per cell: the count of the total of the 11 macro categories in each cell creates the attributes that will be passed as a parameter for the grouping algorithm and for the iconographic analysis.

		Macro categories Count				
Cell Id	**GPS coordinates (lat, lng)**	**Food**	**Finance**	**Administrative**	**Transport**	**Cultural**
1	-23.532288, -46.671019	5	10	3	3	5
2	-23.532288, -46.668570	14	1	0	2	2
3	-23.532288, -46.666120	16	3	1	1	6
4	-23.532288, -46.663671	6	9	4	2	0

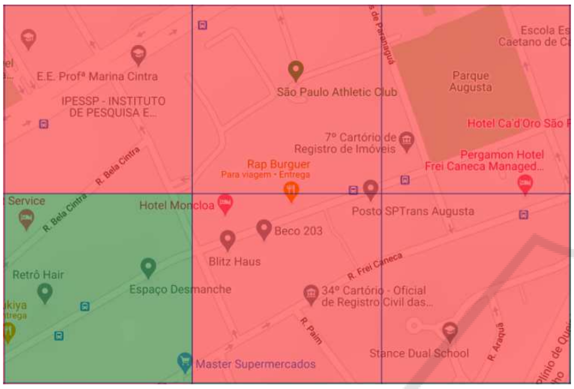

Figure 5: Clustering analysis. Based on the features retrieved during **Data Collect** step, cells are clustered to help analysis based on similarity.

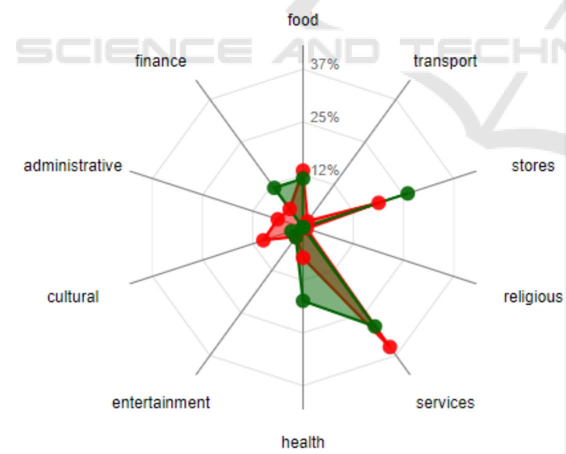

Figure 6: Radar Chart showing the proportion of macro categories for the two clusters presented in Figure 5. Although very similar in *services*, *stores*, and *food*, these two cluster differ in *cultural*, *health*, and *finance* macro categories.

Implementation. The PlaceProfile front-end was developed using HTML5, the website was styled using the CSS3 language and Javascript was used to perform the interactivity actions of the website. Besides, Bootstrap (https://getbootstrap.com/) libraries were used to optimise the styling process, JQuery library (https://jquery.com/) for event handling and interactivity with External APIs. On the PlaceProfile backend, the language used was Python (https://www.python.org/), along with Pandas (https://pandas.pydata.org/) library for the cleaning process and raw data mining and scikit-learn (https://scikit-learn.org/) for the application of clustering algorithms. We used Google Maps API for the construction of the visualizations, graph rendering, and maps.

4 RESULTS

To validate PlaceProfile, this section presents an analysis of POIs collected from the central region of São Paulo city, Brazil. For this analysis, we defined cells with a size of 250×250 meters, generating a grid with 527 cells overlapping the region of interest. After data collection, 27454 POIs were identified. We removed the *miscellaneous* macro category since it is not related to any other macro category, resulting in ten macro-categories for analysis.

Figure 7 shows the most common macro categories for each cell in the region of interest. We can see a lot of patterns in this region. Firstly, the eastern region is represented by the most POIs categorized by *health* (due to the amount of green in the cells). Second, the northeast of the region shows macro categories related to *stores* (reddish cells). Finally, POIs related to the *food* (in yellow) and *services* (in purple) can be seen throughout the whole region, although more concentrated in the center. Cells with POIs below a minimum threshold (in this case, one) receive black color.

Next, we proceed to analyze the similarity among the POIs collected for this region with clustering analysis. Using the same data and the same preprocessing steps discussed above, we performed the k-means clustering algorithm. Figure 8 (top) shows the result of such clustering, where each color corresponds to a different cluster. Notice that the clustering result shows disconnected components in the visual

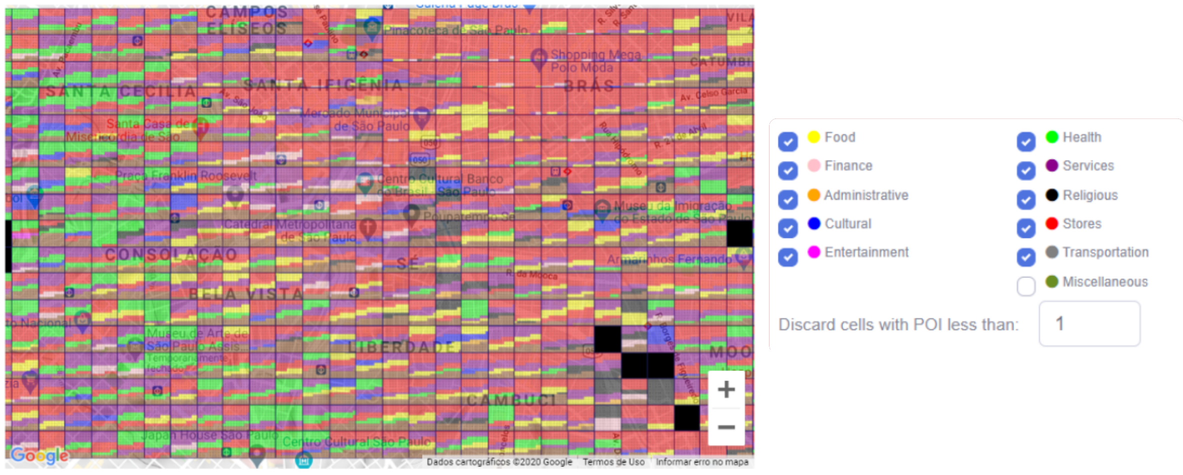

Figure 7: Using PlaceProfile to get an overview of a region of a metropolitan region. We can see mainly a division between POIs related to *stores* and POIs related to *health*. Black cells correspond to regions with number of POIs below a minimum threshold.

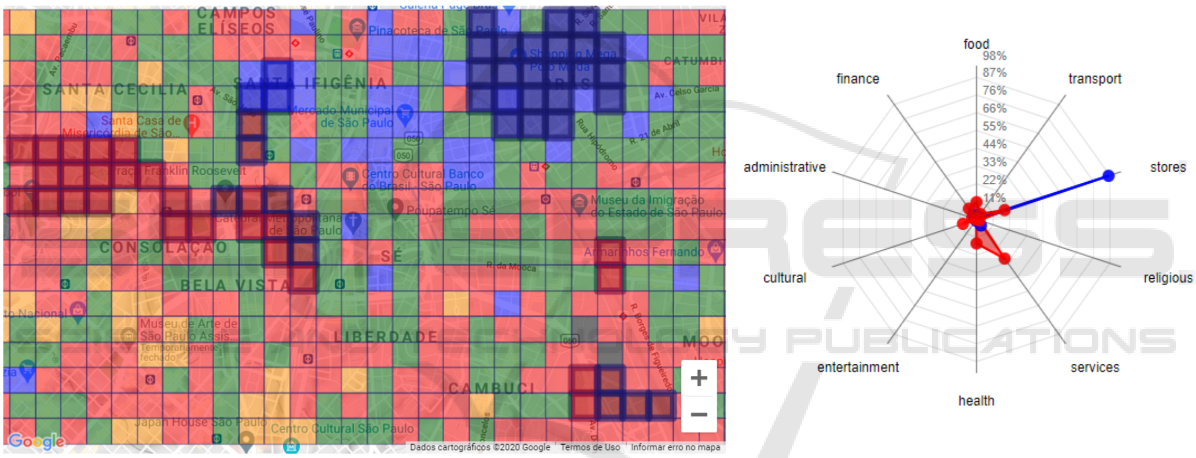

Figure 8: Similarity analysis based on clustering results. Selected cells shows that blue cluster represents POIs related to *store* macro category while red cluster represents POIs related to *health* and *services* macro categories.

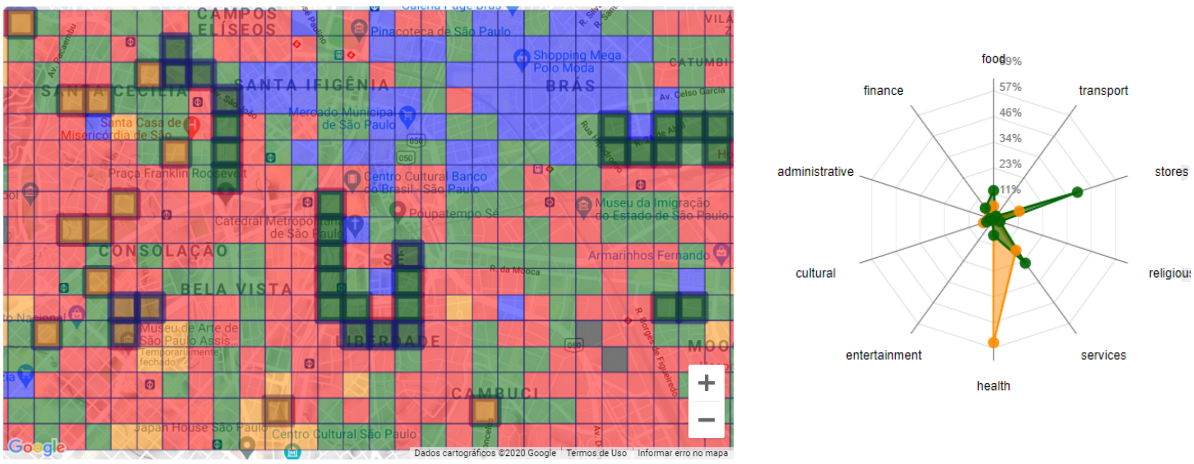

Figure 9: Selecting different clusters to understand patterns of POIs. Cluster green seems to have POIs related to *store* and *services* macro categories, while cluster yellow corresponds to POIs related to the macro category *health*.

representation since the *k*-means was performed in a high dimensional space computed with the macro-categories but without the GPS coordinates.

One of the problems of analysing the result of a clustering algorithm is that information about which features led to the clustering patterns is usually lost.To explain how the information in the regions led to clusters formation, a Radar Chart showing the proportion of POIs in each region can be used through a coordination mechanism. Figure 8 (bottom) shows the Radar Chart encoding the proportion of POIs for the red and blue cells selected and highlighted in the grid visualization with thicker borders. Thus, we can know how those selected cells belonging to the blue cluster present a much higher proportion of POIs related to the macro category *store* than any other, whereas the cells selected from the red cluster present POIs related to the *services* and *health* macro categories. Notice that this analysis is consistent with the iconograph representation seen in Figure 7. That is, while the blue cluster is related to the northeast region of Figure 7 – cells with POIs highly related to macro category *store*, the red cluster is consistent with the cells showing POIs related to *health* and *services* in the same figure.

Figure 9 shows the same clustering result but with different selected cells. Using coordination between the map and the Radar Chart to help on the cluster explainability. Using the Radar Chart, the cells from the green cluster represent POIs related to the macro categories *store* and *services*. Notice that, these cells are the ones located on the boundaries of the blue cluster and the neighborhood of the area concentrated with POIs of *store* macro category in Figure 7. Lastly, we see from the Radar Chart that yellow cluster corresponds to POIs related to *health* macro category. This information is also perceived using the iconograph representation of Figure 7.

In this case study, we showed how the Radar Chart's explainability mechanism could help users understand the cluster formation based on the analysis of the proportion of POIs in cells of clusters of interest. Nevertheless, the iconographic approach also provides analysis improvements since it encodes patterns of cells presenting a similar proportion of POIs with the same category. The iconographic approach can provide an overview of data organization and details about POIs at the same time. Thus, the proposed approach presents new mechanisms to improve the cluster-based analysis.

5 CONCLUSION

In this paper, we present PlaceProfile, a web-based visualization tool to identify the profile of areas in cities or metropolitan regions. PlaceProfile allows the definition of regions, granularity of analysis, and other parameters to assist in analyzing patterns based on points of interest.

The main advantage of our tool consists of its ability to provide understanding about the areas being analyzed. First, we augment clustering results by coordinating a map with a Radar Chart that shows the proportion of points of interest in selected cells. Thus, users can get to know how these cells differentiate or relate to contributing to cluster formation. Second, our iconographic approach extends cluster analysis by showing an overview and detailed information simultaneously. On a higher level, users understand the result and possible cluster formation by inspecting the iconographic design's color patterns. In a detailed analysis, users inspect the proportion of points of interest inside cluster cells. We show the analysis power of these two approaches by inspecting a region in São Paulo, Brazil.

The analysis and interaction mechanisms were validated through a case study in a region of São Paulo, Brazil. The approach helped us understand how the points of interest are organized so that near areas present similar categories. Moreover, the case study also demonstrated how the two analysis strategies (clustering and iconographic) are consistent.

We plan to accommodate mobility data to analyze displacement patterns and recommend points of interest according to the cluster and iconographic profiles in future works.

ACKNOWLEDGMENTS

This work was supported by FAPESP (São Paulo Research Foundation), grant number #2018/17881-3 and #2018/25755-8.

REFERENCES

Albino, V., Berardi, U., and Dangelico, R. M. (2015). Smart cities: Definitions, dimensions, performance, and initiatives. *Journal of Urban Technology*, 22(1):3–21.

Andrienko, G., Andrienko, N., Bak, P., Keim, D., and Wrobel, S. (2013). Visual analytics focusing on spatial events. In *Visual analytics of movement*, pages 209–251. Springer.

Batty, M. (2009). Cities as complex systems: Scaling, interaction, networks, dynamics and urban morphologies.

Claramunt, C., Jiang, B., and Bargiela, A. (2000). A new framework for the integration, analysis and visualisation of urban traffic data within geographic information systems. *Transportation Research Part C: Emerging Technologies*, 8(1-6):167–184.

D'Andrea, E., Ducange, P., Loffreno, D., Marcelloni, F., and Zaccone, T. (2018). Smart profiling of city areas based on web data. In *2018 IEEE International Conference on Smart Computing (SMARTCOMP)*, pages 226–233.

Demissie, M. G., de Almeida Correia, G. H., and Bento, C. (2013). Exploring cellular network handover information for urban mobility analysis. *Journal of Transport Geography*, 31:164–170.

Dunn, J. C. (1973). A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters.

Frigui, H. and Krishnapuram, R. (1997). Clustering by competitive agglomeration. *Pattern recognition*, 30(7):1109–1119.

Hung, C.-C. and Peng, W.-C. (2011). A regression-based approach for mining user movement patterns from random sample data. *Data & Knowledge Engineering*, 70(1):1–20.

Jiang, S., Ferreira, J., and González, M. C. (2012). Clustering daily patterns of human activities in the city. *Data Mining and Knowledge Discovery*, 25(3):478–510.

Jiang, S., Ferreira, J., and Gonzalez, M. C. (2017). Activity-based human mobility patterns inferred from mobile phone data: A case study of singapore. *IEEE Transactions on Big Data*, 3(2):208–219.

Liu, S., Pu, J., Luo, Q., Qu, H., Ni, L. M., and Krishnan, R. (2013). Vait: A visual analytics system for metropolitan transportation. *IEEE Transactions on Intelligent Transportation Systems*, 14(4):1586–1596.

MacQueen, J. et al. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 281–297. Oakland, CA, USA.

Pack, M. L., Wongsuphasawat, K., VanDaniker, M., and Filippova, D. (2009). Ice–visual analytics for transportation incident datasets. In *2009 IEEE International Conference on Information Reuse & Integration*, pages 200–205. IEEE.

Paul, A., Chen, B.-W., Bharanitharan, K., and Wang, J.-F. (2013). Video search and indexing with reinforcement agent for interactive multimedia services. *ACM Transactions on Embedded Computing Systems (TECS)*, 12(2):25.

Pu, J., Liu, S., Ding, Y., Qu, H., and Ni, L. (2013). T-watcher: A new visual analytic system for effective traffic surveillance. In *2013 IEEE 14th International Conference on Mobile Data Management*, volume 1, pages 127–136. IEEE.

Ravenstein, E. G. (1885). The laws of migration. *Journal of the statistical society of London*, 48(2):167–235.

Sagl, G., Loidl, M., and Beinat, E. (2012). A visual analytics approach for extracting spatio-temporal urban mobility information from mobile network traffic. *ISPRS International Journal of Geo-Information*, 1(3):256–271.

Silva, E. (2014). *Meio ambiente & mobilidade urbana*. Série Meio ambiente. Editora Senac São Paulo.

Sobral, T., Galvão, T., and Borges, J. (2019). Visualization of urban mobility data from intelligent transportation systems. *Sensors*, 19(2).

Song, Y. and Miller, H. J. (2012). Exploring traffic flow databases using space-time plots and data cubes. *Transportation*, 39(2):215–234.

Xiong, H., Zhang, D., Zhang, D., Gauthier, V., Yang, K., and Becker, M. (2014). Mpaas: Mobility prediction as a service in telecom cloud. *Information Systems Frontiers*, 16(1):59–75.