Visualizing Functional Regions by Analysis of Geo-textual Data

Yunzhe Wang[1], George Baciu[1] and Chenhui Li[2]

[1]The Hong Kong Polytechnic University
[2]East China Normal University

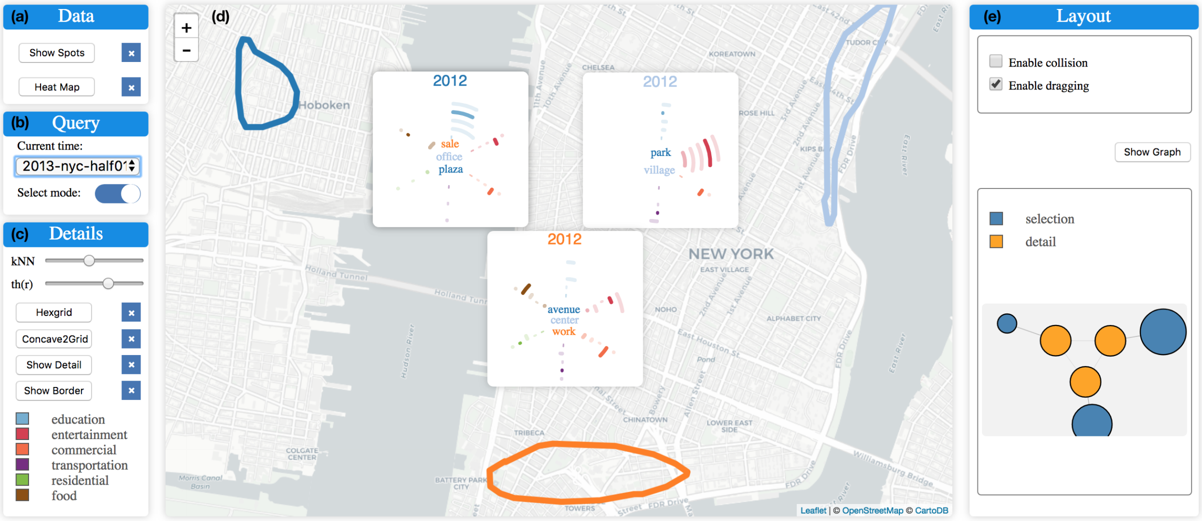


Figure 1: *The interface of our system consists of (a) Data menu. Scatter plot and heatmap are provided for displaying the the distribution of data. (b) Query menu. Users can retrieve data at specific time, or enable the free selection mode to query an AOI (Area Of Interest). (c) The Details menu allows users to change parameters, view details of AOIs and reference the color legend. (d) Map view of the current time step. (e) In the Layout menu, users can adjust positions of details.*

Abstract

Using tremendous geo-textual data collected from social media applications, we facilitate the analysis of region functions. Revealing the evolution of regions is also feasible due to the wide distribution and constant involvement of users. By extracting semantics from textual properties, we aim at classifying geographical locations in terms of their functional types. Hence, we train a classification model with the Support Vector Machine, and apply it to aggregated word embeddings to predict the function of spots. We highly cooperate with techniques in graph analysis. Firstly, regions are segmented based on a latent graph. Then, we propose an adaptive layout solution to deal with situations of multi-AOI queries. The generated layout and interactive metaphor provide convenience for observation and comparison. Experiments are conducted with the YFCC100M dataset to prove the effectiveness of our system.

CCS Concepts

•***Human-centered computing*** → ***Visual analytics;***

1. Introduction

The interplay between human activities and the dynamics of regions affect decisions of various issues, including urban planning, commercial development, maintenance of social order. People doing the same type of activities tend to gather together. This phenomenon impels the formation of functional regions, such as food court, CBD, residential communities. The development and recession of these regions might in turn affect the human activities.

Functional division provides a legible understanding of the composition of a city. Analysts gain insights into the geographical distribution and the influence of various functional regions. Some regions are highly dynamic. Their territories might expand or shrink, and the inner functions even change. Normally, on-spot investigation is not practical. Most of the existing work explores human movements to divide regions. However, mobility data are typically obtained from transportation tools like taxi and metro, which cannot provide a full coverage of focal areas. Besides, the interpretation of region functions still relies on prior knowledge.

Recently, the geo-textual data burst with the prevalence of social media. On photo-sharing platforms, users can record geo-positions(i.e., ***spots***)where photos were taken. They are allowed to attach descriptive texts which imply contexts of the spots. We propose to cluster closely-related spots using a graph-based method. Each cluster is regarded as a ***region***. After inspecting the functional types of inner spots, we find that some regions have ***explicit*** functions, while others possess a mixture of functions. The contributions of our work are as follows,

- We reduce the feature diversity of geo-locations by classifiying semantics extracted from geo-textual data. Semantics are represented by aggregating embeddings of representative words in posts.
- Regions are segmented by maximizing the modularity of a latent graph. The graph integrates both spatial and functional proximity. Therefore, closely located spots with similar context constitute a region.
- Instead of creating a separate view, we place AOI details on the map and optimize their positions. By mapping AOIs and details to nodes in a graph, force layout method can be applied to ensure that the detail position is close to the corresponding AOI and all other details of remaining AOIs simultaneously.

2. Related Work

Add citation 1 [MJR*11], citation 2 [LWL*17].

Region segmentation can be achieved by recognizing different patterns of human activities. People tend to perform similarly within a region. Cranshaw et al. [CSHS12] counted the visiting frequency of a location by all users, and locations of similar visiting patterns are placed into the same region. This work lacks the interpretation of region functions. Main roads can be treated as a reference for segmentation. Yuan et al. [YZX*15] merged small blocks generated by road crossings, and used POI statistics and a topic model to infer the function. One issue is that their method cannot further divide a large region, due to the limitation of road network. MobiSeg [WZC*17] is a visualization system that supports dynamic updating of region segmentation. Data from multiple sources are fused to mitigate the sparsity problem. The segmentation is based on an initial tessellation procedure. Interestingly, both work of Wu et al. [WZC*17] and Yuan et al. [YZX*15] proposed analogies to the context of textual analysis, while our work directly manipulates the textual data.

Extensive research has been conducted to uncover the functions of regions, also known as the land use. Previous studies took remote-sensing satellite images as input and get results at a coarse level. For example, Voorde et al. [VdVJC11] classified areas in terms of residential, commercial, service and green space. To attain fine-grained categories of land use, supervised classification methods as adopted by Pan et al. [PQW*13] need to be applied. We carried out a similar work. The difference is that their training set contains regions and manually labeled functions, while our training set consists of spot-function pairs. We believe that it is more reasonable to classify spots, because some regions may have serveral functions involved. Hence, it is not easy to decide a specific function to label them.

There have been plenty of work to cluster similar semantics of texts. In natural language processing, word embedding aims at converting a word to a numeric vector for further operations like classification and regression. Results produced by topic modeling [FZLC16] are clusters of similar words, but they may suffer from high diversity.

Many visual analytic systems for geographical data prefer to put details that users query in a separate view, mostly because they want to provide organized views to facilitate comparisons. For example, Wu et al. [WZC*17] place the visualization of mobility feature vectors of local districts in a detail view. However, users have to switch between two views to map their queries and corresponding details. Yang et al. [YDGM17] use a leader line to connect locations with visualizations. In this work, we place details next to queried areas to relieve users from mapping.

3. Method

3.1. Data Format

The dataset we used to verify the efficiency of our method is the YFCC100m dataset [TSF*16]. Especially, we focus on data of New York City, and we aggregate data by half of a year. For each data item, we filtered 9 attributes, (***user_id, time, upload_time, title, description, tag, longitute, latitute, url***), where *time* is when the photo was taken. *title, description* and *tag* describe content in the photo. Original texts of the three attributes might involve random characters and stop words, so we clean them and only preserve meaningful English words. In the end, a spot i associates with a list of words, $W_i = \{w_1, w_2, \dots, w_k\}$. *longitute, latitute* give the position where the photo was taken. We can access the photo through the *url* attribute.

3.2. Spot Semantics

The function of a region is revealed by investigating the characteristics of its inner spots. We get the spot(e.g., i) features by analyzing semantics of the corresponding word lists(e.g., W_i). Hence, we can deduce if the place is famous for food, tourism, or commerce. First of all, we need to project words into a vector space. Word2Vec [MCC*14] is adopted to extract feature vectors from words. The distributed representations [MSC*13] generated by Word2Vec embed rich context information. For training, we input a corpus of 3 million words. Finally, a word w_j can be represented by a 300-dimensional vector, $v_j = [f1, f2, \dots, f300]$.

As a spot might associate with a list of words, its semantics S_i can be described by aggregating vectors of words in W_i.

$$S_i = g([t_1, t_2, \dots, t_k] \begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_k \end{bmatrix}), \tag{1}$$

where $g(\cdot)$ is an aggregation operator, which can be min/max, or mean of each dimension in word vectors [DBVCDD16]. In this work, we adopt the max operator. Each vector(i.e., v_j) is weighted by the TF-IDF value(i.e., t_j) of word(i.e., w_j),

$$t_j = \frac{n_{ji}}{\sum_k n_{kj}} \times log \frac{|P|}{|\{i : w_j \in p_i\}|}, \tag{2}$$

where n_{ji} denotes the frequency that w_j occurs in W_i. $\sum_k n_{ki}$ is the sum of all word frequencies in W_i. $|P|$ is the number of posts in one time step, and the denominator in IDF is the number of posts which contain the word w_j.

Our ultimate goal is to classify spots into different types. Therefore, we trained a classification model which is based on SVM and a RBF kernel. In the training set, each spot is manually labeled with an appropriate type. In this work, we have six types ***[residential, commercial, transportation, food, education, entertainment]***. Namely, spot i can be represented by $[S_i, T_i]$ in the training set, where T_i is one of the six types. Functions of all other spots are then predicted by the classification model.

3.3. Region Segmentation

To divide a whole area into multiple regions, we hope to find spots that are both closely located and are of similar functions. For this purpose, we construct a graph G by connecting a spot to its k Nearest Neighbors(kNN). The weight of links are measured by the cosine similarity between semantic vectors of two end spots i and j, $E_{ij} = \frac{S_i \cdot S_j}{||S_i|| ||S_j||}$. Therefore, we can take both spatial and functional proximities into consideration.

Then, we conduct Community Detection on G by maximizing the modularity [RSC*10], which is defined as

$$Q = \frac{1}{2m} \sum_{uv} \left[A_{uv} - \frac{k_u k_v}{2m} \right] \frac{s_u s_v + 1}{2}, \tag{3}$$

where u and v are nodes in the graph. A is the adjacency matrix. k_u is the degree of node u. s indicates the membership of a node to community. In this way, corresponding spots are separated into disjoint communities. We take the geographical coverage of a community as a region, and the boundary can be calculated by finding the concave hull of spots. The function type of a region is decided according to statistics of types of inner spots. Details are given in Section 4.

3.4. Visualization

Conversed to traditional map exploration tools, our system is required to provide heterogeneous utilities: (1) basic, view distribution; (2) extract functions from textual data/project diverse semantics to visual representation of functions; (3) more;

The highlight of our visual representation manifest in three aspects.

Firstly, for the benefit of aesthetic perception [CCW*16] and easy comparison, we convert concave polygons to conjoint hexagon cells. Since original polygons are in irregular shapes, it is difficult to visually compare the size of different regions. After conversion, the region coverage can be denoted by the number of hexagon cells that it contains, because all cells are in uniform size. As shown in Figure 5, we spread a hexagon grid onto the focal area. Then, we use the Scanline method [WREE67] to check all cells in the grid. By applying the ray casting method [Rot82], if any of the six vertices of a cell is found inside a polygon, we assign this cell to the region that the polygon corresponds to.

Secondly, we allow users to freely select an Area Of Interest(AOI) and query the temporal details. When users click the border of an AOI, details will pop up on the map. Multiple kinds of information are encoded to demonstrate the evolution of AOI. The scalable metaphor design can be found in Figure 2. Generally, it looks like a multi-ring donut chart and each ring depicts the data of one time step. From the inside out, rings are placed following a chronological order. At each time step, the ring is equally divided into 6 sections and each denotes the percentage of spots of a particular functional type. For each type, arcs at all time steps are aligned along an axis. Therefore, users can observe the distribution of all types at one time step, and meanwhile, they can find the changes of a certain type over time. To ensure that an arc does not go beyond the corresponding section, the length l is calculated by $l = N_t * (1/6 * 2\pi * r)/N$. N_t is the number of spots of the type and N is the total number of spots. r denotes the radius of the ring. We also preserve a space of circle in the center to put representative words. Those words are chosen from posts in term of frequency. The black dot shows the existence of *explicit regions*. The partial arc (e.g., arc1) from axis to the dot indicates the ratio of spots belonging to the explicit region.

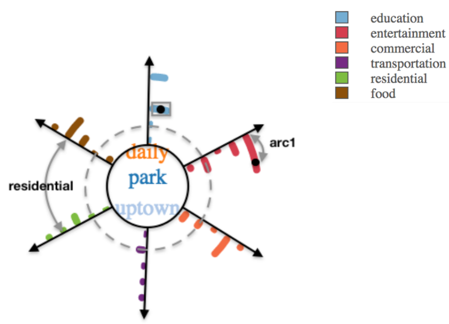


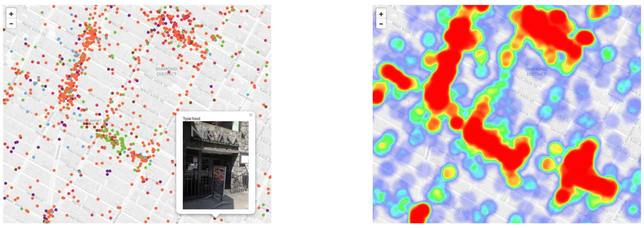
Figure 2: *Visual design of details. It consists of 6 equally-occupied sections. Each ring encodes data at a time step. Arcs of a section align along the black arrow.*

Thirdly, when users select several AOIs and compare their details, we utilize force layout algorithms [Dwy09] to decide the layout of details. Force layout algorithms have been widely used in graph drawing. We abstract AOIs and details as nodes, which are depicted by their smallest enclosing circle. We want details to stay as close as possible to the related AOI, so that users do not need to take a long eye movement for mapping. Besides, all details should get close to facilitate the comparison. Hence, we connect AOIs with corresponding details and set the edge weight to 0.9. All details are fully connected with an edge weight of 0.5. The link length is negatively correlated to the connection weight. Users are free to change positions of details, based on the force layout.

4. Experiment

To evaluate our system, we retrieve all posts in New York city during 2009 to 2014. Time interval is set to be half a year. Users might take several photos at a spot and no textual infomation added. For example, results in this section are based on data in the first half of year 2013, and there are 39099 posts. However, only 10406 spots contribute to region segmentation. We also find that photos taken at the same spot tend to reflect a similar theme.

From Fig. 1, we divide the interface into five parts. In *data* section, there are two ways to display the spot distribution, as shown in Fig. 3. If the user wants to access photos, he needs to click the circle which denotes a spot. Heatmap helps users to locate a potential AOI by density. We also found that explicit regions are more likely to lie in high density areas.



(a) (b)

Figure 3: *Partial distribution of data in the first half of year 2013 in NYC. (a) In the scatter plot, spots are filled with different colors to show their functional types after classification. (b) Heatmap signifies high density areas with warmer colors.*

Then, each spot gets connected to its 10 nearest neighbors. Figure 4 presents the resulted graph. Communities are detected on such a graph and they delimit by bounding concave polygons, as shown in Fig. 4(b). The average number of spots at all time steps is 15978, and the average time of community detection is 0.6802 seconds. We set the explicit ratio to 0.7, which means the region is *explicit* if over 70% of spots have the same function. However, as shown in Fig.5, only a few of them are *explicit*. It implies that many regions in NYC integrate with multiple functions. We can also find regions where only one type dominates. For instance, R1 in Fig.5 turns out to be the Metropolitan Museum of Art, which is defined as *education* related.

In this work, the width of hexagon cells is 100 meters. In Fig. 5, we convert all ***explicit*** regions into combinations of hexagon cells. In NYC, many regions are commercial related(i.e., regions filled with *orange*). In our definition, all themes about shopping, business, finance are categorized as commercial. Most of these regions locate in the center of NYC. After conversion, regions close to each

(a) (b)

Figure 4: *Cluster spots based on kNN graph. (a) Graph constructed by connecting spots with its 10NN neighbors. (b) Communities are delimited by concave boundaries.*

other merge together, such as R2 in Fig. 5. Actually, we can see that R2 stretches along the Fifth Avenue, which is known as one of the most expensive shopping streets in the world.

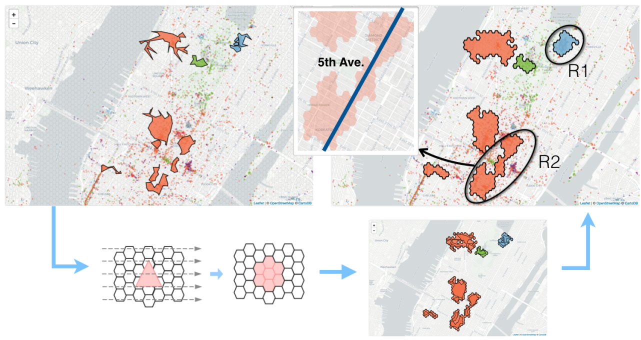
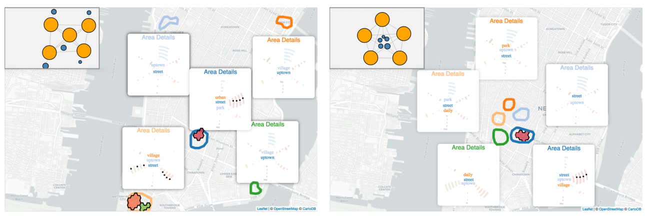


Figure 5: *Convert concave polygons to hexagon cells, for the convenience of region comparison. By applying the ScanLine method, all cells intersected with the polygon will be filled with the same color.*

According to Fig. 6, our system can adaptively place details on the map. Thus, users can get gid of one-to-one mappings between two separate views. All details stay as close as possible to each other. Thus, it is convenient to compare them.



(a) (b)

Figure 6: *Given positions of AOIs selected by users, our system can adaptively put detail views around them. Five AOIs (a) far from each other. (b) close to each other.*

5. Conclusions

In this work, we segment regions by clustering similar spots. The similarity is measured by analyzing semantics of geo-textual data. To decide if a region has an explicit function, we inspect the functions of all inner spots. In our visual system, users can observe the overall distribution of all explicit regions, or they can query a certain area and view the temporal evolution. For the future work, we plan to calibrate the classification by considering image semantics. Also, we need to conduct user studies to evaluate the performance of our adaptive layout method.

Acknowledgements

The authors would like to acknowledge the partial support of IGRF PolyU 152142/15E and Project 4-ZZFF from the Department of Computing, The Hong Kong Polytechnic University.

References

[CCW*16] CHEN S., CHEN S., WANG Z., LIANG J., YUAN X., CAO N., WU Y.: D-map: Visual analysis of ego-centric information diffusion patterns in social media. In *Visual Analytics Science and Technology (VAST), 2016 IEEE Conference on* (2016), IEEE, pp. 41–50. 3

[CSHS12] CRANSHAW J., SCHWARTZ R., HONG J. I., SADEH N.: The livehoods project: Utilizing social media to understand the dynamics of a city. 2

[DBVCDD16] DE BOOM C., VAN CANNEYT S., DEMEESTER T., DHOEDT B.: Representation learning for very short texts using weighted word embedding aggregation. *Pattern Recognition Letters 80* (2016), 150–156. 3

[Dwy09] DWYER T.: Scalable, versatile and simple constrained graph layout. In *Computer Graphics Forum* (2009), vol. 28, Wiley Online Library, pp. 991–998. 3

[FZLC16] FENG K., ZHAO K., LIU Y., CONG G.: A system for region search and exploration. *Proc. VLDB Endow. 9*, 13 (Sept. 2016), 1549–1552. URL: http://dx.doi.org/10.14778/3007263.3007306, doi:10.14778/3007263.3007306. 2

[LWL*17] LIU D., WENG D., LI Y., BAO J., ZHENG Y., QU H., WU Y.: Smartadp: Visual analytics of large-scale taxi trajectories for selecting billboard locations. *IEEE transactions on visualization and computer graphics 23*, 1 (2017), 1–10. 2

[MCC*14] MIKOLOV T., CHEN K., CORRADO G., DEAN J., SUTSKEVER L., ZWEIG G.: word2vec, 2014. 2

[MJR*11] MACEACHREN A. M., JAISWAL A., ROBINSON A. C., PEZANOWSKI S., SAVELYEV A., MITRA P., ZHANG X., BLANFORD J.: Senseplace2: Geotwitter analytics support for situational awareness. In *Visual Analytics Science and Technology (VAST), 2011 IEEE Conference on* (2011), IEEE, pp. 181–190. 2

[MSC*13] MIKOLOV T., SUTSKEVER I., CHEN K., CORRADO G. S., DEAN J.: Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems* (2013), pp. 3111–3119. 2

[PQW*13] PAN G., QI G., WU Z., ZHANG D., LI S.: Land-use classification using taxi gps traces. *IEEE Transactions on Intelligent Transportation Systems 14*, 1 (2013), 113–123. 2

[Rot82] ROTH S. D.: Ray casting for modeling solids. *Computer graphics and image processing 18*, 2 (1982), 109–144. 3

[RSC*10] RATTI C., SOBOLEVSKY S., CALABRESE F., ANDRIS C., READES J., MARTINO M., CLAXTON R., STROGATZ S. H.: Redrawing the map of great britain from a network of human interactions. *PloS one 5*, 12 (2010), e14248. 3

[TSF*16] THOMEE B., SHAMMA D. A., FRIEDLAND G., ELIZALDE B., NI K., POLAND D., BORTH D., LI L.-J.: Yfcc100m: The new data in multimedia research. *Communications of the ACM 59*, 2 (2016), 64–73. 2

[VdVJC11] VAN DE VOORDE T., JACQUET W., CANTERS F.: Mapping form and function in urban areas: An approach based on urban metrics and continuous impervious surface data. *Landscape and Urban Planning 102*, 3 (2011), 143–155. 2

[WREE67] WYLIE C., ROMNEY G., EVANS D., ERDAHL A.: Half-tone perspective drawings by computer. In *Proceedings of the November 14-16, 1967, fall joint computer conference* (1967), ACM, pp. 49–58. 3

[WZC*17] WU W., ZHENG Y., CAO N., ZENG H., NI B., QU H., NI L. M.: Mobiseg: Interactive region segmentation using heterogeneous mobility data. In *Pacific Visualization Symposium (PacificVis), 2017 IEEE* (2017), IEEE, pp. 91–100. 2

[YDGM17] YANG Y., DWYER T., GOODWIN S., MARRIOTT K.: Many-to-many geographically-embedded flow visualisation: an evaluation. *IEEE transactions on visualization and computer graphics 23*, 1 (2017), 411–420. 2

[YZX*15] YUAN N. J., ZHENG Y., XIE X., WANG Y., ZHENG K., XIONG H.: Discovering urban functional zones using latent activity trajectories. *IEEE Transactions on Knowledge and Data Engineering 27*, 3 (2015), 712–725. 2