**Paper Reproduce: Urban Pulse - Capturing the Rhythm of Cities**Final Project | sl9404 **|** 2021 Fall CS-GY 6313/CUSP-GX 6006 Information Visualization

GitHub: <http://losiyu.github.io/paper-reproduce-urban-pulse>Chart

Description automatically generated

**About Paper**

The goal of this work is to instead understand the city in the context of multiple urban data sets. They define the concept of an “urban pulse” which captures the spatio-temporal activity in a city across multiple temporal resolutions.

PDF Link: <https://fmiranda.me/publications/urban-pulse/tvcg-2017-urbanpulse.pdf>  
Demo: <https://github.com/VIDA-NYU/urban-pulse>

**Related Work**

There are many related visualization projects. I divided them with four categories, includes dataset, target, and technique. Urban-Pulse is using Flickr photo upload location and date. Similar media also include Sina Weibo [5] and Facebook [10]. Except for media data, government data is another good choice. US government transparency program published a lot of datasets from the department, which is valuable and reliable [2]. The second step is data analysis. "Beyond Transparency: Open Data and the Future of Civic Innovation" is a book that introduced the whole framework of using data [1]. The Shadbolt team shows how to search, categorize, and analyze government data [3]. The paper's target is to describe a new concept of the phenomenon and introduce a toolkit to detect it. A similar article includes people's movement [5] and human activity [4]. They Both created some new visualization methods to facilitate the detection and analysis. However, the heat map is still the best way to describe the activity flow using urban data. So, it is the essential plot in all of the dashboards. The calculation method is the core of the density map. Scheepens introduces MANet, which performs well in human activity and has a smaller MSE. Wang publishes a tool to visualize the traffic data. The Mesoscopic simulation models, which combine the properties of both macroscopic and microscopic simulation models, have high performance when visualizing high volume density data. Analysis of geometry Social Media Data Massive research has been done to exploit social media data to infer and predict people's spatial and temporal behaviors. For example, geotags are used to improve situation awareness in emergency response [11, 12] and disease control [13], and to understand cities [14, 15].

**Problems and Difficulties**

1. Find the data

[Flickr data](https://webscope.sandbox.yahoo.com/catalog.php?datatype=i&did=67), published by Yahoo, need a formal application to it. Then I tried to find the data set in their GitHub. I successfully find them in their [lab database](http://vgc.poly.edu/files/urban-pulse/data/), but because the data has been cleaned, we can only see 3 columns includes longitude, latitude, and timestamp and only in NYC and SF. The size of dataset is around 600,000. Another dataset I used are also in the lab database are pre-calculated critical point, both locations have around 100 points.

1. A picture containing graphical user interface

   Description automatically generatedDensity Function  
   Density map is one of the most important features in the visualization dashboard. The intensity of color was calculated by density function using Gaussian weighted sum.

However, I could not find the framework using the same function for render, so I choose [deck.gl heatmap layer](https://deck.gl/docs/api-reference/aggregation-layers/heatmap-layer) to visualize the human activity. Deck.gl use [Gaussian kernel function](https://en.wikipedia.org/wiki/Kernel_(statistics%29#Kernel_functions_in_common_use)) to calculate the color intensity. This is the most time-cost step. According to the paper It need 18 minutes to calculate if using a regular PC. Also, this is a high-pressure for website I made, 600K is nearly reach the limitation of deck.gl. So the website is stuck for a few second for rendering sometimes. Maxima Beats indicate

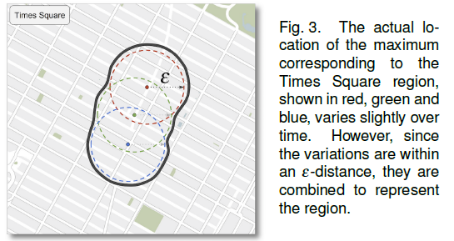
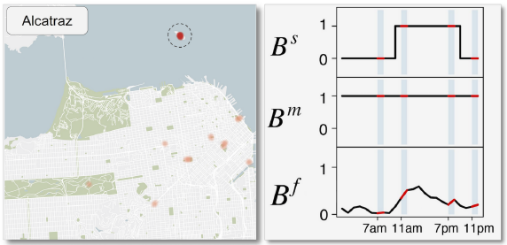
1. Topological persistence

The paper uses topological persistence to evaluate the urban pulse. The author purpose a new concept of persistence different from the original [16]. The new persist function pairs the global maximum with the global minimum. Also, it can still be computed efficiently in O(nlogn) time.

1. Urban Pulse Location

Urban Pulse Density function evaluates the human activity based on the mesh of 600K point; then, we use Euler distance to calculate each point’s pulse (Figure 1)`. The function also depends on how grouping the data. This is also a step of density function; even though the author has improved the efficiency, it still costs a lot of time to calculate the rank of the points.

1. Urban Pulse beats

 Capturing the urban pulse is the core concept of the paper. The question is how we can identify it after getting the activity curve over time. The author determines it based on significant beats, Maxima Beats, and Function Beats. Significant Beats are 0/1 sequence to indicate the absence/presence of a persistent high maximum at the location. Maxima Beats the absence /presence of a maximum at the location over the different time steps. These parameters record the activity around the significance beat. Function Beats stored the scalar function of variation of scalar function (Figure 2). The line chart is based on the functional beats, which show human activity. The green circle is based on the Significant Beats, which shows the urban pulse rhythm.

Figure

Figure

**Implementation**

1. Geo Map

Map

Description automatically generatedAlgorithm: density map with Gaussian weighted sum

Interaction: zoom, pan

Framework: Mapbox, Deck.gl

Code: <https://github.com/Losiyu/paper-reproduce-urban-pulse/blob/master/src/components/Map.js>

1. Chart, scatter chart

   Description automatically generatedRank Scatter plot

Algorithm: root mean square of significant, maximum and functional beat  
Interaction: regular and lasso select, zoom, pan  
Framework: Plotly.js  
Code: <https://github.com/Losiyu/paper-reproduce-urban-pulse/blob/master/src/components/Scatterplot.js>

1. Distance Chart

Chart, scatter chart

Description automatically generatedAlgorithm: Euclidean distance between two time series

Interaction: highlight, zoom

Framework: Plotly.js

Code: <https://github.com/Losiyu/paper-reproduce-urban-pulse/blob/master/src/components/DistanceChart.js>

1. Chart, line chart

   Description automatically generatedPulse Chart

Algorithm: maximum beat? (Significant beat? High: Low): 0

Interaction: zoom, select strong pulse time

Framework: Plotly.js, d3

Code: <https://github.com/Losiyu/paper-reproduce-urban-pulse/blob/master/src/components/PulseChart.js>

1. Banner and Modal

****

Graphical user interface, application

Description automatically generatedInteraction: select, drag

Framework: Material-UI

Code: 1. <https://github.com/Losiyu/paper-reproduce-urban-pulse/blob/master/src/components/Banner.js>

2. <https://github.com/Losiyu/paper-reproduce-urban-pulse/blob/master/src/components/Locations.js>

**Stories**

1. Focus on a part of critical points – Select The rest of plot will only show selected critical point.

Chart, scatter chart

Description automatically generated After

Map

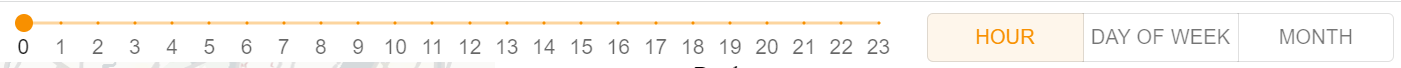
Description automatically generated A picture containing background pattern

Description automatically generatedChart, line chart

Description automatically generated Chart, scatter chart

Description automatically generated

1. Change the date – Select a different time or cluster method



Ex. All the plots will calculate critical point and density base on Day of week rather than hour

Map

Description automatically generatedChart, line chart

Description automatically generated Chart, scatter chart

Description automatically generated Chart, scatter chart

Description automatically generated

1. Highlight the location on the map and rank chart by – Hover on pulse chart

Chart, line chart

Description automatically generated Chart, scatter chart

Description automatically generated A picture containing diagram

Description automatically generated Chart, scatter chart

Description automatically generated

1. Select the location want to focus and show critical points only in an area – Select a location or an area

Graphical user interface, application, PowerPoint

Description automatically generated Graphical user interface, application

Description automatically generated

1. Filter the location with the pulse time we interest – select the date/weekday/month

Ex. Show locations has a strong pulse all the time from January to May.

Chart, line chart

Description automatically generated Chart, line chart

Description automatically generated

1. Find the similar urban pulse

Chart, scatter chart

Description automatically generated Chart

Description automatically generatedChart, scatter chart

Description automatically generated

**Works Cited**

[1] B. Goldstein and L. Dyson. Beyond Transparency: Open Data and the Future of Civic Innovation. Code for America Press, 2013.

[2] A. Noulas, S. Scellato, R. Lambiotte, M. Pontil, and C. Mascolo. A tale of many cities: universal patterns in human urban mobility. PloS one, 7(5):e37027, 2012.

[3] N. Shadbolt, K. O’Hara, T. Berners-Lee, N. Gibbins, H. Glaser, H. Wendy, and M. Schraefel. Linked Open Government Data: Lessons from Data.gov.uk. IEEE Intelligent Systems, 27(3):16–24, 2012.

[4] W. Wu, J. Xu, H. Zeng, Y. Zheng, H. Qu, B. Ni, M. Yuan, and L. M. Ni. Telcovis: Visual exploration of co-occurrence in urban human mobility based on telco data. IEEE TVCG, 22(1):935–944, 2016.

[5] S. Chen, X. Yuan, Z. Wang, C. Guo, J. Liang, Z. Wang, X. Zhang, and J. Zhang. Interactive visual discovering of movement patterns from sparsely sampled geo-tagged social media data. IEEE TVCG, 22(1):270–279, 2016.

[6] R. Scheepens, N. Willems, H. van de Wetering, G. Andrienko, N. Andrienko, and J. van Wijk. Composite density maps for multivariate trajectories. IEEE TVCG, 17(12):2518–2527, 2011.

[7] Z. Wang, M. Lu, X. Yuan, J. Zhang, and H. v. d. Wetering. Visual Traffic Jam Analysis Based on Trajectory Data. IEEE TVCG, 19(12):2159–2168, 2013.

[8] J. Poco, H. Doraiswamy, H. T. Vo, J. a. L. D. Comba, J. Freire, and C. T. Silva. Exploring traffic dynamics in urban environments using vector valued functions. CGF, 34(3):161–170, 2015.

[9] Y. Zheng, F. Liu, and H. Hsieh. U-air: when urban air quality inference meets big data. In Proc. KDD, pages 1436–1444, 2013.

[10] Jiří Hladík, Dajana Snopková, Marek Lichter, Lukáš Herman & Milan Konečný (2021) Spatial-temporal analysis of retail and services using Facebook Places data: a case study in Brno, Czech Republic, Annals of GIS, DOI: [10.1080/19475683.2021.1921846](https://doi.org/10.1080/19475683.2021.1921846)

[11] S. Vieweg, A. L. Hughes, K. Starbird, and L. Palen. Microblogging during two natural hazards events: what twitter may contribute to situational awareness. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 1079–1088. ACM, 2010.

[12] J. Yin, A. Lampert, M. Cameron, B. Robinson, and R. Power. Using social media to enhance emergency situation awareness. IEEE Intelligent Systems, 27(6):52–59, Nov. 2012.

[13] M.-H. Hwang, S. Wang, G. Cao, A. Padmanabhan, and Z. Zhang. Spatiotemporal transformation of social media geostreams: a case study of twitter for flu risk analysis. In Proceedings of the 4th ACM SIGSPATIAL International Workshop on GeoStreaming, pages 12–21. ACM, 2013.

[14] A. Vaccari, M. Martino, F. Rojas, and C. Ratti. Pulse of the city: Visualizing urban dynamics of special events. In Proceedings of 20th International Conference on Computer Graphis and Vision, 2010.

[15] C. Xia, R. Schwartz, K. Xie, A. Krebs, A. Langdon, J. Ting, and M. Naaman. Citybeat: Real-time social media visualization of hyper-local city data. In Proceedings of the companion publication of the 23rd international conference on World wide web companion, pages 167–170, 2014.

[16] T. Dey, K. Li, C. Luo, P. Ranjan, I. Safa, and Y. Wang. Persistent heat signature for pose-oblivious matching of incomplete models. CGF, 25:1545–1554, 2010.