Network Science-based Urban Forecast Dashboard*

Jiaxin Du
Department of Landscape Architecture & Urban Planning
Texas A&M University, College Station, Texas
Jiaxin.du@tamu.edu

Xinyue Ye*
Department of Landscape Architecture & Urban Planning
Texas A&M University, College Station, Texas
xinyue.ye@tamu.edu

Galen Newman
Department of Landscape Architecture & Urban Planning
Texas A&M University, College Station, Texas
gnewman@tamu.edu

David Retchless
Department of Marine and Coastal Environmental Science
Texas A&M University at Galveston, Galveston, Texas
retchled@tamug.edu

ABSTRACT

The urban environment is a highly dynamic and complex system. Urban dynamics in this complex system is largely reflected by the movement of people to and from Places of Interest (POIs) in the urban area. To better understand and plan for the city's various scenarios, there is a need to forecast urban dynamic conditions in terms of the possible movements of people across POIs. However, such predictions are not easy because an interdependent and living system is hard to forecast. In addition, the commuting and shopping of individuals in urban environments will show distinct patterns at various stages of disasters as compared to normal situations.

This paper presents a network science-based urban forecast dashboard, in order to monitor urban events and identify the interdependencies that characterize urban dynamics. Behind the dashboard is a deep learning model that incorporates the network dynamics between POIs. The dashboard powers the prediction of urban dynamics from a network science perspective. This research calls for a unified framework to model the flow and network in the city. The dashboard visualizes how network science and urban science can mutually benefit from each other.

KEYWORDS

Network Science, Urban Forecast, Dashboard, Events, Resilience

ACM Reference format:

Jiaxin Du, Xinyue Ye, Galen Newman, and David Retchless. 2022. Network Science-based Urban Forecast Dashboard. In 5th ACM SIGSPATIAL International Workshop on Advances in Resilient and Intelligent Cities (ARIC'22), November 1, 2022, Seattle, USA. ACM, New York, NY, USA, 4 pages. https://doi.org/XXXX

1 Introduction

The urban environment is a dynamic living system. People are constantly moving across the built environment to work, study, entertain, and engage in other activities. Rationale for travel decisions are often complex and diverse, and may be based on many factors, such as location, people's mood, and weather [1,2]. The movement of people makes cities alive. Understanding human mobility is the key to depict the fine-scale human-environment interaction. Modeling complex urban systems often requires abstraction and simulation [3,4]; the more data and details modeled, the better the prediction.

The ability to forecast human dynamics is desired in smart city applications [5–7]. For example, restaurants may hope to know food demands so they can purchase enough ingredients while creating less waste. Hotels may wish to adjust room prices based on predictions of the number of guests, and governments may need to be notified of over-crowded situations so they can issue appropriate restrictions. Yet it is a challenging problem for all of these situations as current forecasts are not yet easily attained.

People typically make their predictions about numbers of visitors based on experience. For instance, a grocery store can anticipate more crowds on a good weather weekend than on a rainy weekday. However, newly created prediction models offer an increasingly viable alternative to such experience-based inference [8]. For example, ARIMA (autoregressive integrated moving average) is a statistical model that can be used to predict the visit count based on historical data but also has several limitations: it can only be used for a single unique location and it ignores the place-based connections in the complex urban system. In other words, it assumes that the

Link to the dashboard: https://tamu.maps.arcgis.com/apps/dashboards/dd2722d7a7e440bc95f69ea716b12b8c
Link to the code:
https://github.com/UrbanDS/Network-Science-based-Urban-Forecast-Dashboard

visits to a store are independent from other stores and community environments. Instead, graph neural networks offer a more flexible alternative, since they can model place connections and characterize areas based on traffic flows; however, studies using this method have generally been conducted only at a specific time and are not easily updated.

Here, we present a forecast dashboard for modeling and predicting human mobility in the urban environment. A dashboard helps convey timely decision-relevant data to the public [9, 10]. It also provides public access to real-time information and acts as an interface for users to search based on their personal needs. During the Pandemic, one of the most well known dashboards has been the COVID-19 dashboard from John Hopkins University, which allows global users to view active COVID-19 cases counts, deaths, and recoveries, both in their community and around the world. Such a dashboard was initially designed for traffic or emergency management purposes [11, 12]. To our best knowledge, there is no dashboard that aims to predict human dynamics from a network science perspective.

Our web-based dashboard visualizes the past and predicts the future place visits (see the overview in section 2). The visualization and prediction are based on human trajectories data from sampled mobile phone records (see section 3) and the prediction model is built on graph neural networks (see section 4).

2 Dashboard Overview

Dashboards are web services that can be opened and interacted via multiple web browsers. They provide easy access to anyone who wants to explore a data-driven view of a topic of interest; e.g., in this case, human dynamics. A screenshot of our dashboard is presented in Figure 1.

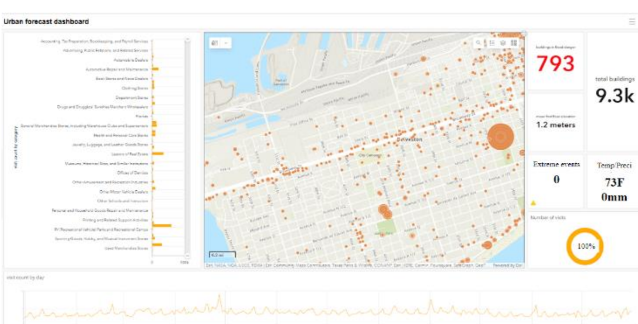

Figure 1: **Overview of the urban forecast dashboard**

2.1 Center Map

At the center of the dashboard is a map showing all the places of interests (POIs). Those places were defined by our data provider including restaurants, shops, parks and other facilities (see the data provider website in the next section for details). Here, we zoom into a random place as a showcase (Galveston TX, USA). The yellow dots on the map indicate the POI locations, while the dot sizes show the number of people who visited the POI during 2019. The size of the point is proportional to the visit count. The larger the dot, the more people visited that POI in that particular time period. As the time period and map scale change, this magnitude scale is repeatedly readjusted.

Users can click on the POI to see detailed information, such as the POI name, brand, and category, as Figure 2 shows.

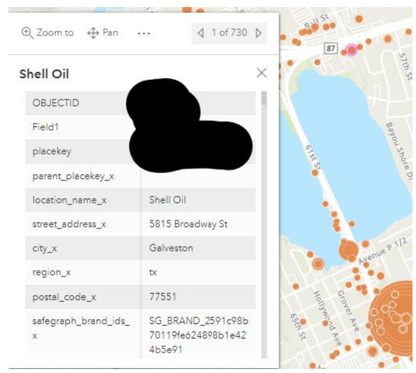

Figure 2: **POI detailed information window (sensitive information is redacted)**

At the top left of the map is a tool used to select elements in the map. If multiple POIs are chosen, the left, right, and bottom panels change to show the aggregated information on these POIs instead of the map view. This allows the user to compare groups of POI features. The top right of the map offers several useful tools, such as place search, map legends, and toggles for visible layers and basemaps. With more data added to the dashboard, users would be able to explore multiple relevant layers for a variety of purposes. We currently have weather watches and warnings, wind station data, etc. These weather data provide some context for human mobility. Users can visualize weather (or other) data with the visit count data to understand conditions during the time period of interest, allowing them to form hypotheses about possible relationships between weather (or other contextual layers) and the human mobility patterns shown.

2.2 Surrounding panels

On the right panel, there are a few widgets showing statistics for the current view, including the number of POIs, the number of visits in that period, weather information and forecasts, embedded news alerts, and other related information. This content changes if the user zooms or moves the map frame, so that it always presents summary statistics for the data in the map view. (As described above, the one exception is when multiple POIs are selected, in which case the surrounding panels show the aggregated information on these POIs instead of a summary for the map view.)

The left side panel shows the number of visits by POI category. This panel also adjusts with the map view, so it only shows the information within the current map view.

At the bottom of the map is the line plot of the visit count in the map view. The x-axis shows the time, and the y-axis shows the number of visits. Due to data license limitations, we cannot show the raw visit count for each POI. Instead, the y-axis shows the percentage fluctuation of the visit count, with 100% meaning the historical mean average, zero meaning no visits, 200% meaning twice as many people as the mean average, and so on.

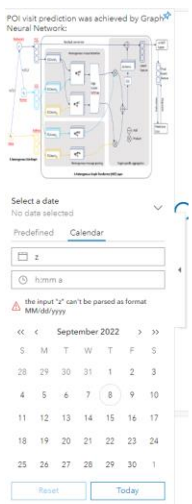

Figure 3: **Expand panel with time selection**

Users can expand a hidden left panel to select the desired time. At the top of this panel is the introduction of the dashboard. The bottom of the dashboard has a time selection tool. Users can choose a time period for which they wish to know visit counts. This panel allows any dates from the past to the future. However, if the relevant data are missing in our database, the user will receive an error message and will be unable to see that period on the map.

2.3 Other technical details

The dashboard is built on an ArcGIS Dashboard. The data are hosted on the ESRI cloud service. For forecasting, we built and ran our model on a private server. The details of the forecasting model will be presented later.

3 Data Sourcing

The human mobility data utilized came from Safegraph (www.safegraph.com). Safegraph sampled mobile phone trajectories and aggregated them to form their data product. The product includes visit count data for POIs starting from 2018. The data coverage is now worldwide, and is updated monthly.

We normalized the data because the data license prohibits us from sharing the raw data and due to privacy concerns. The raw visit data show the number of visits to each POI daily. We normalized the data by calculating the daily visit count for each POI over the mean of the entire period of the record, denoting that as one. Then we used a percentage to show the visit count, with 0.8 as 80% of the mean and 2 denoting 200% of the mean. It can be denoted as equation (1)

$$normalized\ visit\ count = \frac{raw\ visit\ count}{mean\ visit\ count} \qquad (1)$$

We assume that the POI owners would know how many people they hosted daily, so showing the normalized visit count would be as informative as showing the raw data.

4 Forecasting Models

In the backend of the system, we developed several forecasting models. Initially, we used the ARIMA model for each POI. It provides a strong baseline for each POI. However, as noted, it cannot model the connections in the urban environment. Therefore, we would not know how the POIs influence each other, e.g., whether they compete for customers or corporate to draw visitors. This requires a network science view that captures the relationships between POIs and the urban environment. The forecasting model was developed based on graph neural networks. The graph neural networks are deep learning models designed for graph data.

Network. A network consists of two major components {Nodes, Edges}. The nodes model the study objects and the edges model their connections. There are many alternatives to modeling the city as a graph [13], [14]. We treat the POIs as nodes, with edges between nodes created if the Euclidean distance between two POI is less than 5 kilometers [15].

Deep learning model. Deep learning models contain large numbers of learnable parameters, so they can fit the data very well. We chose graph neural networks as they were specifically designed for graph data. We chose the graph convolutional network utilized in much current literature [14].

The training separates everyday situations from disaster situations. We retrieved the extreme weather information from the National Oceanic and Atmospheric Administration (www.noaa.gov). Then we divided the data into two sets -- normal and abnormal. The normal period contains days without significant weather events. The abnormal period contains the extreme event days and the period one-week before the event. By doing so, we have two separate models for predicting daily visit counts, enabling forecasts under normal conditions and for extreme events.

5 Discussion

There are certain limitations to the current urban forecast dashboard.

1. The human mobility data are sampled from mobile phone users. This can result in multiple sampling biases. The people who do not have mobile phones are ignored in such data. Also, even for people who have a phone, the GPS locations might not be accurate.
2. The prediction accuracy is not perfect. We are, however, developing more robust, highly calibrated, and accurate models for prediction. There are many model alternatives, and we will continue to update models on the backend.
3. The network behind the dashboard is relatively simple currently. It only consists of the POIs and their spatial relationships. There are opportunities to expand the network, including by using the demographic background of the POIs and the visitors. We plan to incorporate more data into the network to model more urban complexities.

6 Conclusion

We demonstrate an urban forecast dashboard in this paper. This dashboard visualizes past and future urban dynamics by presenting visitor counts using POIs. Multiple panels and widgets around the map help users navigate the visit count data. The data were sampled from the mobile phones and associated with extreme events and forecasts data. Our forecast models were built using a network science perspective. This connects the POIs and predicts the future visits based on the network. Multiple scenarios were considered, so the model can differentiate between everyday and disaster conditions. We are fine tuning the prediction models to improve their accuracy and efficiency.

There is a to-do list for the urban forecast dashboard. We can link more data in our backend network, so we can present more information. Such information can include visitor demographics, modes of travel (public bus, car, etc.), and other related attributes. These can raise privacy and ethical concerns, and we would like to tackle them in the future.

Because we train our model on several urban events and contexts in the backend, the dashboard can predict human dynamics under multiple scenarios. This dashboard may benefit business management, resilience planning, and emergency response. We will collect the feedback from users to refine the tool interface and functionality.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant Nos. 2112356 and 2122054 as well as the start-up grant 241117 from Texas A&M University. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation and Texas A&M University. We also acknowledge that safegraph provides the data.

REFERENCES

[1] M. Schläpfer *et al.*, “The universal visitation law of human mobility,” *Nature*, vol. 593, no. 7860, pp. 522–527, May 2021, doi: 10.1038/s41586-021-03480-9.

[2] W. Li, Q. Wang, Y. Liu, M. L. Small, and J. Gao, “A spatiotemporal decay model of human mobility when facing large-scale crises” Proceedings of the National Academy of Sciences, vol. 119, no. 33, p. e2203042119, Aug. 2022, doi: 10.1073/pnas.2203042119.

[3] M. Batty and P. A. Longley, “The fractal simulation of urban structure,” *Environ. Plan. A*, vol. 18, no. 9, pp. 1143–1179, 1986.

[4] R. Lempert *et al.*, “Meeting Climate, Mobility, and Equity Goals in Transportation Planning Under Wide-Ranging Scenarios,” *J. Am. Plann. Assoc.*, vol. 86, no. 3, pp. 311–323, Jul. 2020, doi: 10.1080/01944363.2020.1727766.

[5] J. Wu, K. Leung, and G. Leung, “Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study,” *Lancet*, Apr. 2020, doi: 10.1016/S0140-6736(20)30260-9.

[6] M. Chen, Y. Liu, and X. Yu, “NLPMM: A next location predictor with Markov modeling,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2014, vol. 8444 LNAI, no. PART 2, pp. 186–197. doi: 10.1007/978-3-319-06605-9_16.

[7] S. Feng, G. Cong, B. An, and Y. M. Chee, “Poi2vec: Geographical latent representation for predicting future visitors,” 2017.

[8] S. Mashi, A. Inkani, O. Obaro, and A. Asanarimam, “Community perception, response and adaptation strategies towards flood risk in a traditional African city,” *Nat. HAZARDS*, vol. 103, no. 2, pp. 1727–1759, 2020, doi: 10.1007/s11069-020-04052-2.

[9] E. Dong, H. Du, and L. Gardner, “An interactive web-based dashboard to track COVID-19 in real time,” *Lancet Infect. Dis.*, vol. 20, no. 5, pp. 533–534, 2020.

[10] X. Ye, J. Du, X. Gong, S. Na, W. Li, and S. Kudva, “Geospatial and Semantic Mapping Platform for Massive COVID-19 Scientific Publication Search,” *J. Geovisualization Spat. Anal.*, vol. 5, no. 1, p. 5, Jan. 2021, doi: 10.1007/s41651-021-00073-y.

[11] M. Zhao and X. Liu, “Development of decision support tool for optimizing urban emergency rescue facility locations to improve humanitarian logistics management,” *Saf. Sci.*, vol. 102, pp. 110–117, 2018, doi: 10.1016/j.ssci.2017.10.007.

[12] Y. Zhao, T. Eynon, F. Kamw, and D. Sheets, “[Workshop] TrajAnalytics : A Web-Based Visual Analytics Software of Urban Trajectory Data,” 2016. [Online]. Available: http://geoanalytics.net/and/.

[13] B. Yan, K. Janowicz, G. Mai, and S. Gao, “From ITDL to Place2Vec: Reasoning About Place Type Similarity and Relatedness by Learning Embeddings From Augmented Spatial Contexts,” in *Proceedings of the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, New York, NY, USA, Nov. 2017, pp. 1–10. doi: 10.1145/3139958.3140054.

[14] D. Zhu *et al.*, “Understanding Place Characteristics in Geographic Contexts through Graph Convolutional Neural Networks,” *Ann. Am. Assoc. Geogr.*, vol. 110, no. 2, pp. 408–420, Mar. 2020, doi: 10.1080/24694452.2019.1694403.

[15] G. Mai, K. Janowicz, B. Yan, R. Zhu, L. Cai, and N. Lao, “Multi-Scale Representation Learning for Spatial Feature Distributions using Grid Cells,” 2020. [Online]. Available: https://openreview.net/forum?id=rJljdh4KDH