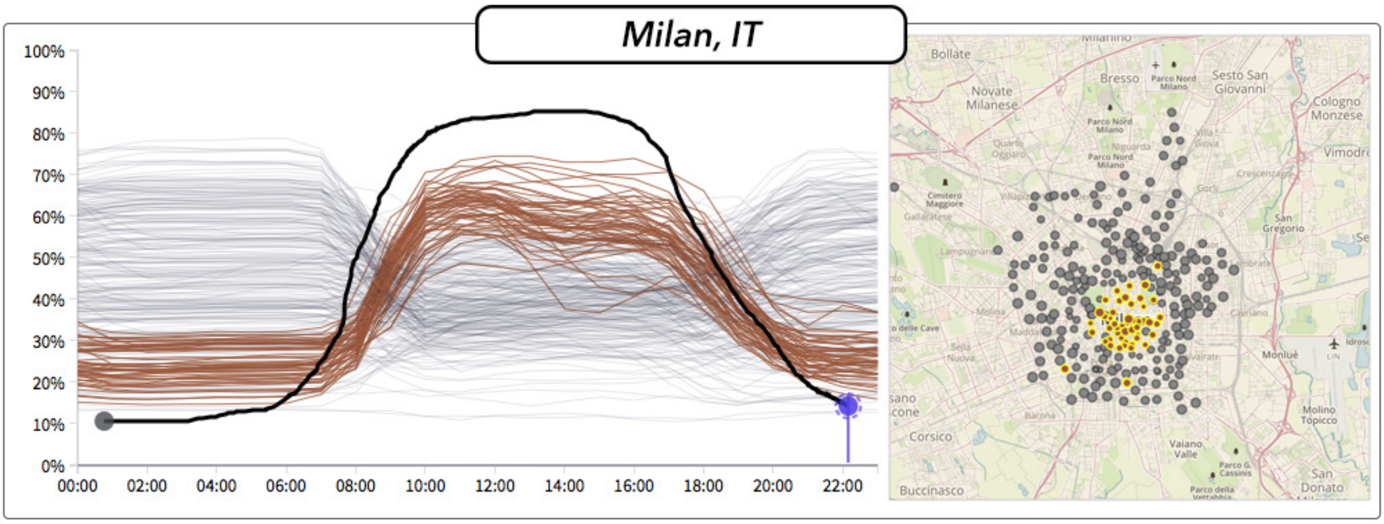

See Figure 1.

Michael Oppermann[1] (opperman@cs.ubc.ca) & Tamara Munzner[1] (tmm@cs.ubc.ca)
[1] University of British Columbia

UNCOVERING SPATIOTEMPORAL DYNAMICS FROM NON-TRAJECTORY DATA

Keywords: non-trajectory data; spatiotemporal patterns; visual exploration; bike sharing; space utilization

1. DESIGN PROBLEM

In a previous project [5], we analyzed and compared hundreds of bike sharing networks worldwide, predominantly based on station fill levels that we recorded continuously over a period of 17 months. We illustrated how we can support users with a wide range of expertise to understand and intelligently leverage this type of data in their decision-making. Our interactive visualization can reveal interesting insights, not only into patterns of bicycle usage but also into underlying spatiotemporal dynamics of a city (see Figure 1).

We have begun a new project that is focused on creating visual and predictive decision-support tools centered around building occupancy data. Previously this data has been used for the automatic control of heating, ventilation and air conditioning (HVAC) systems and now, by following the design study methodology [6], we are opening it up to a broader set of stakeholders in facility management. Initial experiments indicated that making this data accessible and visually explorable can lead to a better understanding how space is actually being used and will ultimately improve space utilization and resource management.

The intriguing underlying similarity between these projects lies in the data characteristics. We are using status changes at distinct locations (non-trajectory), such as the number of available bikes at a docking station or the number of people occupying a certain room, to investigate spatiotemporal patterns. In contrast, much of the previous work has been focused on individual movements (trajectory) or on origin-destination (OD) data.
By noting the similarity in the data, we can take what we learned in both projects to discuss general implications of spatiotemporal non-trajectory data in terms of ethics, data preprocessing, tasks, and visual encodings. Our goal

is to generalize our findings in the context of urban data visualization with the hope to inspire other researchers and designers.

2. IMPLICATIONS FOR ETHICS

Increasingly digitized cities and massive accumulations of data, often containing geo-tagged personal information, pose risks and raise privacy concerns. Especially fine-grained trajectories can disclose sensitive information, for example, a person's home, workplace and daily commuting pattern. Anonymization and aggregation can only help to a certain degree if the number of trajectories is small. Although there are several methods to address this issue, such as spatiotemporal generalization [3], we should reconsider if the storage and analysis of trajectories is necessary in the first place. We have experienced many barriers and delays when dealing with raw trajectory data and this trend will be reinforced as organizations are implementing the General Data Protection Regulation (GDPR).

On the other hand, generalized and location-based counts are often willingly shared or can be easily accessed in open data libraries.

3. IMPLICATIONS FOR DATA PREPROCESSING

The raw data contains meta information (e.g. geographic location) and a corresponding spatial time series for each sensor. The number of sensors can vary from a few dozens to many thousands. Typically, the sensor states are recorded every few minutes but this can be further shortened to generate more detailed time series.

In our bike sharing project we collected and preprocessed live data from more than 20.000 docking stations in a 15-minute interval. In general, if you have enough knowledge about the data, we suspect that rather than using a uniform interval, you can distinguish between day and night times and adjust the interval accordingly.

4. IMPLICATIONS FOR TASKS

Clearly, the analysis of movement flows is more difficult, if not even impossible, without trajectories or OD data but other tasks are very well supported indeed.

The task of investigating temporal footprints of locations (sensor level) or whole regions (i.e. district aggregation level) is often in the center of the analysis. We can compare sensors at different time resolutions and identify trends, outliers and repeating patterns. Existing approaches are often focused on daily, weekly, and monthly resolutions and lack the ability to capture seasonality. In our latest project, we recognize a strong need from multiple stakeholders for inspecting the data at custom time intervals, for instance, to analyze the utilization of meeting rooms on Mondays between 9am and 5pm. High-level overviews are necessary but they are not sufficient. In addition, users want to see the original time series for a specific location or compare multiple time series on a detailed level. A benefit of this data type is, for the task of capturing local and global variations, that it is much easier to group sensors by regions and to normalize the time series.

Back to the bike sharing example, we realized that the balancing of docking station fill levels poses the biggest challenge for operators. Individual routes taken by customers are usually negligible but operators need to know when a station gets full or empty or what the ideal distance between stations is. They want to understand the distribution of sources and sinks and if behaviors are limited to specific stations or if they are affecting adjacent neighborhoods. We noted many specific questions in this project that could be well addressed without requiring trajectories.

5. IMPLICATIONS FOR VISUAL ENCODINGS

Non-trajectory data can be visually encoded with the same techniques that have been proposed for visualizing time-oriented data. In particular, small multiples, interactive linked views or superimposed perspectives allow us to capture both spatial and temporal dimensions.

In our bike sharing system, we have used an interaction technique to visualize flows indirectly (see Figure 1). Users can draw a custom fill level curve on top of the multi-series line chart and stations that follow a similar pattern are highlighted in the linked map view. By drawing multiple curves and observing the change of active stations, users can get a quick intuition about commuting patterns.

Wood et al. [7] used geographic small multiples to visualize docking stations of London's bike sharing scheme at different points in time. Miranda et al. [2] proposed a multi-view pulse monitor to inspect Flickr activity data. Morphocode [4] discussed different visual encodings for visualizing pedestrian counts in an online blog post.

The obvious common method that people choose when working with trajectories is to display flows as a superimposed layer on a geographic map. This may seem aesthetically pleasing, but the result is often cluttered and difficult to analyze. Although many techniques have been proposed to mitigate this problem [1], we propose an alternative, a data transformation to location/sensor-based counts which would then afford other types of visual encodings.

REFERENCES

[1] W.Aigner,S.Miksch,H.Schumann,andC.Tominski. Visualizationof time-oriented data. Springer Science & Business Media, 2011.

[2] F.Miranda,H.Doraiswamy,M.Lage,K.Zhao,B.Gon ̈ıalves,L.Wilson, M. Hsieh, and C. T. Silva. Urban Pulse: Capturing the Rhythm of Cities. IEEE Trans. Visualization and Computer Graphics, 23(1):791– 800, 2017.

[3] A. Monreale, G. Andrienko, N. Andrienko, F. Giannotti, D. Pedreschi, S. Rinzivillo, and S. Wrobel. Movement data anonymity through generalization. Trans. Data Privacy, 3(2):91–121, 2010.

[4] Morphocode. Location + time: urban data visualization. https://morphocode.com/location-time-urban-data-visualization/. Accessed: 2018-06-11.

[5] M.Oppermann,T.Mo ̈ller,andM.Sedlmair.BikeSharingAtlas:Visual Analysis of Bike-Sharing Networks and Urban Commuting Patterns Worldwide. Int. Journal of Transportation, 6(1):1–14, 2018.

[6] M. Sedlmair, M. Meyer, and T. Munzner. Design Study Methodolgy: Reflections from the Trenches and the Stacks. IEEE Trans. Visualization and Computer Graphics, 1(12):2431–2440, 2012.

[7] J.Wood,A.Slingsby,andJ.Dykes.Visualizingthedynamicsof London's bicycle-hire scheme. Cartographica: Int. Journal for Geographic Information and Geovisualization, 46:239–251, 2011.

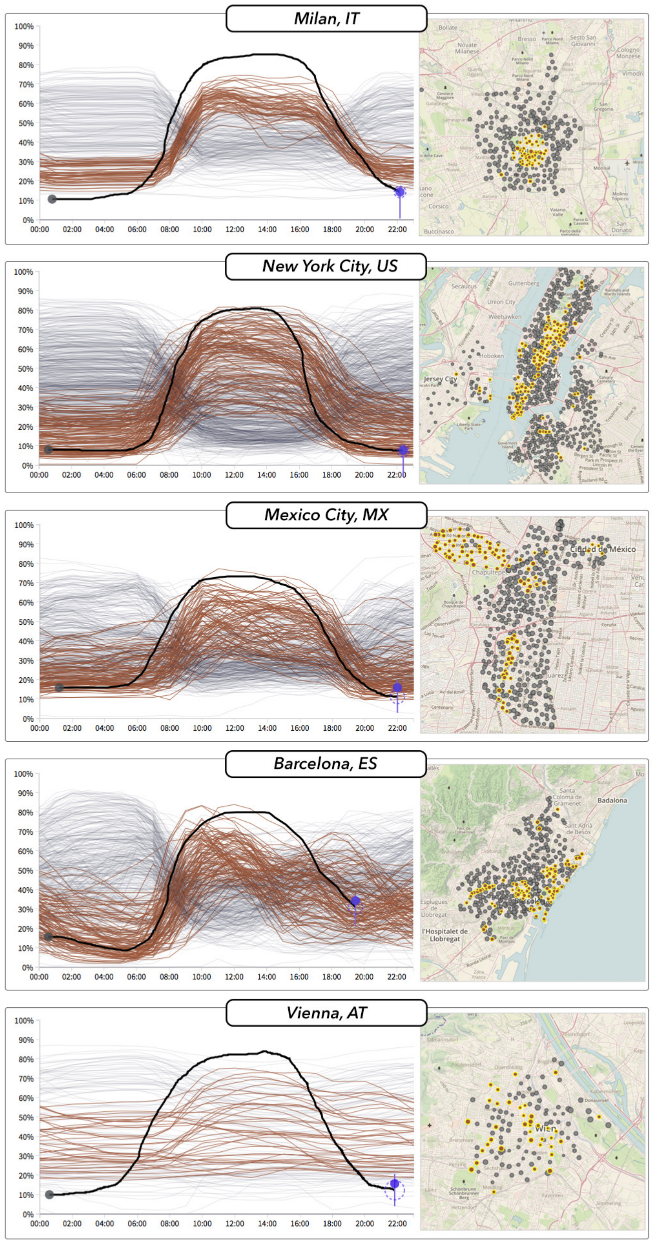

Figure 1: The visualization of average docking station fill levels exposes bike-sharing commuting behaviors in different cities around the world [5]. Each line represents an individual bike sharing station. Users can draw a line—a hypothetical profile—on top of the chart to select similar stations.