Analysis of Mobility Behaviors in Geographic and Semantic Spaces

Natalia Andrienko[1,2,3], Gennady Andrienko[1,2,3], Georg Fuchs[1,3]

[1] Fraunhofer Institute IAIS; [2] City University London; [3] University of Bonn

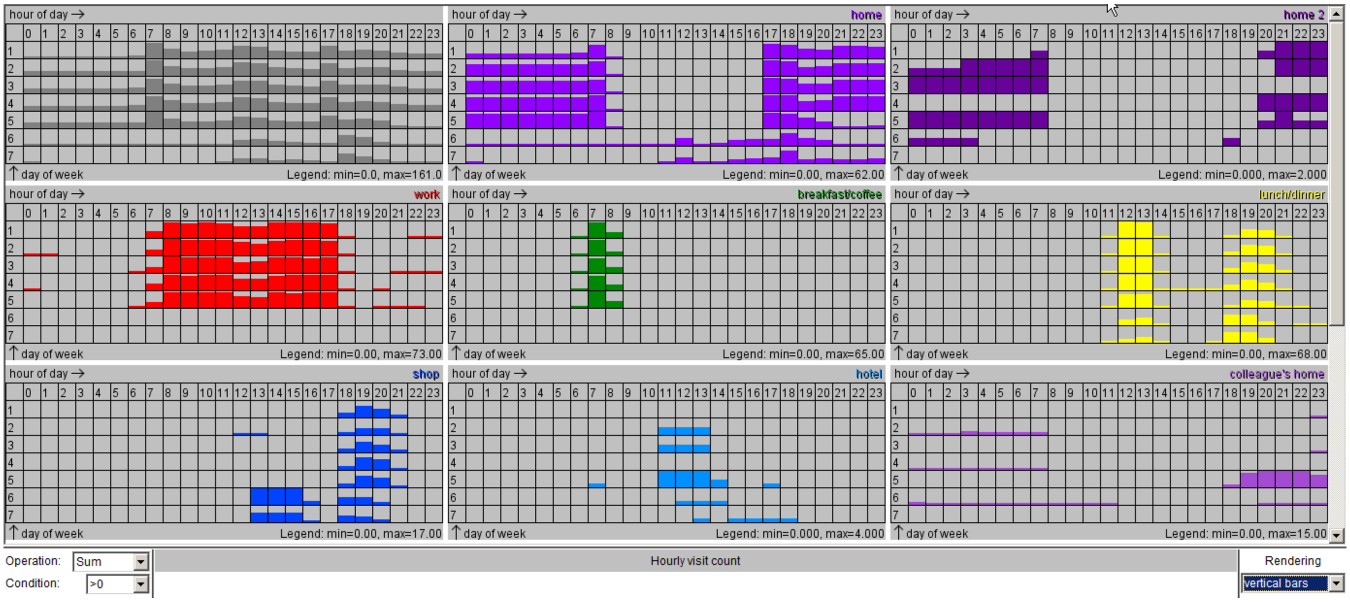

Figure 1. 2D time histograms represent the total counts of visits to different place categories by hourly intervals in the weekly cycle.

Abstract

Repeatedly visited personal and public places were extracted from trajectories by finding spatial clusters of stop points. Temporal patterns of people's presence in the places resulted from spatio-temporal aggregation of the data by the places and hourly intervals within the weekly cycle. Based on these patterns, we identified the meanings or purposes of the places: home, work, breakfast or coffee, lunch and dinner, and dinner or shopping. Meanings of some places could be refined using the credit card transaction data. By representing the place meanings as points on a 2D plane, we built an abstract semantic space and transformed the original trajectories to trajectories in the semantic space. Spatio-temporal aggregation of the transformed trajectories into flows between the semantic places and subsequent clustering of time intervals by the similarity of the flow situations allowed us to reveal the routine movement behaviors. To detect anomalies, we (a) investigated the visits to the places with unknown meanings, and (b) looked for unusual presence times or visit durations at different semantic places. The analysis is scalable since all tools and methods can be applied to much larger data.

1 Extraction and Interpretation of Places

We used an automated tool that extracts repeatedly visited personal and public places by spatial clustering of points from trajectories. Relevant points can be previously selected by interactive filtering. We selected the points of stopping for at least one minute. The tool's work is based on finding groups of points fitting in circles with a chosen maximal radius and uniting close groups. We chose a sufficiently big maximal radius (100 m) to account for the noise in the data. Personal places are extracted separately from the trajectory of each person. Public places visited by at least a given minimal number of distinct persons (2 in our analysis) are extracted from all trajectories together.

Through spatio-temporal aggregation of the trajectories [1], we obtained the visit counts for the extracted places by hourly intervals within the weekly time cycle. For the personal places, only the visits of the place owners were counted. We analyzed the temporal distributions of the place visits using 2D time histograms (Fig. 1), where the rows correspond to 7 days the week, columns to 24 hours of the day, and marks in the cells represent aggregated visit counts for the whole set or subsets of places. By clustering the places according to similarity of their visit distributions, we found groups of places with prominent temporal patterns of visits, which could be attributed to certain categories of places or activities: home, work, breakfast or coffee, lunch, lunch and dinner. We combined card transaction data with the extracted stop points to (1) determine the geographic locations of the businesses and (2) among the places visited in the evenings and on the weekend, distinguish places for eating, shopping, sport, etc.

We could not determine the meanings of 5 public places visited mostly in hour 11 of the week days. The temporal pattern did not hint at any usual people's activity, and no card transaction records could be matched with the place visit times. We found that these places were attended by particular people and gave them a label "BFMO place", where BFMO consists of the initials of the last names of these people.

2 Analysis of Routine Behaviors in Semantic Space

After the assignment of the semantic labels to the places, we transformed the original trajectories into sequences of place visit records, each record containing the semantic label of the visited place and the start and end times of the visit. The intermediate trajectory points between the place visits were omitted.

We created an abstract semantic space where the semantic categories of places are represented as points on a 2D plane; we call them "semantic places". Then, we transformed the sequences of place visit records to trajectories in the semantic space. For this purpose, the place visit records were complemented with the coordinates of the semantic places. Spatio-temporal aggregation can be applied to trajectories in abstract spaces in the same way as to trajectories in geographic space [1]. We aggregated the transformed trajectories into flows (aggregate moves) between the semantic places by hourly intervals. Then we clustered the intervals by similarity of the respective flow situations, i.e., vectors composed of the magnitudes of the flows for all ordered pairs of semantic places. We applied k-means clustering algorithm using Manhattan distance between the vectors as the similarity measure. In the calendar display (bottom center of Fig. 2), pixels representing the hourly intervals are colored by their cluster membership. Periodic patterns with regard to the daily and weekly time cycles could be observed for different values of the parameter k (number of clusters). In the small multiple maps in Fig. 2, the average hourly flows for the time clusters are represented by the widths of the flow symbols (curved lines with the curvature increasing towards the destination).

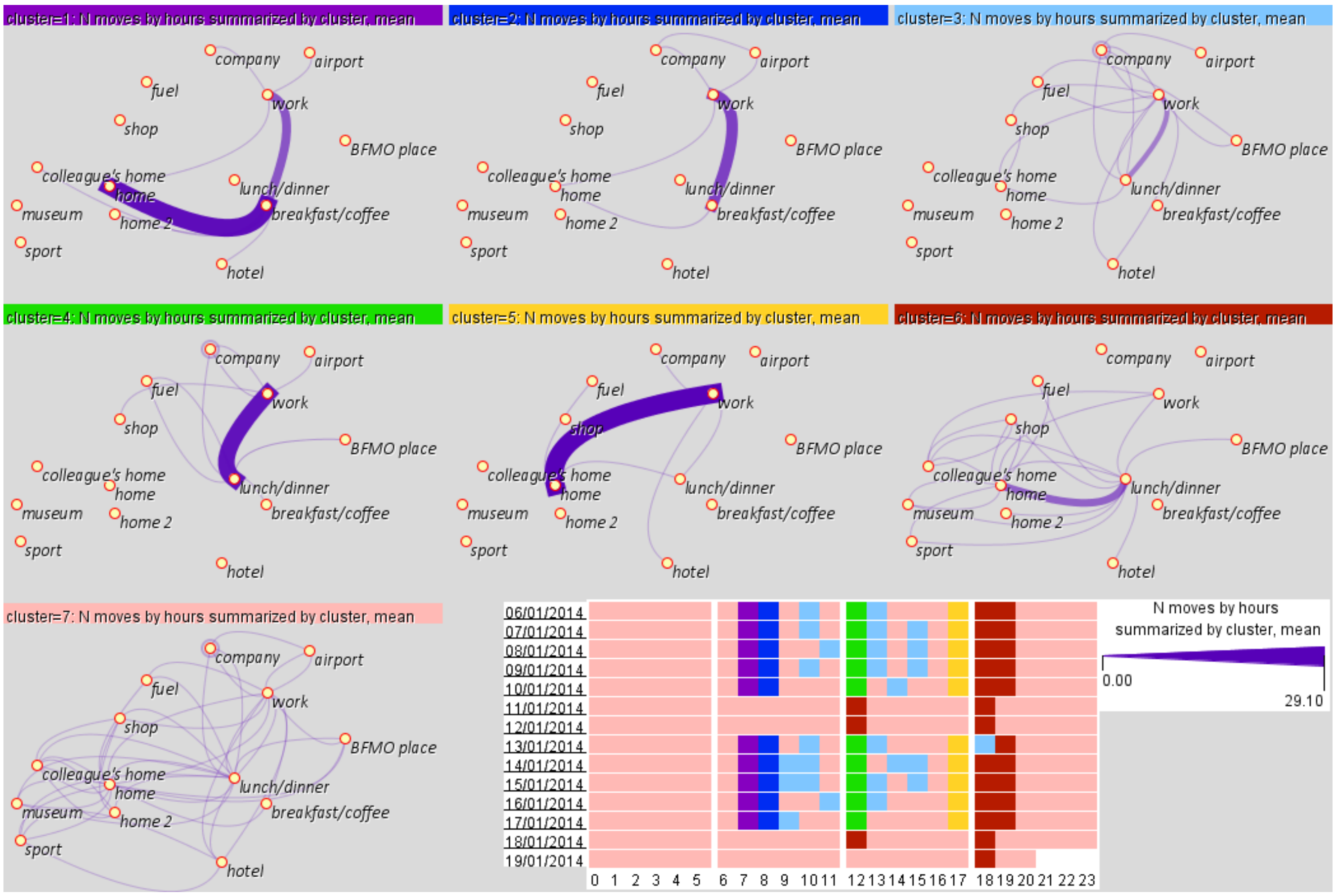

Figure 2.Clustering of hourly intervals by the similarity of the flows in the semantic space reveals high regularity of the movements.

3 Detection and Analysis of Anomalies

By interacting with the map display of the semantic space and transformed trajectories, we selected the daily trajectories visiting BFMO places (explained in section 1) and saw that people went to these places from the work place. From BFMO places, the visitors almost always moved to lunch/dinner places and then returned to the work place. Hence, BFMO places were not visited for having lunch. By extracting the corresponding place visits records by means of spatial filtering, we examined who was when in which place and found 5 cases when two or three people met in the same place. All four visitors of BFMO places were security employees.

For each semantic place, we analyzed the temporal distribution of the visits using a 2D time histogram similar to those shown in Fig.1 but with 14 rows corresponding to the consecutive days of the two week period of the data (as in the calendar display in Fig. 2). We paid attention to place visits in unusual times. For each such case, we extracted (by spatio-temporal filtering) the place visit records including the visitors' names, exact visit times, and place names. In this way, we detected that four persons sometimes attended the homes of their colleagues in night time. We also detected night visits to work and some other anomalies.

4 Conclusion

All methods and tools that we used for our analysis are scalable with regard to the number of individuals, number of places, and length of the time period covered by the data. We mostly used aggregated views, which could also be applied to much larger data. Detailed data (place visit records) were accessed only for analyzing anomalies. The analysis greatly relied on computational data processing: stop and place extraction, data aggregation, and clustering. These operations are also scalable. Apart from the examination of the anomalies, the analysis was done in a way respecting personal privacy, without accessing personal data.

References

[1] Andrienko, N., Andrienko, G.: Spatial generalization and aggregation of massive movement data. *IEEE Trans. Visualization and Computer Graphics*, 17(2): 205-219, 2011.