Behavior Visual Analysis Supporting Spatial temporal Data with Uncertainty

Siming Chen Zuchao Wang Xiaoru Yuan

Key Laboratory of Machine Perception (Minister of Education), and school of EECS, Peking University, Bejing, China

Introduction

We present a case study for uncertainty handling in human behavior visual analytics. Our data from IEEE VAST Challenge 2014 concerns integration of GPS trajectory logs together with a few other location related properties. The data have many errors, data missing and conflicts. In this work, we identify and fix several kinds of uncertainties in the dataset.

Methodology

1) Data-driven approach: We use the distribution of the data points, to further improve the automatic calculation of the POI position.
2) Cross reference of heterogeneous data set: By investigating the shared attributes from different data source, we can mark different level of uncertainty and reduce uncertainty in a reasonable level.
3) Combined visual exploration with automatic calculation: Together with automatic calculation and batch processing, we also use the visual interface to investigate the uncertain data, which provides the feedback loop to further improve the data quality.

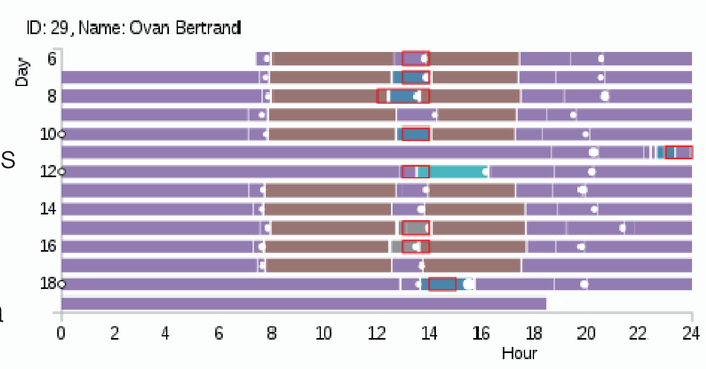
Overview

Uncertainty Type	Dealing with Uncertainty	Map	Trajectory Data	Transaction data (low resolution)	Transaction data (high resolution)	Description
POI information	POI region range extraction	√				Extract a range with uncertainty from the JPEG map
	POI region range refinement	√	√			Based on common pattern extracted from the trajectory data
Temporal information	Resolution refinement			√	√	Mark the detail time based on the matched event with high resolution
	Transaction time error and conflict		√	√	√	Mark the conflict time for each transaction events
Transaction attribute	Data missing			√	√	Add and mark the missing records
	Price conflict			√	√	Mark two values
Location information	Location conflict		√	√	√	Mark the conflict location
	Location missing, shift and noise	√	√	√	√	Mark the 'jump' position and shift based on common pattern and POI, using the transaction for verification
People information	Missing car assignment		√	√	√	Find people who have the matched transaction and location of the car

Event Timeline

Event Definition

We combine different datasets with the concept of movement event. The events are first defined from the GPS logs, as stops above 1 minute. Each event is naturally associated with a car, a time span and a location.

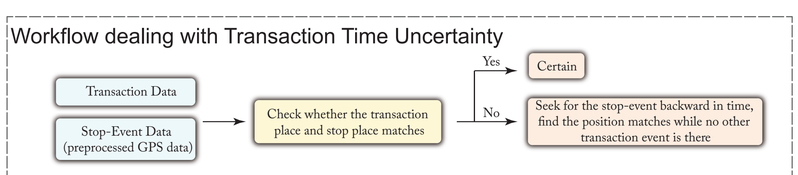


Visual Mapping

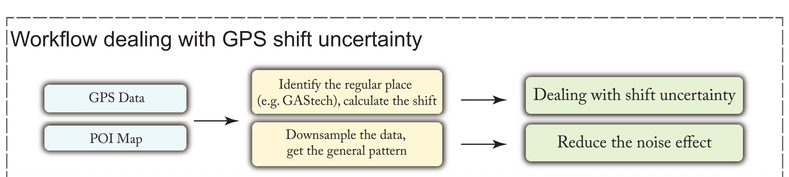
The fused event data can be visualized with a event. The X axis represents hour of day, while the Y axis represents each day. Each event is represented as a rectangle, with the color showing POI categories, and position showing its start and end times. White circle indicates the transaction record, the size of which is the price paid. Red rectangle indicates the result of outlier detection methods, to help people analyze the data.

Uncertainty Processing

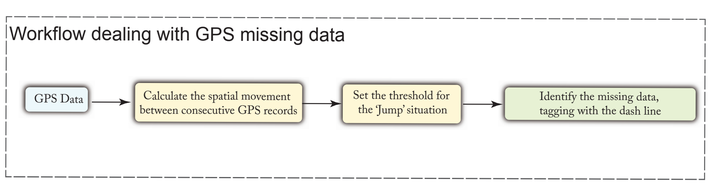
Time Uncertainty

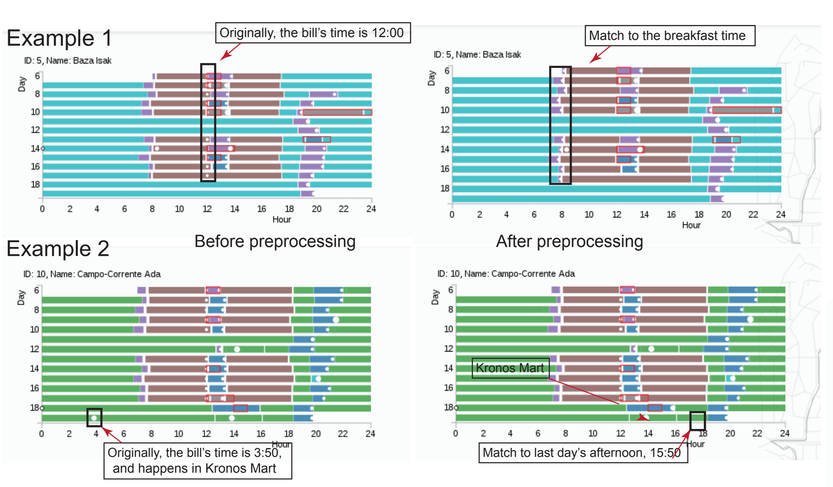


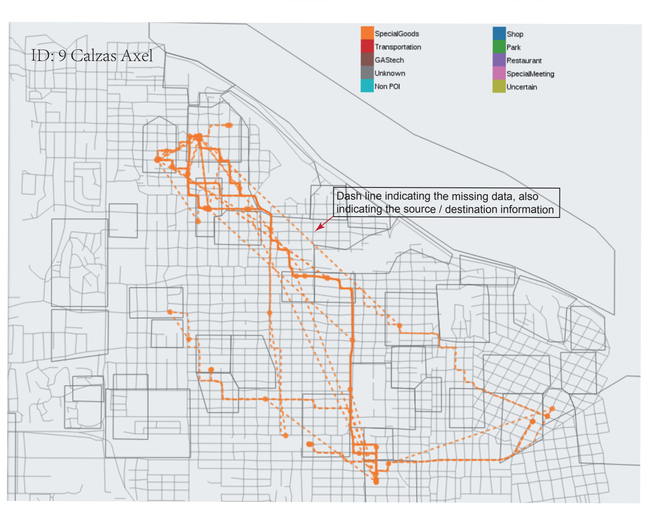
GPS Shift



Missing Data



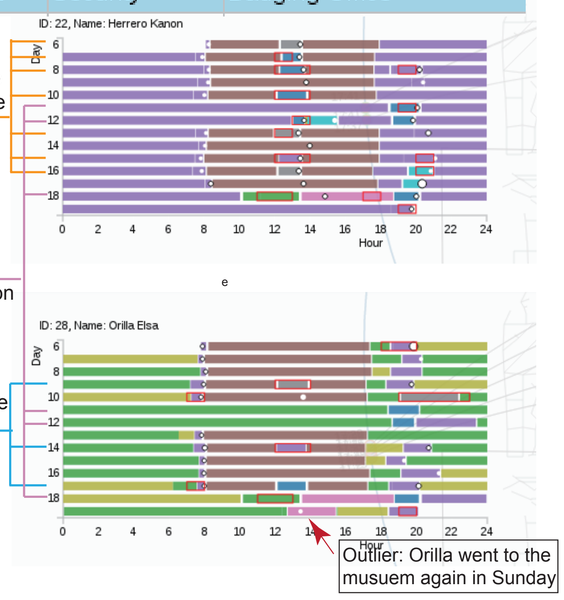




Behavior Analysis based on Processed Data

Correlation Pattern

Orilla	Elsa	28	Engineering	Drill Technician
Herrero	Kanon	22	Security	Badging Office



Exploration Step

The case is done based on the uncertainty processing, the exploration step in al is as follows.
Step 1: Visualize the movement in the spatial view, and detect the shift and errors. And also find outliers in the event timeline view.
Step 2: Find some places that have high ratio of apperance, and compare with the corresponding routes in the same time of other people.
Step 3: Calculate the shift values based on the movement on places of interest. Shift back the routes. Use downsampling to reduce the noises.
Step 4: Verify the shifted routes with the event timeline view, which matches the regular patterns.
Step 5: Identify some places of interest this people usually go, and filter for people with the same route, and find there are very similar routing behaviors of the other person.
Step 6: Further check with the transaction data, identify the correlation between two people.

Acknowledgements

This work is supported by National NSFC Project (No. 61170204) . The authors want to thank IEEE VAST Challenge 2014 Committe and reviewers. The authors want to thank IEEE PacificVis 2014 Committe and reviewers.

Contact
xiaoru.yuan@pku.edu.cn http://vis.pku.edu.cn
VisPKU