

# ChinaVIS 2016

July 21-23, Changsha

## Uncertainty-aware Visual Analytics for Exploring Human Behaviors from Heterogeneous Spatial Temporal Data

Siming Chen<sup>1</sup>, Zuchao Wang<sup>1,2</sup>, Jie Liang<sup>1,3</sup>, Xiaoru Yuan<sup>1</sup>

1) Peking University, Beijing, China

2) SQiHoo 360 Technology Co. Ltd.

3) Faculty of Engineer and Information Technology, The University of Technology, Sydney

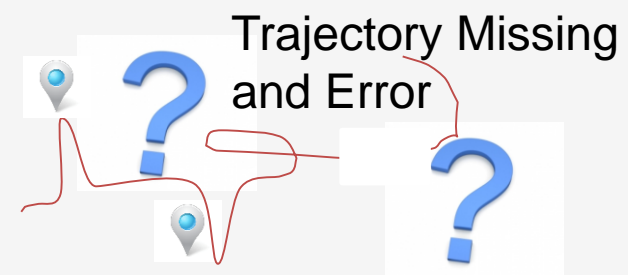
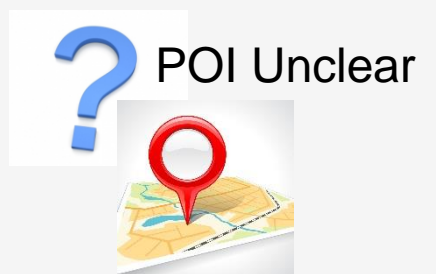
## If you are a detective




Someone missing?

Would you please help us?

# Spatial and temporal data with uncertainty



Attribute Conflict 

## Content

- ⊙ Introduction
- ⊙ Task and Data Description
- ⊙ MovementFinder
  - ⊙ Design Goal
  - ⊙ Dealing with Uncertainty
  - ⊙ Visual Design and Investigative Interaction
- ⊙ Case Study
- ⊙ Conclusion and Future Work

ChinaVis 2016 长沙

# Introduction

## Background

### GAStech

- Tethy-based international company
- Several employees missing
- Recording the GPS logs and transaction logs from the employees

### Abila

- Kronos capital

## Data

- GPS records, 685,170 items, 33MB
- Transaction records (loyalty card, credit (debt) card), around 1,491 records
- Employee information (Name, title, position and department)

## Tasks

- Describe daily behavior of the employees
- Identify up to twelve unusual events or patterns that you see
  - Who, when, where
  - Pattern, confidence and evidence

## Dealing with uncertainty

- Missing data
- Conflicting data
- Data of varying resolutions
- Outliers
- Other kinds of confusing data



Urname	Ward	Age	Administration	Assistant to CEO
Mies Haber	Ruscella	115	Administration	Assistant to CEO
Balas	Felix	3	Engineering	Engineer
Dedos	Lidelse	14	Engineering	Engineering Group Manager
Calzas	Axel	9	Engineering	Drill Technician
Orilla	Elsa	28	Engineering	Drill Technician
Tempestad	Brand	33	Engineering	Drill Technician
Frente	Birgitta	18	Engineering	Geologist
Frente	Vira	19	Engineering	Hydraulic Technician
Cazar	Gustav	11	Engineering	Hydraulic Technician
Borrasca	Isande	7	Engineering	Drill Technician
Azada	Lars	2	Engineering	Engineer
Orilla	Kare	27	Engineering	Drill Technician
Onda	Marin	26	Engineering	Drill Site Manager
Nubarron	Adra	25	Engineering	Geologist
Alcazar	Lucas	1	Information Technology	IT Helpdesk

### Spatial Overview

### Entity View

### Selection for individual illustration

White dots indicating the payment

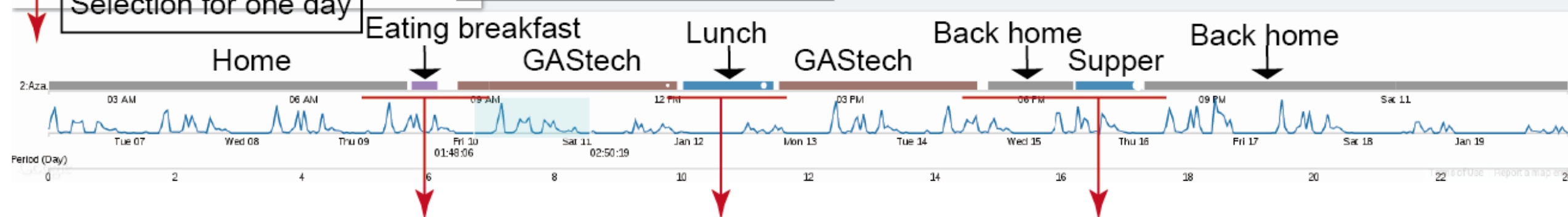
### Stop-Event View, showing every day's behavior

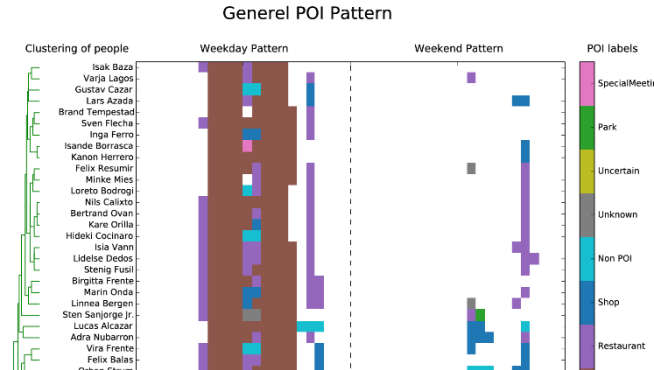
### Selection for one day

■ SpecialGoods  
■ Transportation  
■ GASTech  
■ Unknown  
■ NonPOI  
■ Shop  
■ Restaurant  
■ SpecialMeeting  
■ Uncertain

### Control Panel

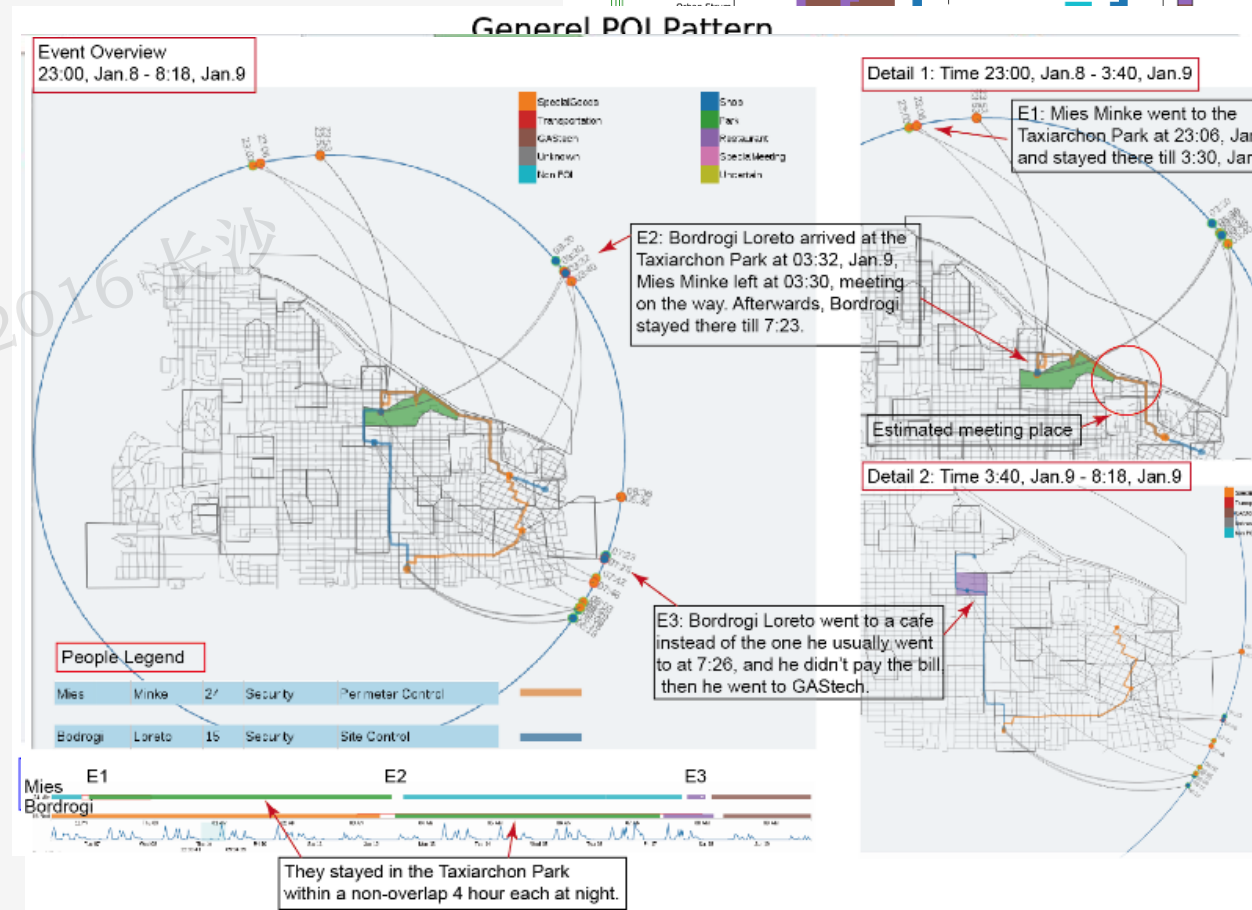
Rendering Parameters  
 Line  
 linewidth   
 opacity   
 visible   
 opacity   
 Density  
 Timeline Parameters  
 type   
 timeSte...   
 play  
 Analysis Parameters  
 Filter Parameters  
 Spatial Filter  
 filtering...   
 linkSrc   
 linkDest   
 clearDetail   
 viewSt...   
 stroke...   
 stroke0...   
 ST-Distance Matrix Filter  
 ST-Curve Density Filter  
 Close Controls





# Working Pipeline

- Data processing
  - POI matching and labeling
  - Pattern extraction
  - Outlier calculation
- Data cleaning – dealing with uncertainty
- Visualization
  - GPS tracks
  - Event sequence for each person
- Event detection with multiple filters visual analysis





## Data cleaning

### Dealing with uncertainty in the following aspects

#### POI data

- Manually extracted from tourist map and neither precise nor complete.

#### GPS data

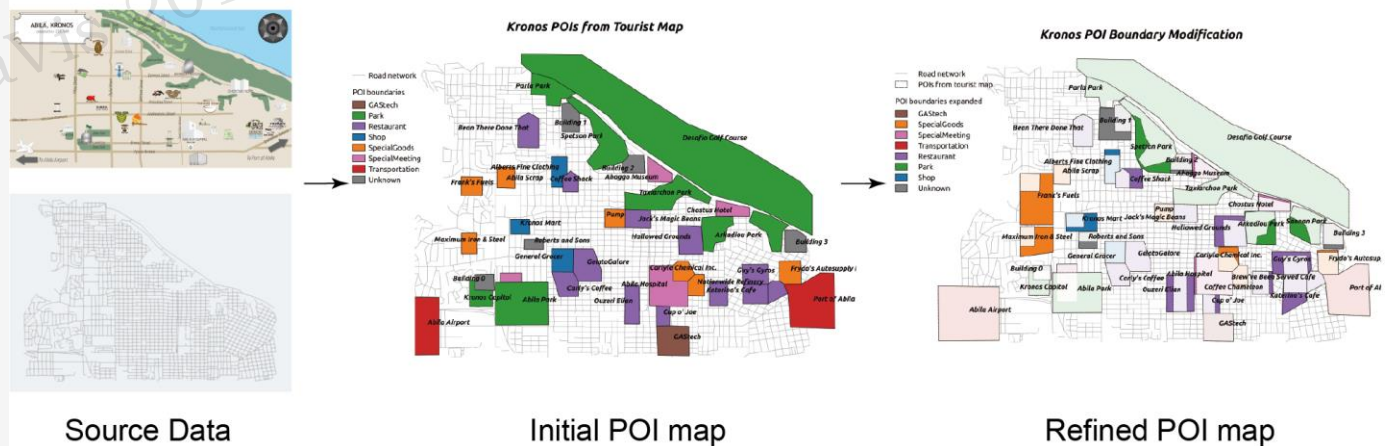
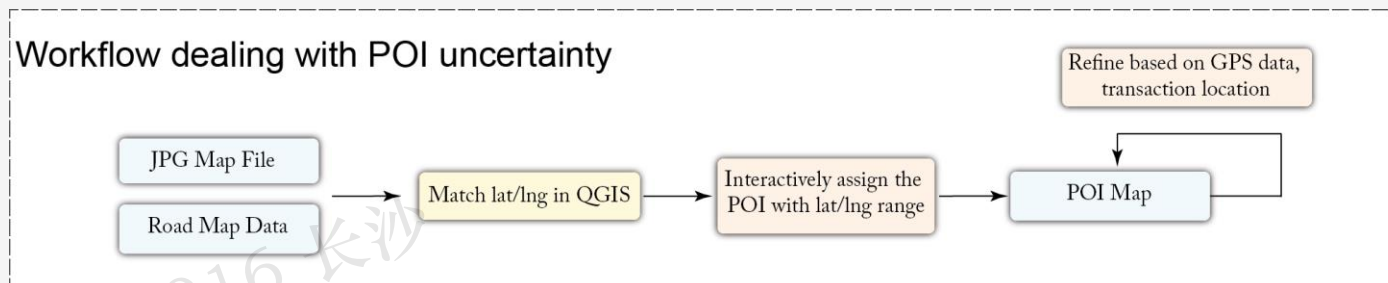
- A track can be systematically shifted.
- Noise exists
- Distance between two consecutive records can be far away.

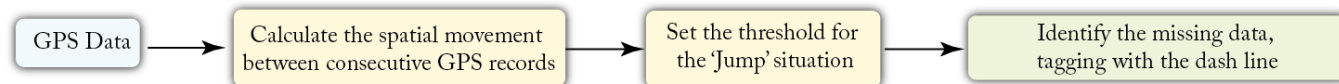
#### Transaction data

- In case of price conflicts, we show both prices.
- In case of transaction time conflicts, we select the time that better matches the GPS logs.

#### Car assignment data

- No employee ID provided





## Data cleaning

### Dealing with uncertainty in the following aspects

#### POI data

- Manually extracted from tourist map and neither precise nor complete.

#### GPS data

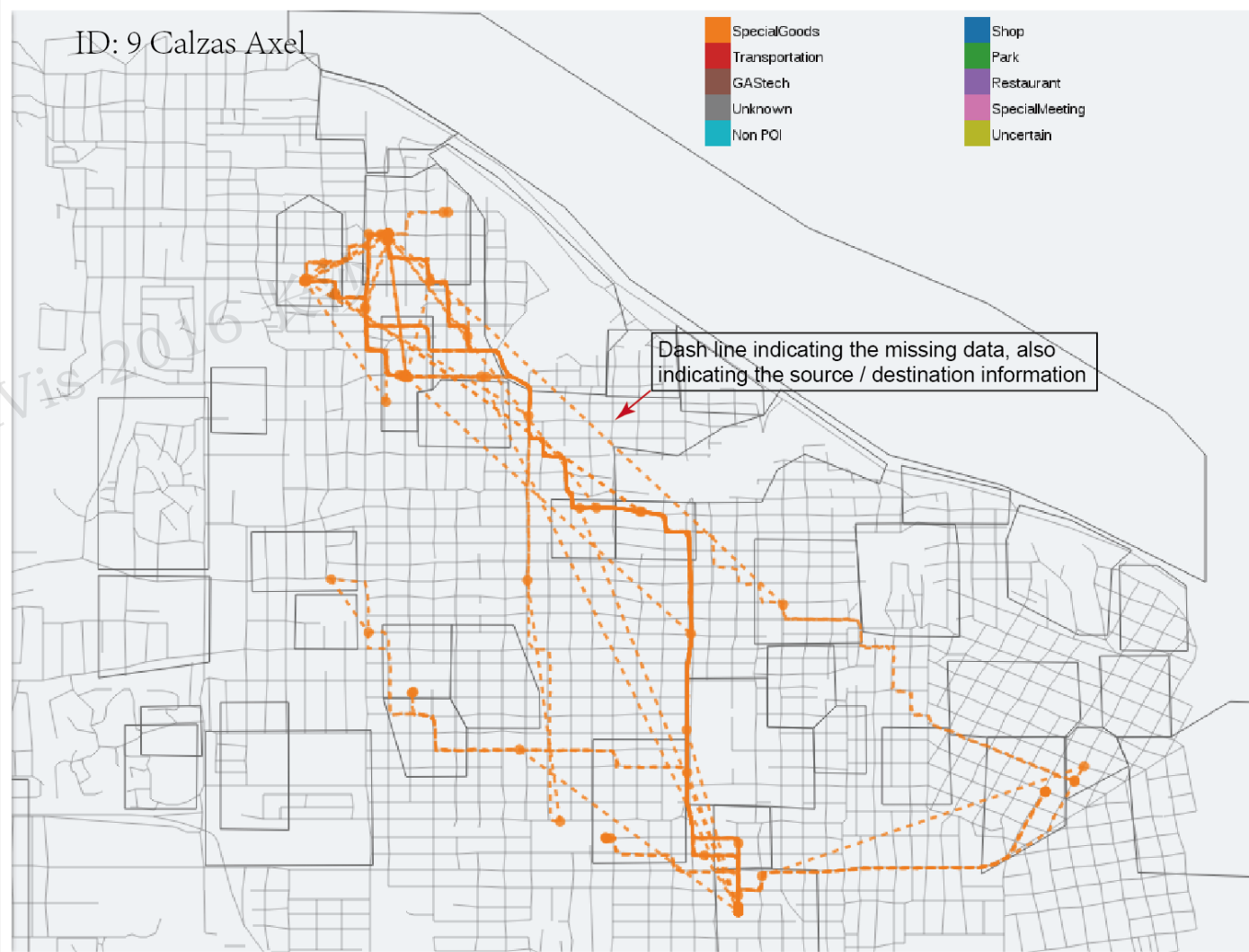
- A track can be systematically shifted.
- Noise exists
- Distance between two consecutive records can be far away.

#### Transaction data

- In case of price conflicts, we show both prices.
- In case of transaction time conflicts, we select the time that better matches the GPS logs.

#### Car assignment data

- No employee ID provided



# Data cleaning

## Dealing with uncertainty in the following aspects

### POI data

- Manually extracted from tourist map and neither precise nor complete.

### GPS data

- A track can be systematically shifted.
- Noise exists
- Distance between two consecutive records can be far away.

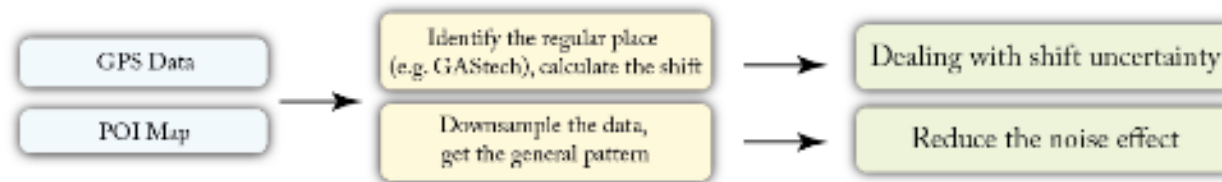
### Transaction data

- In case of price conflicts, we show both prices.
- In case of transaction time conflicts, we select the time that better matches the GPS logs.

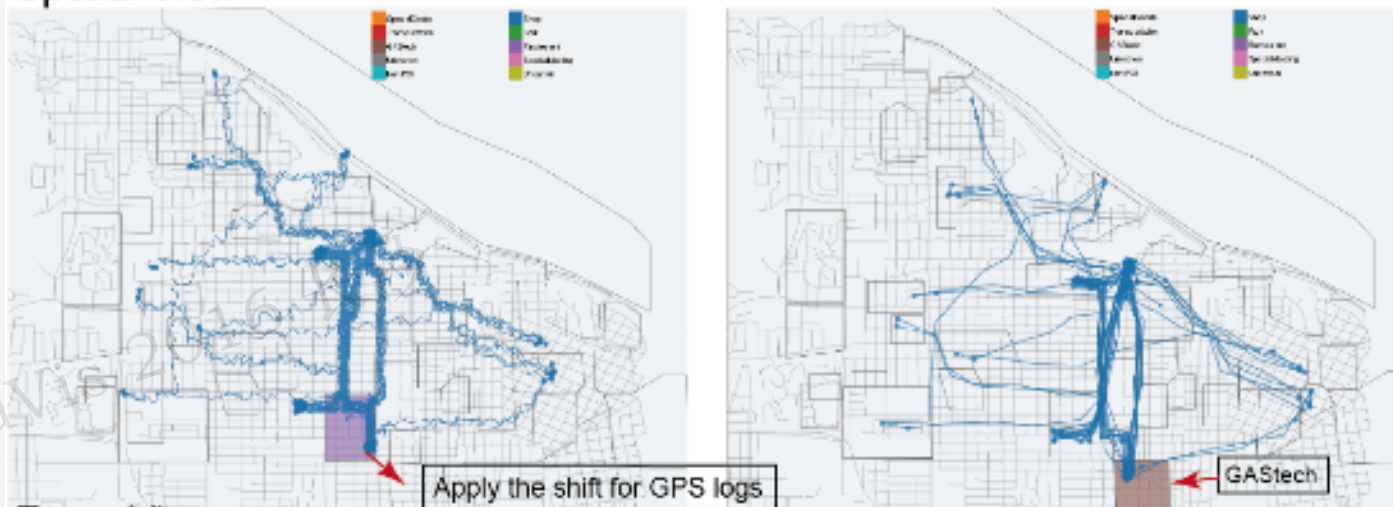
### Car assignment data

- No employee ID provided

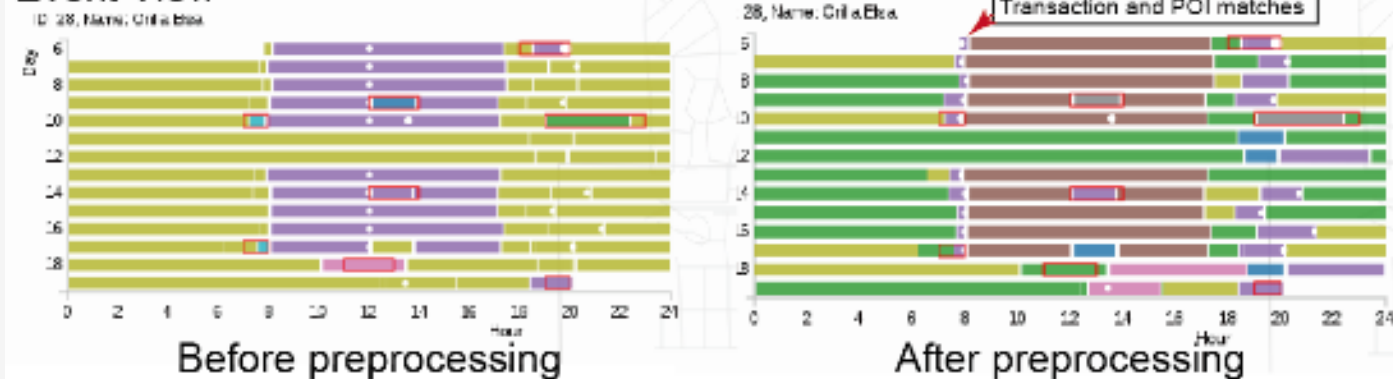
## Workflow dealing with GPS shift uncertainty



## Spatial View



## Event View



# Data cleaning

Dealing with uncertainty in the following aspects

POI data

- Manually extracted from tourist map and neither precise nor complete.

GPS data

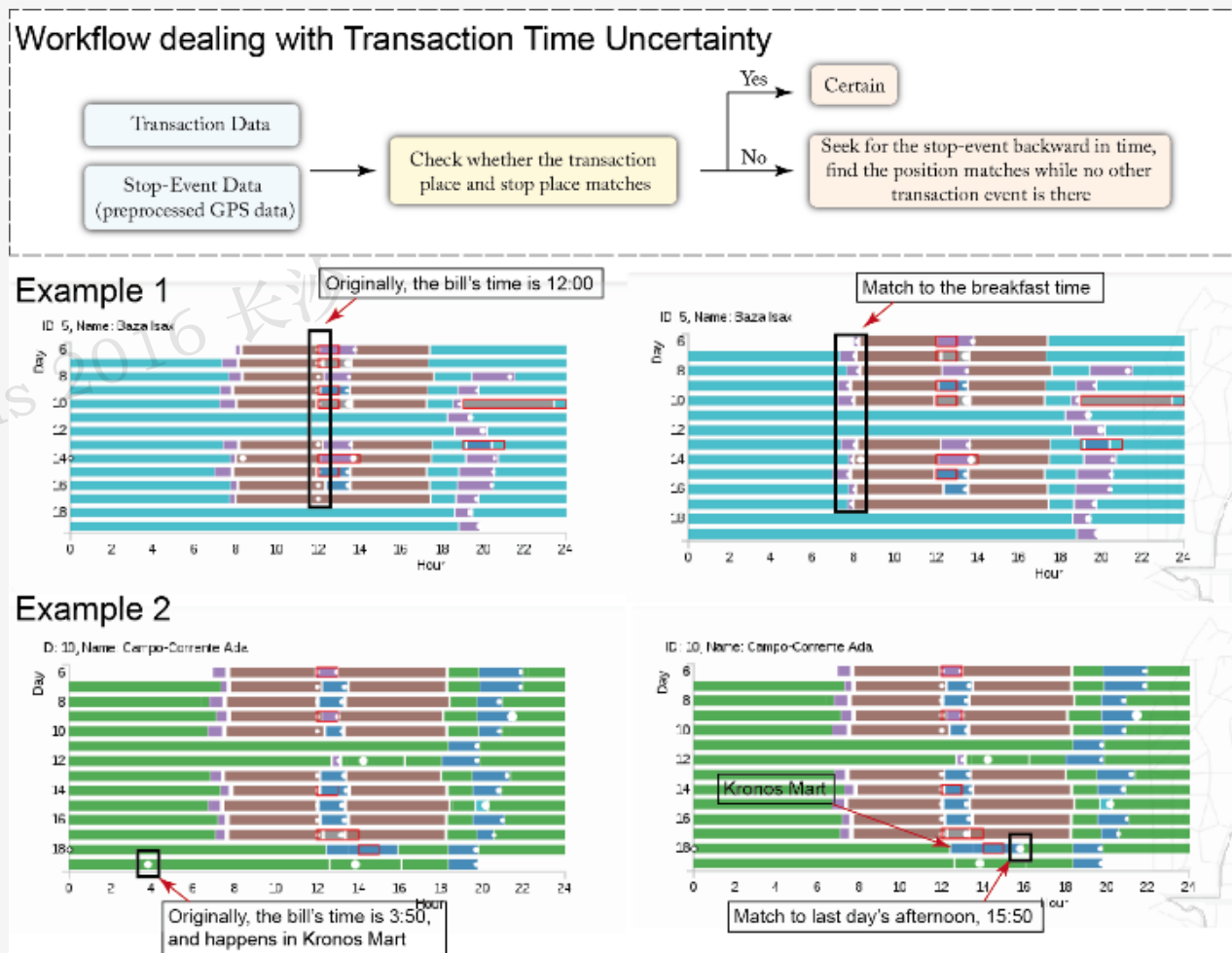
- A track can be systematically shifted.
- Noise exists
- Distance between two consecutive records can be far away.

Transaction data

- In case of price conflicts, we show both prices.
- In case of transaction time conflicts, we select the time that better matches the GPS logs.

Car assignment data

- No employee ID provided



## Data cleaning

### Dealing with uncertainty in the following aspects

#### POI data

- Manually extracted from tourist map and neither precise nor complete.

#### GPS data

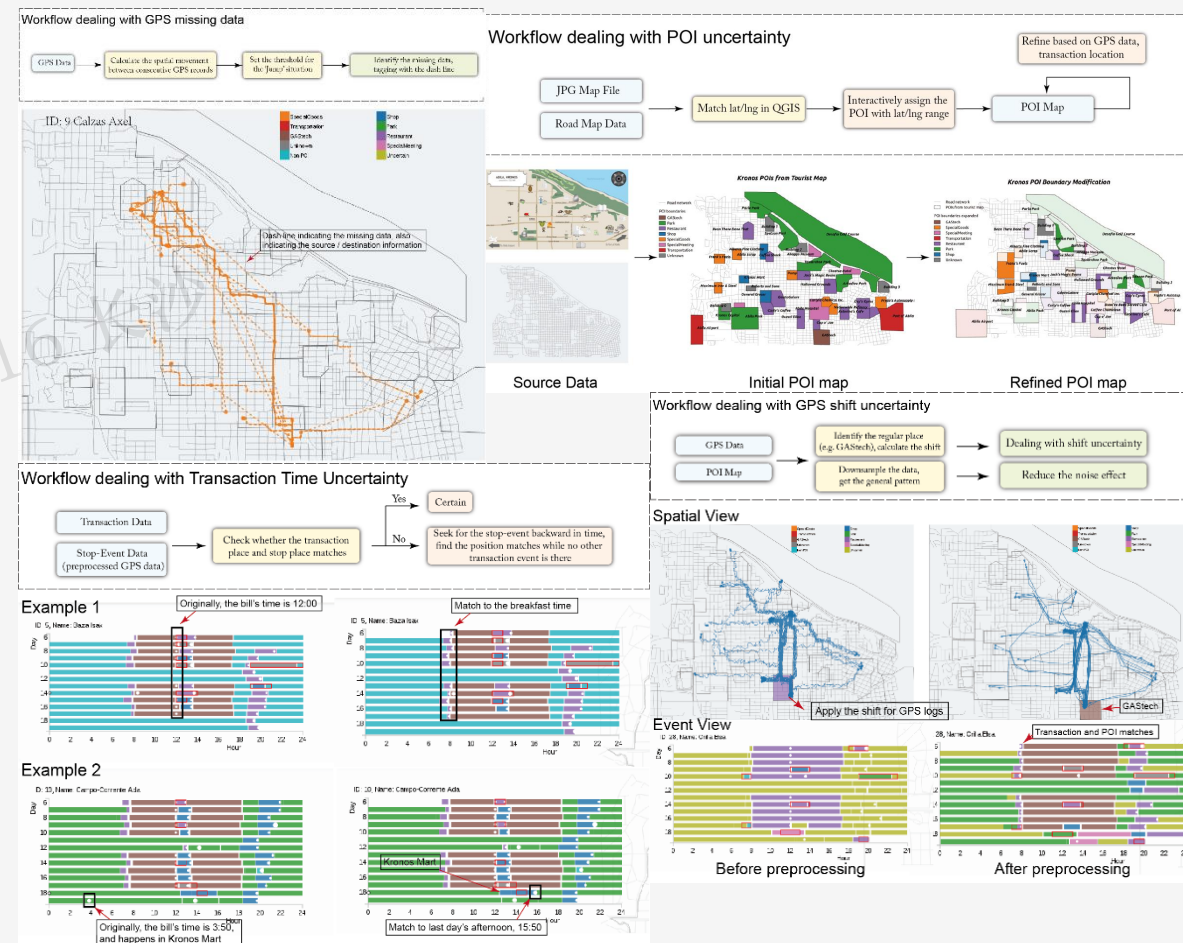
- A track can be systematically shifted.
- Noise exists
- Distance between two consecutive records can be far away.

#### Transaction data

- In case of price conflicts, we show both prices.
- In case of transaction time conflicts, we select the time that better matches the GPS logs.

#### Car assignment data

- No employee ID provided




# Uncertainty classification and dealing methods

## ◎ Methods

- ◎ Data based reliability improvement
- ◎ Cross reference of multiple source of data to enrich the data
- ◎ Visually identification and automatic methods

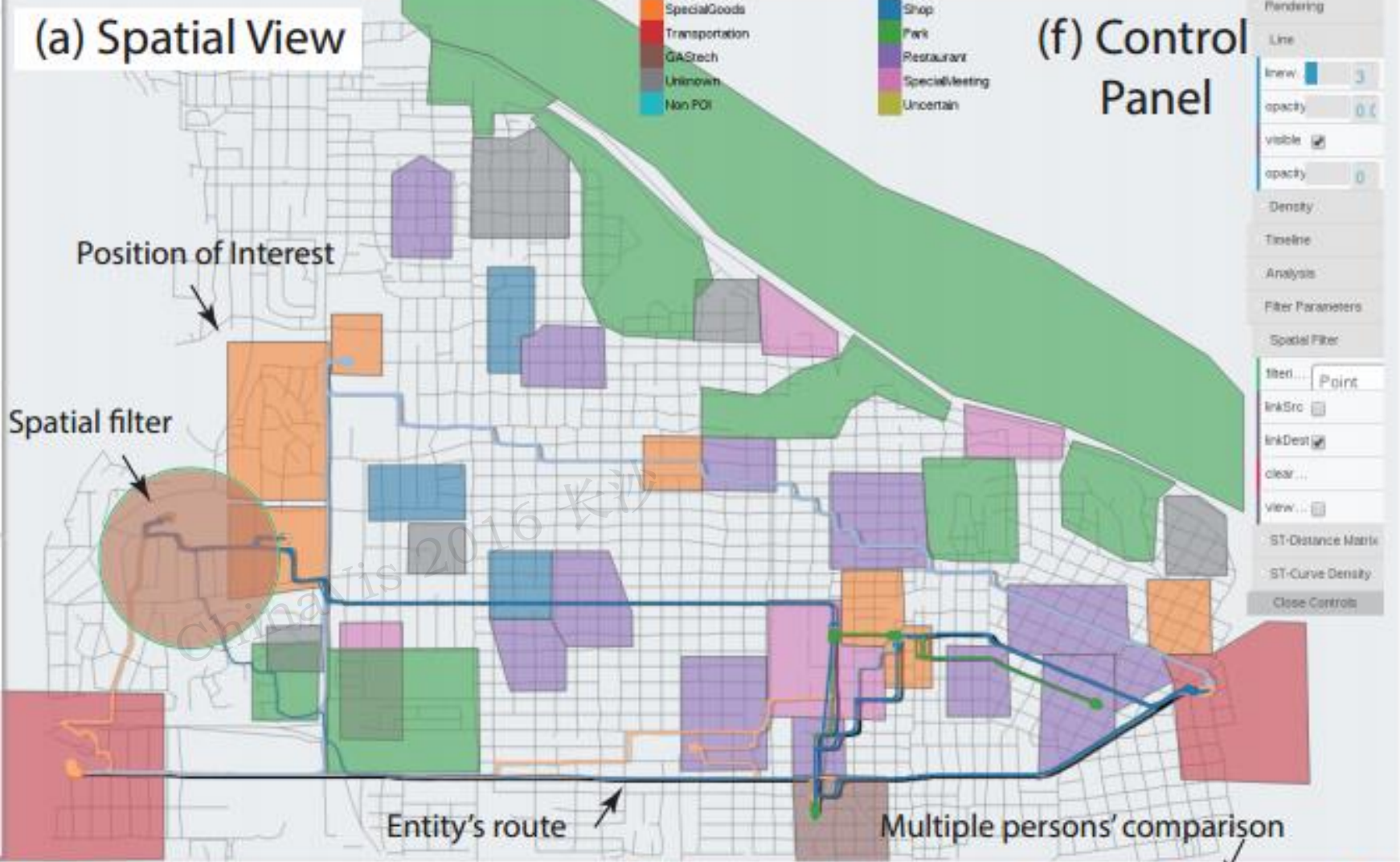
## ◎ Benefits

- ◎ Precise information and confidence level of information
- ◎ Foundation for event detection

Example: Geo information shift and downsampling 

**(c) Entity View**

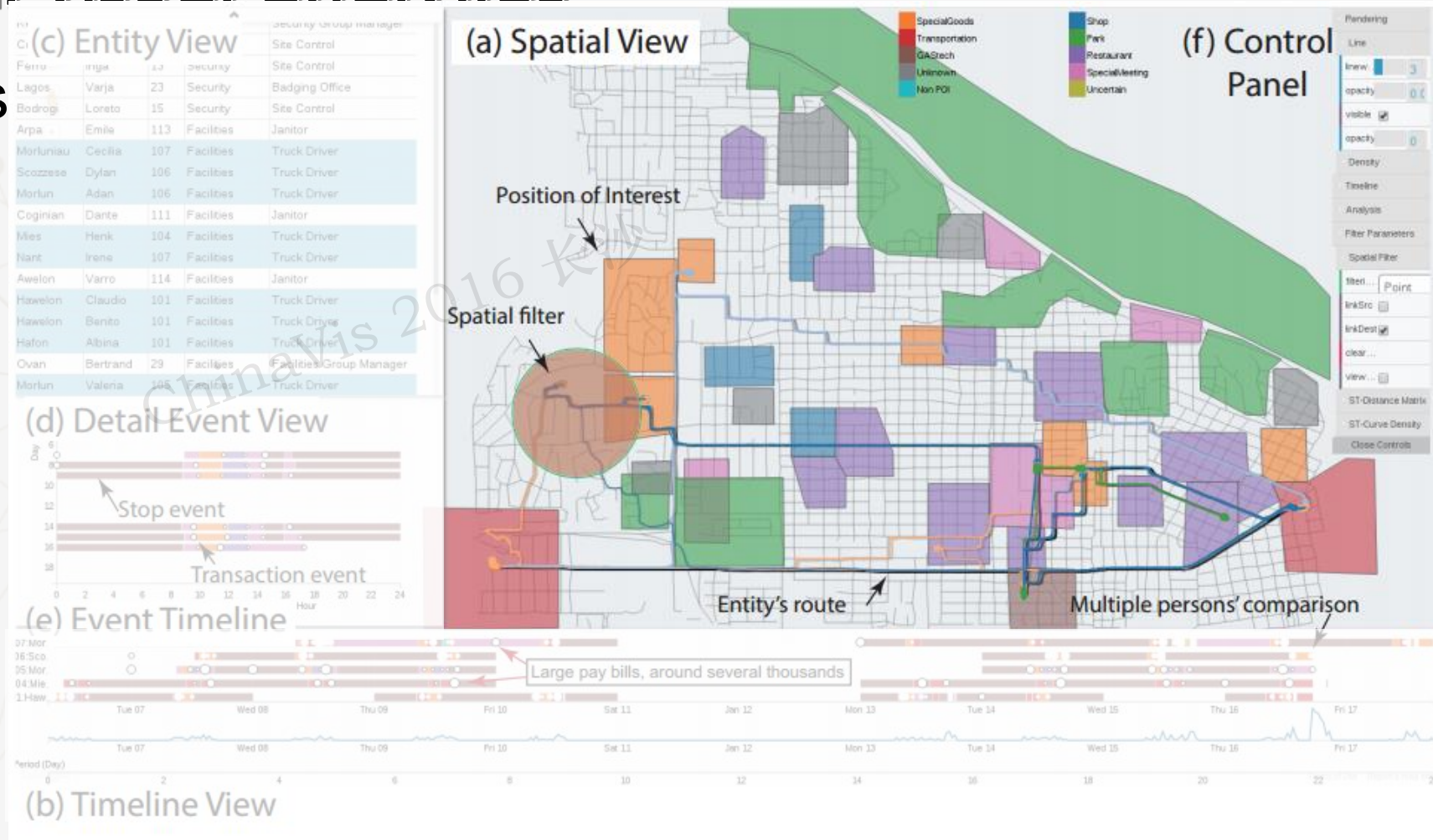
Entity	Name	Age	Category	Role
Ferru	Inga	43	Security	Security Group Manager
Lagos	Varja	23	Security	Badging Office
Bodrogi	Loreto	15	Security	Site Control
Arpa	Emile	113	Facilities	Janitor
Morluniau	Cecilia	107	Facilities	Truck Driver
Scozzese	Dylan	106	Facilities	Truck Driver
Morlun	Adan	106	Facilities	Truck Driver
Coginian	Dante	111	Facilities	Janitor
Mies	Henk	104	Facilities	Truck Driver
Nant	Irene	107	Facilities	Truck Driver
Awslon	Varro	114	Facilities	Janitor
Hawelson	Claudio	101	Facilities	Truck Driver
Hawelson	Benito	101	Facilities	Truck Driver
Hafon	Albina	101	Facilities	Truck Driver
Ovan	Bertrand	29	Facilities	Facilities Group Manager
Morlun	Valena	105	Facilities	Truck Driver



**(b) Timeline View**

# Spatial temporal visual analytics

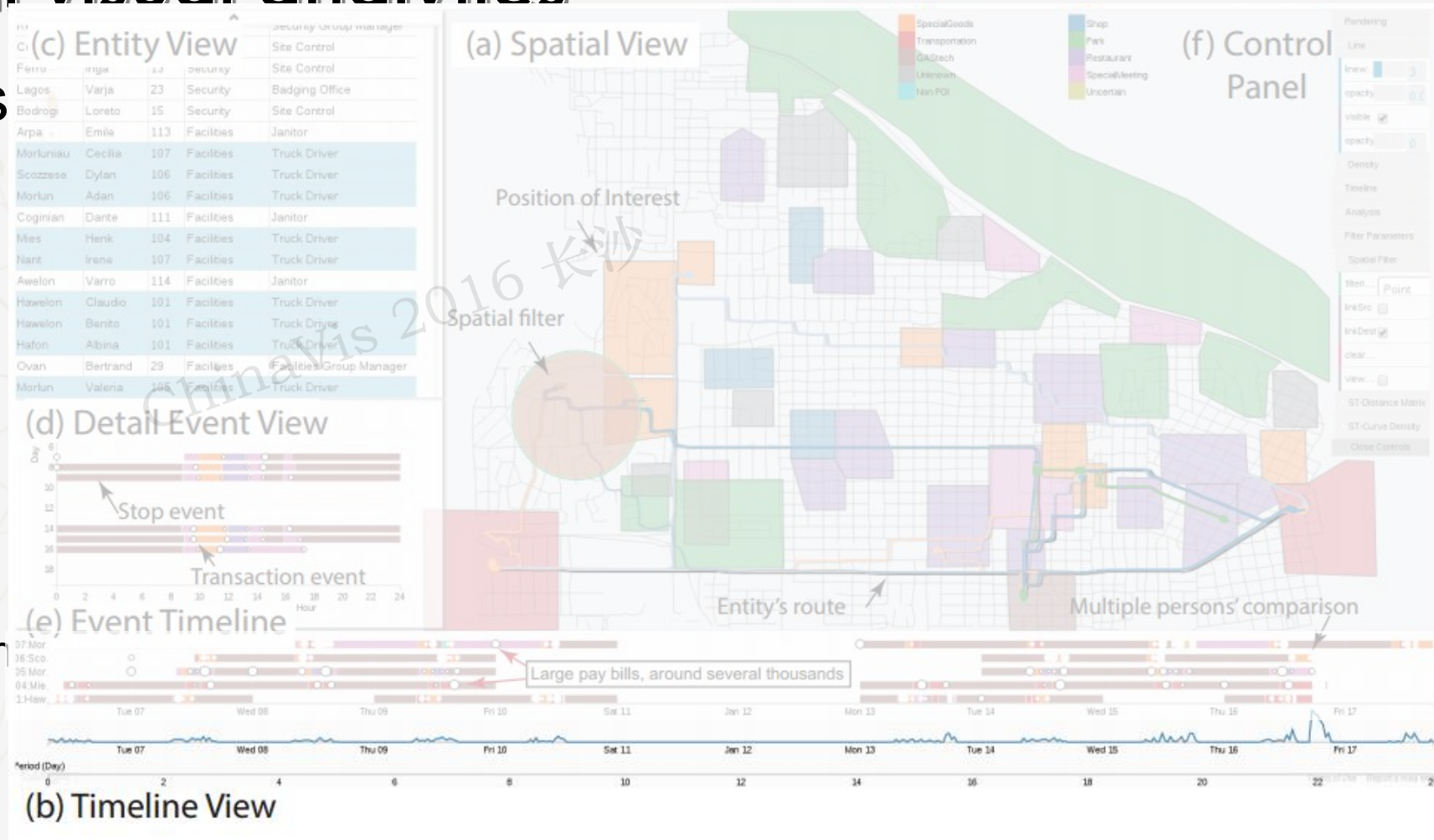
- Visual Components
  - Spatial View
    - POIs
    - GPS tracks
    - Spatial filtering





# Spatial temporal visual analytics

- Visual Components
  - Spatial View
    - POIs
    - GPS tracks
    - Spatial filtering
  - Temporal View
    - GPS logs distribution
    - Temporal filtering
    - Periodical selection



# Spatial temporal visual analytics

## • Visual Components

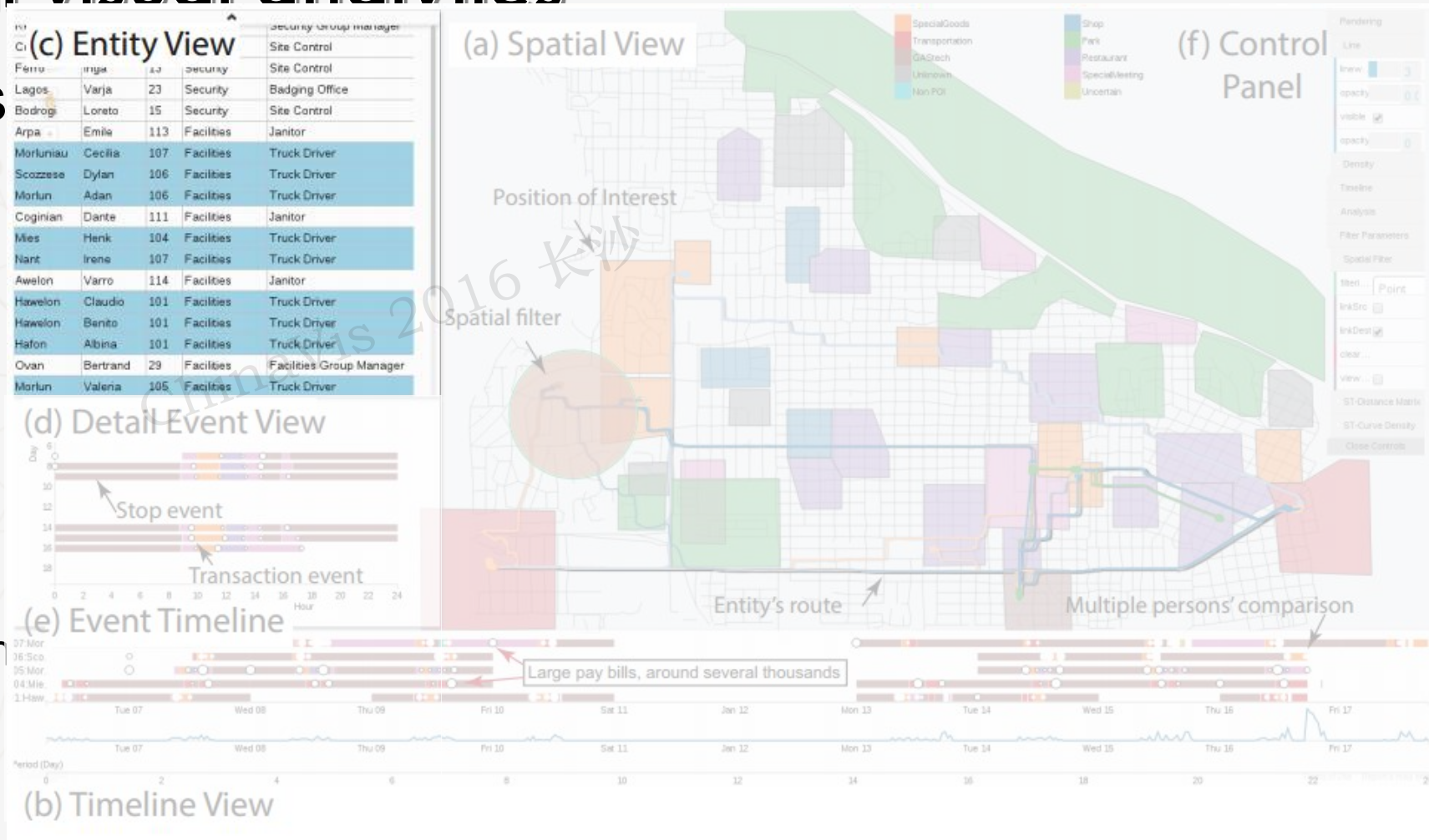
### • Spatial View

- POIs
- GPS tracks
- Spatial filtering

### • Temporal View

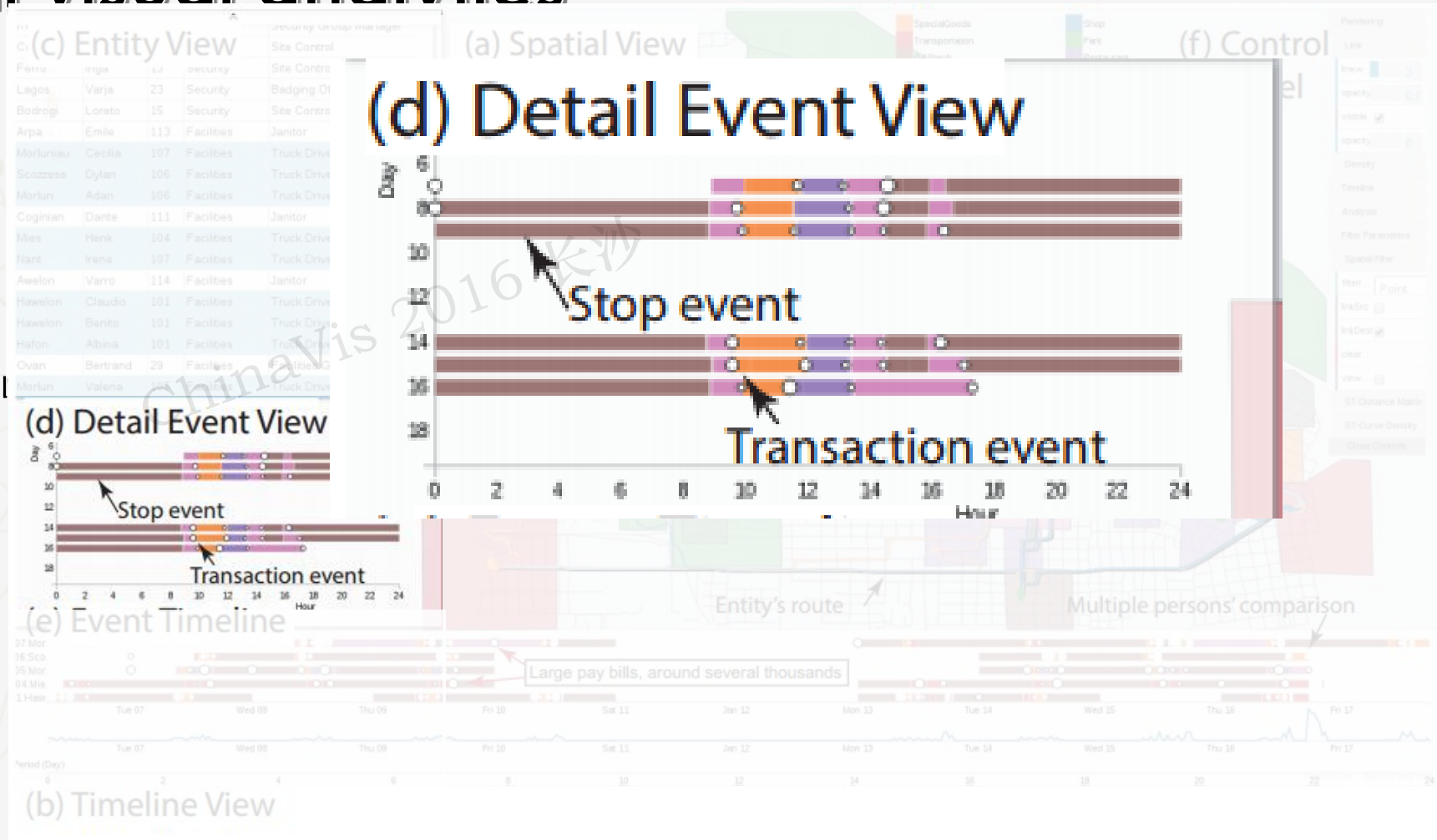
- GPS logs distribution
- Temporal filtering
- Periodical selection

### • Entity View



# Spatial temporal visual analytics

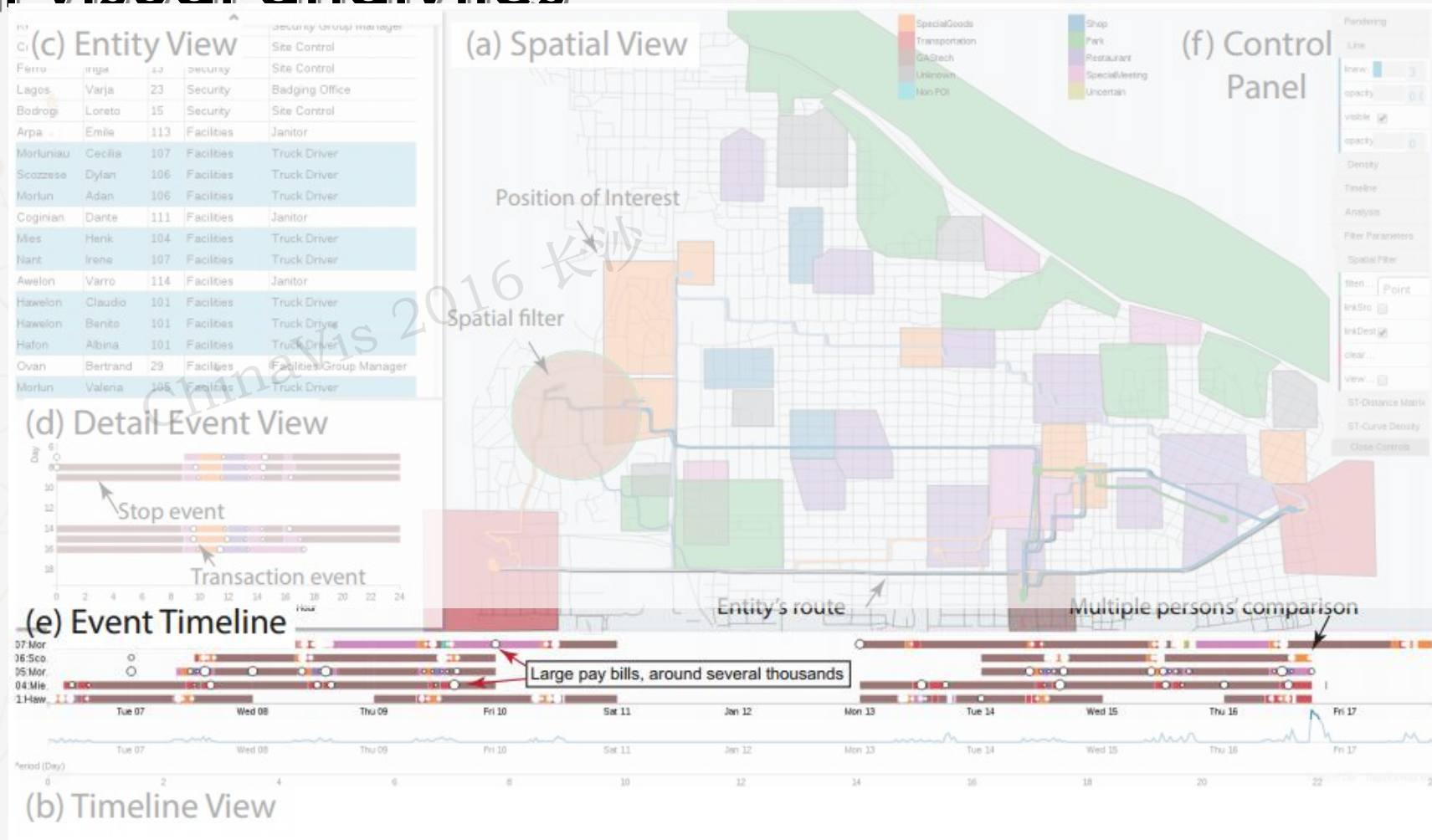
- Visual Components
  - Spatial View
    - POIs
    - GPS tracks
    - Spatial filtering
  - Temporal View
    - GPS logs distribution
    - Temporal filtering
    - Periodical selection
  - Entity View
  - Detail Event View
    - Event sequence visualization
    - Comparison and outlier detection



# Spatial temporal visual analytics

## • Visual Components

- Temporal View
  - GPS logs distribution
  - Temporal filtering
  - Periodical selection
- Spatial View
  - POIs
  - GPS tracks
  - Spatial filtering
- Entity View
- Detail Event View
  - Event sequence visualization
  - Comparison and outlier detection



# Spatial temporal visual analytics

## • Visual Components

### • Temporal View

- GPS logs distribution
- Temporal filtering
- Periodical selection

### • Spatial View

- POIs
- GPS tracks
- Spatial filtering

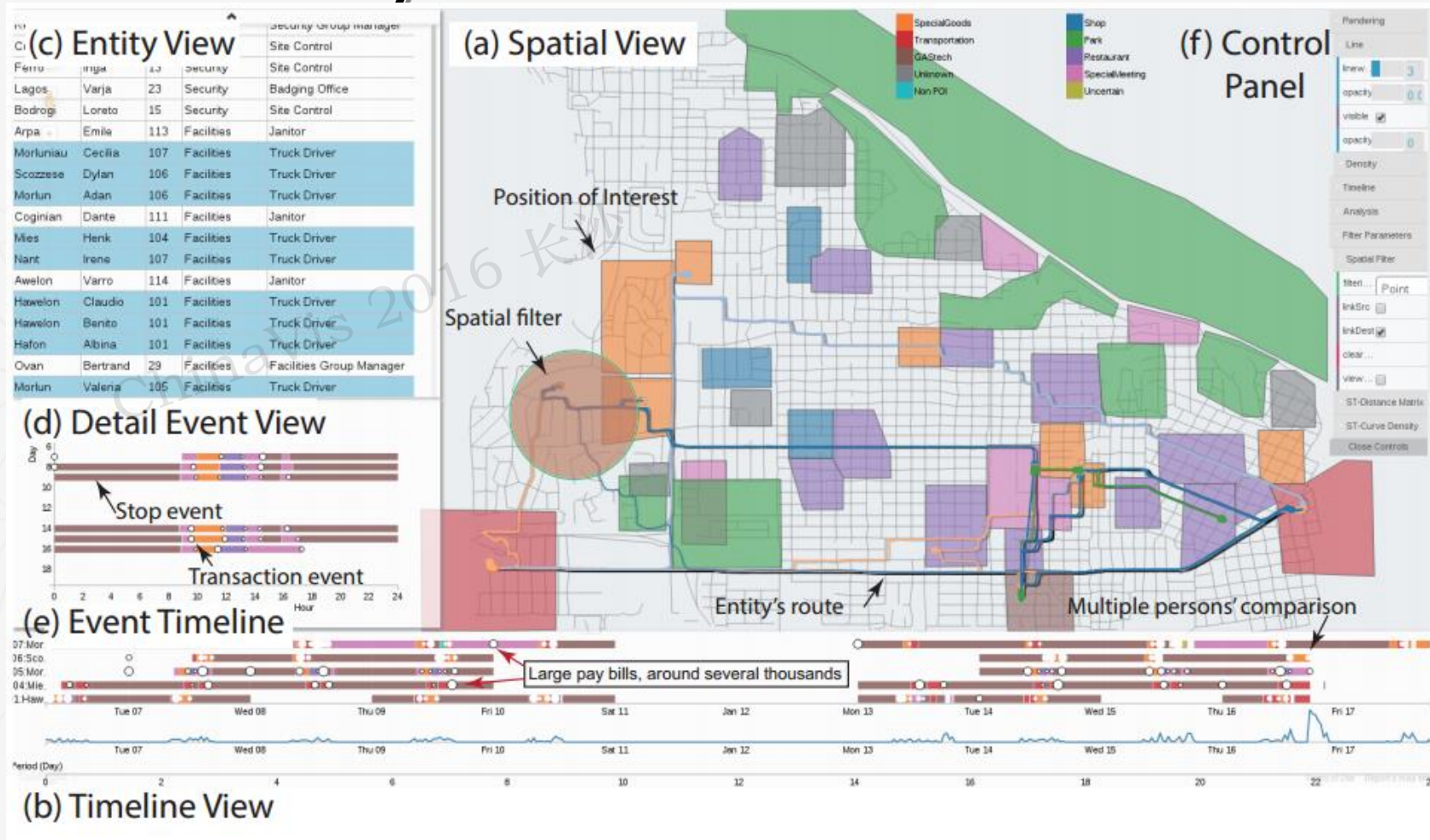
### • Entity View

### • Detail Event View

- Event sequence visualization
- Comparison and outlier detection

### • Spatial Temporal Circular Design

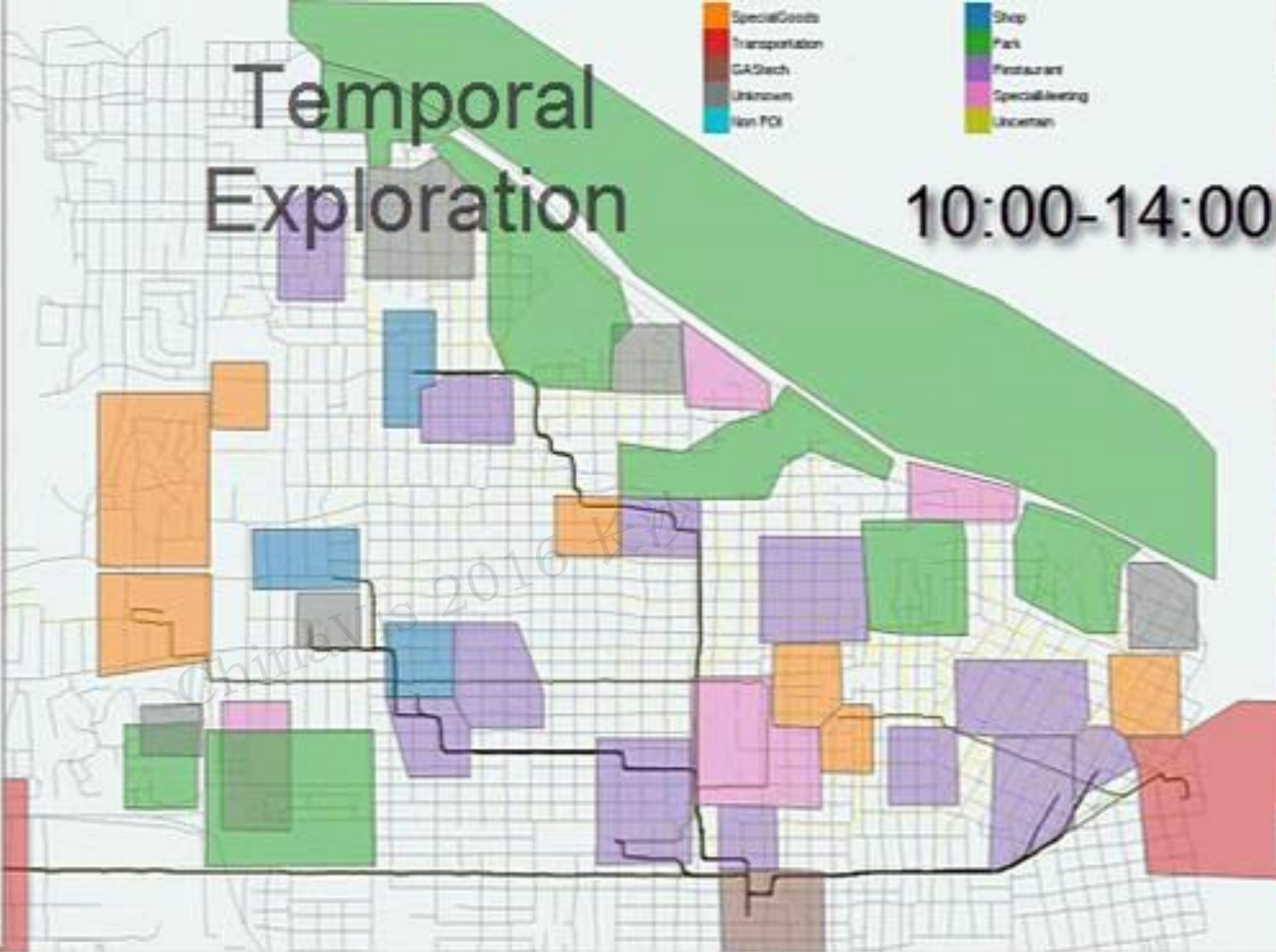
- Detail event exploration
- Spatial temporal correlation



Lastname	Firstname	CardID	Type	Title
Sarjorge Jr.	Sten	31	Executive	President/CEO
Campo-Corrente	Ada	10	Executive	SVP/COO
Barranco	Ingrid	4	Executive	SVP/CFO
Vasco-Pais	Wilem	35	Executive	Environmental Safety Advisor
Strum	Orhan	32	Executive	SVP/COO
Ribera	Anda	119	Administration	Assistant to CFO
Forjuniau	Carla	118	Administration	Assistant to IT Group Manager
Lagos	Linda	112	Administration	Assistant to COO
Las	Comelia	110	Administration	Assistant to Security Group Manager
Pantanal	Rachel	117	Administration	Assistant to COO
Bramar	Mat	116	Administration	Assistant to CEO
Mies Haber	Ruscella	115	Administration	Assistant to Engineering Group Manager
Datta	Eric	7	Executive	Executive

# Temporal Exploration

10:00-14:00



- Special Goods
- Transportation
- GA/Sech
- Unknown
- Non POI
- Shop
- Park
- Restaurant
- Specializing
- Uncertain

Rendering Parameters

Line

linewidth

opacity

visible

opacity (top)

Density

Timeline Parameters

type

lineOpacity

map

Analysis Parameters

putCross

crossPad

enableFilter

operation

viewChart

visible

viewEmail

map

map

map

Filter Parameters

Close Controls

Day

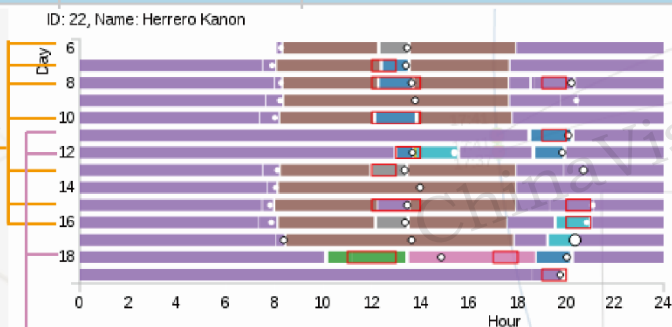
Map



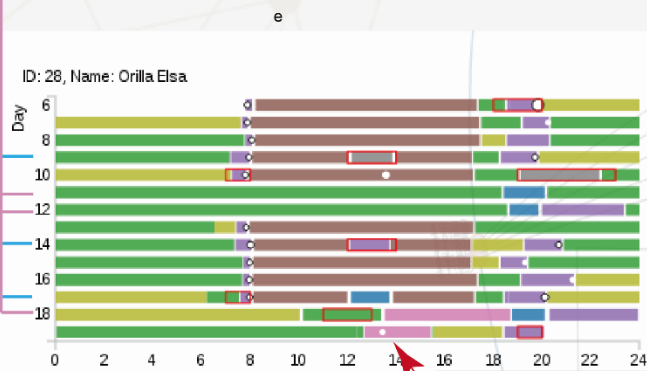
### Case: Investigation after uncertainty dealing

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

Pattern 1: They go together in one car. Herreron Kanon drove, and he paid for two in lunch while Orilla Elsa used her Loyalty Card

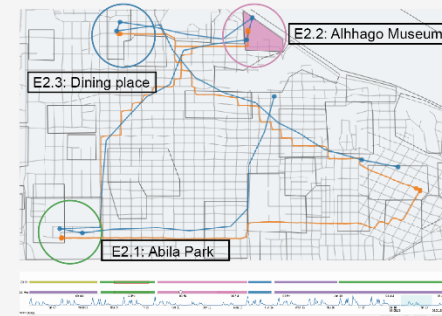
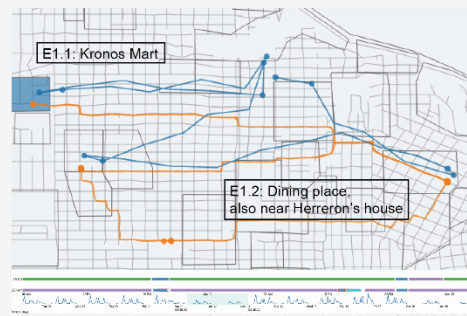
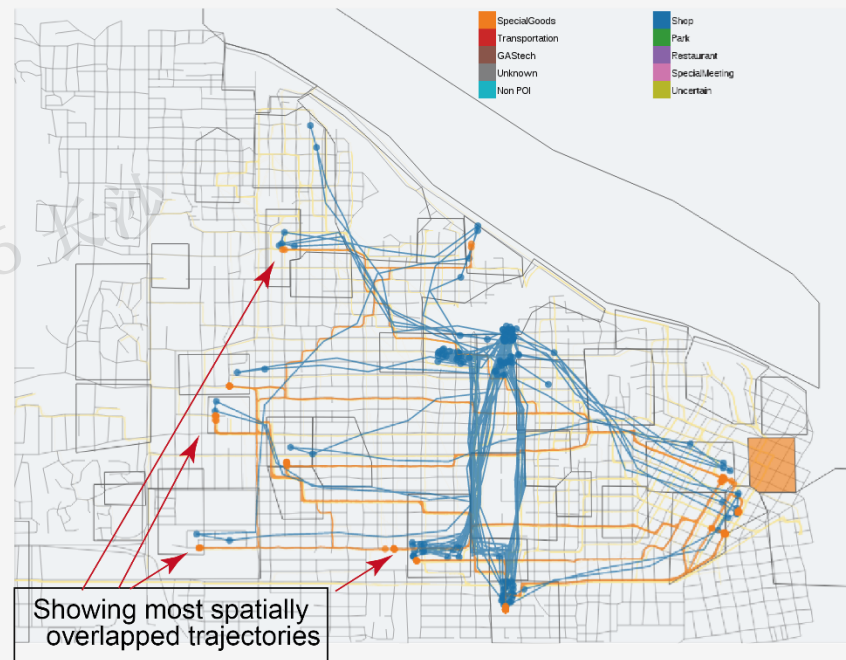


Pattern 3: In the weekend, they drove together, went to eating, park or museum. Herreron Kanon paid for the bill.



Pattern 2: They go together in one car. Orilla Elsa drove, Herrero Kanon paid for two in lunch while Orilla Elsa used her Loyalty Card

Outlier: Orilla went to the musuem again in Sunday



## Conclusion

- ◉ Dealing with uncertainty
  - ◉ Combining several types of data, referring to the correct and higher resolution ones
  - ◉ Dealing with uncertainty in multiple dimension
- ◉ Combination of multiple filters
  - ◉ Visual summary of daily patterns
  - ◉ Provide flexible investigation tool
  - ◉ Support comparison for further event detection
- ◉ Future work
  - ◉ Try data with larger scale
  - ◉ Provide more aggregation methods for group analysis



**ChinaVIS**

07.21~23 Changsha

2016

The 3rd China Visualization and Visual Analytics Conference  
第三届中国可视化与可视分析大会

Thank you !

ChinaVis 2016 长沙