MC2: Finding Anomalies from Geospatio-temporal Patterns

Honorable Mention for Effective Detailed Analysis

Georgia Tech
Hannah Kim, Jaegul Choo, Francesco Poggi, James Nugent, Yi Han, Mengdie Hu, Haesun Park, John Stasko
Overview

Goal: Find unusual events/patterns
Who, where, when, why?

- GPS coordinates (latitude, longitude)
- Shops
  - e.g., Jack’s Magical Beans
  - Nationwide Refinery
- Employees
  - e.g., Sten Sanjorge Jr.
- Cars
  - e.g., Car 1, 2
  - Truck 101
Data

GPS coordinates

Shops

Cars

Employees

Car-employee assignment

<table>
<thead>
<tr>
<th>Name</th>
<th>Car ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcazar Lucas</td>
<td>1</td>
</tr>
<tr>
<td>Azada Lars</td>
<td>2</td>
</tr>
</tbody>
</table>
Data

GPS coordinates

Shops

Cars

Employees

Car tracking data

<table>
<thead>
<tr>
<th>Timestamp</th>
<th>Car ID</th>
<th>lat</th>
<th>long</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/6/2014 6:28:01</td>
<td>35</td>
<td>36.076</td>
<td>24.875</td>
</tr>
<tr>
<td>1/6/2014 6:28:03</td>
<td>35</td>
<td>36.076</td>
<td>24.874</td>
</tr>
</tbody>
</table>

Car-employee assignment

<table>
<thead>
<tr>
<th>Name</th>
<th>Car ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcazar Lucas</td>
<td>1</td>
</tr>
<tr>
<td>Azada Lars</td>
<td>2</td>
</tr>
</tbody>
</table>
### Data

#### GPS coordinates

<table>
<thead>
<tr>
<th>Timestamp</th>
<th>Car ID</th>
<th>lat</th>
<th>long</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/6/2014 6:28:01</td>
<td>35</td>
<td>36.076</td>
<td>24.875</td>
</tr>
<tr>
<td>1/6/2014 6:28:03</td>
<td>35</td>
<td>36.076</td>
<td>24.874</td>
</tr>
</tbody>
</table>

#### Car tracking data

<table>
<thead>
<tr>
<th>Timestamp</th>
<th>Car ID</th>
<th>lat</th>
<th>long</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/6/2014 6:28:01</td>
<td>35</td>
<td>36.076</td>
<td>24.875</td>
</tr>
<tr>
<td>1/6/2014 6:28:03</td>
<td>35</td>
<td>36.076</td>
<td>24.874</td>
</tr>
</tbody>
</table>

#### Cars

<table>
<thead>
<tr>
<th>Name</th>
<th>Car ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcazar Lucas</td>
<td>1</td>
</tr>
<tr>
<td>Azada Lars</td>
<td>2</td>
</tr>
</tbody>
</table>

#### Employees

<table>
<thead>
<tr>
<th>Name</th>
<th>Car ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcazar Lucas</td>
<td>1</td>
</tr>
<tr>
<td>Azada Lars</td>
<td>2</td>
</tr>
</tbody>
</table>

#### Shops

<table>
<thead>
<tr>
<th>Timestamp</th>
<th>location</th>
<th>price</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/6/2014</td>
<td>Jack's Beans</td>
<td>18.59</td>
<td>Elsa</td>
</tr>
</tbody>
</table>

#### Card/loyalty spending data
Missing/Conflicting Data

- **Shops**
  - Card/loyalty spending data
  - Credit card vs. loyalty amount mismatch

- **Employees**
  - Car-employee assignment
  - No 1-1 matching between Trucks – truck drivers

- **Cars**
  - Car tracking data
  - Missing chunk

- **GPS coordinates**

?
Identifying Shop Locations

- Shop locations
- GPS coordinates
- Shops
- Employees
- Cars
Identifying Shop Locations

For each spending record, pick a coordinate that matches spending timestamp

<table>
<thead>
<tr>
<th>Timestamp</th>
<th>Car ID</th>
<th>lat</th>
<th>long</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/6/2014 11:30</td>
<td>1</td>
<td>36.0762</td>
<td>24.87469</td>
</tr>
<tr>
<td>1/6/2014 12:10</td>
<td>1</td>
<td>36.0762</td>
<td>24.87469</td>
</tr>
</tbody>
</table>

Table:

<table>
<thead>
<tr>
<th>Timestamp</th>
<th>location</th>
<th>price</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/6/2014 12:00</td>
<td>Jack's Beans</td>
<td>18.59</td>
<td>Lucas</td>
</tr>
</tbody>
</table>
Identifying Shop Locations

- Golf course
- Hotel
- Mart
- Gelato
- Coffee shops

Diagram showing the locations of various shops.
Identifying Shop Locations
Other Locations Identified
Processed Data

Tracking + Spending data
Vis → meaningful info

e.g. The executives play golf together at weekends

- Shop locations
- GPS coordinates
- Cars
- Car tracking data
- Employees
- Card/loyalty transaction
- Car-employee assignment
Our Tool

Space-time cube view

Geospatial trajectory over time

Timeline view

Staying (+spending) pattern at a particular location

Matrix view

Spending amount (Shop or Employee) x Day
Our Tool
Space-time cube view

- Employees’ trajectory over space and time
- Employee-focused
Demo

Brand and Isande at Chostus Hotel
Our Tool
Timeline view

- Staying pattern (duration) for a particular location
- Location-focused
Demo

Suspicious visits on Gastech executives
Our Tool Matrix view

- Spending amount heatmap
- (Shop x Day) for a particular employee
- (Employee x Day) at a particular shop
Demo

Flexible interactions between three views

Lucas’s unusual spending at Frydos Autosupply
Thank you!

Summary
MATLAB’s built-in interaction capabilities

Future Work
Interactive Visual MATLAB

Acknowledgements
This work is supported by
- NSF grants CCF-0808863 and IIS-0915788
- DARPA XDATA grant FA8750-12-2-0309
- The DHS VACCINE Center of Excellence