Understanding Travel Patterns of Commuting Private Cars using Big data of Electronic Registration Identification of Vehicles

Presenter: Junchao Lv
Member: Dr. Linjiang Zheng
Junchao Lv  Yuhui Ye
Chenglin Ye

July 9th, 2020
1. Problem
2. Method
3. Experiment
4. Conclusion
1. Problem

- **Private car ownership grows rapidly**

  As shown in Figure 1, from 2013 to 2019, China's private car ownership increased rapidly, and its share in motor vehicle ownership increased from **35.20%** in 2013 to **59.48%** in 2019.

---

**Figure 1. The number and growth rate of motor vehicles, cars and private cars in China from 2013 to 2019**
1. Problem

 Residents travel more towards private cars

• In 2017, a survey conducted in Ningbo, China revealed that private cars are the preferred mode of transportation for Ningbo residents, as shown in Figure 2.

Figure 2. The main travel modes of Ningbo residents
1. Problem

☐ Private cars account for a high proportion of urban road traffic

• Taking Chongqing as an example, Figure 3 shows the distribution of Chongqing's traffic volume for a week.

Figure 3. Chongqing's traffic volume distribution for a week
1. Problem

- Private cars account for a high proportion of urban road traffic
  - In order to more intuitively reflect the proportion of private cars, on the basis of Figure 3, we further draw the private car traffic ratio chart, as shown in Figure 4.

![Figure 4. Proportion of private cars](image)

- Morning peak: 06:00-10:00
- Evening peak: 16:00-20:00
- Off-peak period: other hours
1. Problem

- **Private car has exacerbated traffic problems**

  - The growth of the number of private cars and the preference of residents for private cars, although satisfying the residents' personalized travel and providing convenience for residents to travel, but also aggravated the problem of urban road traffic congestion, especially the morning and evening peaks, as shown in Figure 5.

  ![Figure 5. Morning peak traffic congestion](image-url)
1. Problem

- **Private cars travel more densely in the morning and evening peak**
  - In Figure 4, the highest percentage of private cars is located in the morning and evening peak period.
  - The private car traffic in the morning and evening peak period (8 hours) is more than the off-peak period (16 hours), as shown in Figure 6.

<table>
<thead>
<tr>
<th></th>
<th>Peak period</th>
<th>Off-peak period</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time:</strong></td>
<td>33.33%</td>
<td>66.67%</td>
</tr>
<tr>
<td><strong>Traffic volume</strong></td>
<td>50.68%</td>
<td>49.32%</td>
</tr>
</tbody>
</table>

Figure 6. Comparison of peak hours and off-peak hours
1. Problem

- A large number of private cars are used for commuting

  - Commuting:
    
    ✓ Definition: Commuting is the process of going there and back between home and workplace.
    
    ✓ Feature:
      
      • Commuting mainly occurs in the morning and evening rush hours.
      • Commuting reflects the long-term travel behavior of people.
      • Commuting relates to the home and workplace of individual commuters.
    
    ✓ Negative impact: Commuting engendered a series of problems, such as traffic congestion and long commuting time.

  - Commuting mainly occurs in the morning and evening rush hours and private cars travel more densely in the morning and evening rush hours, indicating that there is a connection between commuting and private cars.

  - Therefore, exploring the connection between the two, that is, discovering the private car group used for commuting, is very meaningful for taking appropriate measures against this group to mitigate the negative impact of commuting.
1. Problem

- **Difficulties in obtaining private car data, ERI data provides support for related research**
  - Data sources for private car data:
    - GPS data
    - ERI data
  - GPS Data Limits on Private Car Research:
    - GPS data on private cars is difficult to obtain;
    - The obtained GPS data on private cars only covers some private cars, and cannot reflect the characteristics of city-level private cars.
  - Advantages of ERI data:
    - Full vehicle coverage. It can cover all private cars at the city level;
    - It can accurately identify the vehicle, and rarely misses or misreads.
1. Problem
2. Method
3. Experiment
4. Conclusion
2. Method

Introduction to ERI data

- **ERI definition:** Electronic Registration Identification (ERI) based on Radio Frequency Identification (RFID), is an emerging vehicle identification technology to collecting travel data of individual vehicle.

- **ERI principle:** ERI is based on RFID. In ERI system, the collection of ERI data mainly relies on two devices: RFID tags attached to the vehicle windshield and RFID readers deployed on key urban road sections. The tag stores the vehicle registration information, such as unique electronic identification (i.e. EID), vehicle type, usage, etc. When a vehicle passes through an RFID reader, the information in the vehicle’s RFID tag is read and a travel record is generated. Figure 7 shows this process.

![Figure 7. ERI data generation process](image)
## 2. Method

### Introduction to ERI data

- **ERI data format**: As shown in Table I.

<table>
<thead>
<tr>
<th>Field name</th>
<th>Sample of field value</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>RID</td>
<td>R228</td>
<td>Identification of the RFID reader</td>
</tr>
<tr>
<td>EID</td>
<td>838326</td>
<td>Identification of the vehicle</td>
</tr>
<tr>
<td>passtime</td>
<td>2016-03-06 15:15:58</td>
<td>Timestamp</td>
</tr>
<tr>
<td>carType</td>
<td>K33</td>
<td>Vehicle Type, “K33” means a small car</td>
</tr>
<tr>
<td>plateType</td>
<td>02</td>
<td>Plate type, “02” means a compact car</td>
</tr>
<tr>
<td>useProperty</td>
<td>A</td>
<td>Usage of the vehicle, “A” means the vehicle is non-operating.</td>
</tr>
</tbody>
</table>

- **ERI dataset size**: More than 70 million data generated in a week.
2. Method

- **Basic definition based on ERI data**

  - **ERI Record:**
    - ✓ A record is a three-tuple consisting of (EID, RID, Passtime), called \( R \).

  - **ERI Segment:**
    - ✓ A segment is composed of two adjacent Rs with the same EID. It is a six-tuple consisting of (EID, Ot, Dt, ORID, DRID, Interval), called \( Seg \).
    - ✓ “adjacent”: If the EIDs of the two Rs are the same, and the other Rs of this EID do not have a Passtime between the Passtimes of the two, the two Rs are said to be adjacent.

  - **ERI Trajectory:**
    - ✓ The trajectory of the car is consisting of all its Rs or all Segs, called \( Tra \). That means a car has at most one \( Tra \).

  - **ERI Trip:**
    - ✓ The \( Tra \) of the car can be divided into multiple Trips after trajectory segmentation. A Trip is called \( Trip \).

  - **ERI Travel behavior**
    - ✓ The elements of a \( Trip \) are different in importance to individuals. We define the essential part of each \( Trip \) as travel behavior, called \( Tb \).

---

Table II. All Rs for vehicles with Eid = 838326 on March 6, 2016

<table>
<thead>
<tr>
<th>EID</th>
<th>RID</th>
<th>Passtime</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R_1 )</td>
<td>838326 R228</td>
<td>2016-03-06 07:48:58</td>
</tr>
<tr>
<td>( R_2 )</td>
<td>838326 R212</td>
<td>2016-03-06 07:52:13</td>
</tr>
<tr>
<td>( R_3 )</td>
<td>838326 R200</td>
<td>2016-03-06 07:58:35</td>
</tr>
<tr>
<td>( R_4 )</td>
<td>838326 R224</td>
<td>2016-03-06 08:04:14</td>
</tr>
<tr>
<td>( R_5 )</td>
<td>838326 R240</td>
<td>2016-03-06 08:07:56</td>
</tr>
<tr>
<td>( R_6 )</td>
<td>838326 R236</td>
<td>2016-03-06 08:09:24</td>
</tr>
<tr>
<td>( R_{18} )</td>
<td>838326 R227</td>
<td>2016-03-06 20:41:21</td>
</tr>
</tbody>
</table>
2. Method

Basic definition based on ERI data

- Sort all the $R$ of the vehicle 838326 in ascending time order, and then display it on the map, as shown in Figure 8, this is the $Tra$ of vehicle 838326.

- Each line of different colors in the figure represents one $Trip$.

- The process of getting $Trip$ from $Tra$ is called trajectory segmentation [1].

- Figure 9 shows the relationship between $R$, $Seg$, $Tra$, $Trip$.

- Figure 10 shows the relationship between $Tb$ and $Trip$.

---

2. Method

IDENTIFYING COMMUTING PRIVATE CARS BY CLUSTERING IN SPATIAL AND TEMPORAL DIMENSIONS

- **Commuting private cars**: *Tbs* of private cars have spatiotemporal similarity.
  - Spatial similarity: If the set of RIDs in two *Tbs* are similar, we call these two *Tbs* are spatial similar.
  - Temporal similarity:
    \[
    \forall i, \text{\text{\text{Tb}}}_1 \cdot \text{\text{\text{R}}}_i \cdot \text{\text{\text{R}}D} = \text{\text{\text{Tb}}}_2 \cdot \text{\text{\text{R}}}_i \cdot \text{\text{\text{R}}D},
    \]
    \[
    \text{\text{\text{Tb}}}_2 \cdot \text{\text{\text{R}}}_i \cdot \text{\text{\text{Pa}}s\text{\text{\text{st}}\text{\text{\text{time}}}}-\alpha} \leq \text{\text{\text{Tb}}}_1 \cdot \text{\text{\text{R}}}_i \cdot \text{\text{\text{Pa}}s\text{\text{\text{st}}\text{\text{\text{time}}}}} \leq \text{\text{\text{Tb}}}_2 \cdot \text{\text{\text{R}}}_i \cdot \text{\text{\text{Pa}}s\text{\text{\text{st}}\text{\text{\text{time}}}}} + \alpha
    \]

- **The detailed procedures**: The method includes two processes:
  - Extract *Tbs* from *Trips*;
  - Measure spatiotemporal similarity of *Tbs*. 
2. Method

- IDENTIFYING COMMUTING PRIVATE CARS BY CLUSTERING IN SPATIAL AND TEMPORAL DIMENSIONS

Measure spatiotemporal similarity of $T_b$s

- Binary $T_b$ to form binary sequence.
- We use Hamming distance to measure the similarity between binary sequences corresponding to $T_b$. 

Spatial similarity measurement

Temporal similarity measurement
2. Method

IDENTIFYING COMMUTING PRIVATE CARS BY CLUSTERING IN SPATIAL AND TEMPORAL DIMENSIONS

- Measure spatiotemporal similarity of $Tb$s
  - Binary $Tb$ to form binary sequence.
  - We use Hamming distance to measure the similarity between binary sequences corresponding to $Tb$.
  - We use Hamming distance as a distance metric to cluster the $Tb$s of each traveler. The clustering result is a $Tb$ set with spatiotemporal similarity. Then we identify travelers with $Tb$ sets similar in temporal and spatial as commuting private cars.
1. Problem
2. Method
3. Experiment
4. Conclusion
3. Experiment

**ANALYSIS OF THE CHARACTERISTICS OF SPATIOTEMPORAL**

- After experiment, our method identified 215716 commuting private cars from 1082991 private cars. We analyze the travel patterns from the perspectives of individuals and groups.
3. Experiment

ANALYSIS OF THE CHARACTERISTICS OF SPATIOTEMPORAL

- **Individuals:**
  Here shows the 5-day trajectory of one commuter identified by our method.

The traveler’s travel is very regular, and his travels occur in the morning peak. Combining with the actual scene, the origin (Sigongli Interchange) is a residential area, and the destination (Banan Interchange) is a work place.
# 3. Experiment

## ANALYSIS OF THE CHARACTERISTICS OF SPATIOTEMPORAL

### Groups:
- 1082991 private cars are divided into two groups
  - Commuting private cars (CPC)
  - Non-commuting private cars (NCPC)
- The statistical information of the CPC and NCPC.

<table>
<thead>
<tr>
<th>Category</th>
<th>Number of Trips</th>
<th>Average travel time (min)</th>
<th>Average travel distance (km)</th>
<th>Average travel speed (km/h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPC</td>
<td>2293342</td>
<td>28.7</td>
<td>12.67</td>
<td>26.49</td>
</tr>
<tr>
<td>NCPC</td>
<td>4307712</td>
<td>38.7</td>
<td>15.90</td>
<td>24.65</td>
</tr>
</tbody>
</table>

- Temporal characteristics analysis.

![Figure 11. Distributions of departure times and arrival times of CPC](image1)

![Figure 12. Distributions of departure times and arrival times of NCPC](image2)
3. Experiment

- ANALYSIS OF THE CHARACTERISTICS OF SPATIOTEMPORAL

- Groups:
  - ✓ Spatial characteristics analysis.

Figure 13. The hot spot areas distribution at the origins of CPC in rush hours and traffic transfer information in rush hours.

Figure 14. The typical origins and destinations and the land use property map of Chongqing.

According to the definition of commuting, we identify the typical origins in morning rush hours and the typical destinations in evening rush hours as residential land, and identify the typical destinations in morning rush hours and the typical origins in evening rush hours as commercial land. For areas that are both residential land and commercial land, we identify them as Mixed commercial and residential land. The result is shown in Figure 14 (a).
1. Problem
2. Method
3. Experiment
4. Conclusion
4. Conclusion

- Some issues to be solved in the next month

- In summary, we identify commuting private cars based on ERI data.

- We propose a novel method to measure the spatiotemporal similarity of $Tb$s to mine regular travel behavior, then identify private cars with regular travel behavior as commuting private cars.

- According to the result of experiment, we analyze the temporal characteristics of CPC and NCPC at departure time and arrival time.

- Finally, we verify the nature of land use of CPC at origin and destination during rush hours, and it proves that the travel patterns of commuting private cars identified by our proposed method can reflect the residence-workplace relationship.
THANKS