A Constraint-Based Bicycle Origin-Destination Estimation Procedure

Report by: Ziqi Song, Ph.D., Anthony Chen, Ph.D.
Graduate Research Assistant: Zhaocai Liu
Utah State University
Department of Civil and Environmental Engineering
Introduction

• Cycling is an active and green transportation mode

• Origin-destination (O-D) trip matrix estimation plays a key role in transportation planning

• O-D matrix estimation methods developed for motorized traffic cannot be directly used to estimate bicycle O-D trip matrices

• It is critically important to develop an innovative bicycle O-D matrix estimation procedure that utilizes bicycle data from multiple sources
Research Objectives

• To explore bicycle data that are useful for bicycle O-D estimation

• To develop a constraint-based bicycle O-D estimation procedure that utilizes bicycle data from multiple sources

• To conduct a case study using the proposed methodology
Bicycle Data from Multiple Sources

• Bicycle count data
  – Manual count data: easy to collect and flexible, but accuracy is subject to human error/bias and fatigue of the data collector
  – Automated count data: suitable for continuous, long-term bicycle monitoring, but requires significant financial and technical resources

• Emerging crowdsourcing bicycle data
  – Capable of collecting detailed bicycle trip information, such as trip purpose and trip trajectories.
  – Potential issues include sample bias due to participation barriers, demographic bias, and other issues, such as quality control and recruitment

• Bike-sharing data
  – Similar with crowdsourcing data

Source: Siemens (2015)
Methodology

• A constraint-based bicycle O-D estimation procedure that consists of two major stages:
  
  – First generate an efficient bicycle route set that contains a set of Pareto optimal routes
  
  – Then develop a bicycle Path Flow Estimator (PFE) based on the path-size logit (PSL) route choice model
Cyclists’ Route Choice Criteria

• Cyclists’ route choice criteria include:

  – Route distance: consists of the sum of link distance along the route and the sum of turning movement penalties at intersections

  – Route bicycle level of service: serves as a surrogate measure to account for different attributes contributing to the safety of bicycle routes

  – Route pollution: consider carbon monoxide as an important indicator of the level of atmospheric pollution
Multi-Criteria Shortest Path

- A two-phase procedure is adopted to solve the multi-objective shortest path problem

\[
\text{do } rs = 1 \text{ to } RS \\
K^{rs} = \emptyset \quad \text{// Initialize route set} \\
\text{while } \left( Z_{1k}^{rs} \leq \bar{Z}_{1k}^{rs} \right) \\
K^{rs} = K^{rs} \cup k \quad \text{// Generate all possible routes for the first objective} \\
\text{end while} \\
\text{if } \left( z_{2k}^{rs} > z_{2}^{rs} \right) \text{ or } \left( z_{3k}^{rs} > z_{3}^{rs} \right) \\
K^{rs} = K^{rs} - k \\
\text{end if} \quad \text{// Exclude dominated routes by comparing with other objectives} \\
\text{do } n = 1 \text{ to Criteria } # -1 \\
\quad \text{Ascending order with } Z_{1k}^{rs} \\
\quad \bar{K}^{rs} = \{1\} \quad \text{// Initialize efficient route set with the first route} \\
\quad \text{do } k = 2 \text{ to } |K^{rs}| \quad \text{// Update efficient route set with other routes} \\
\quad \quad \text{do } l = 1 \text{ to } |\bar{K}^{rs}| \\
\quad \quad \quad \text{if } \left( z_{n+1,k}^{rs} < z_{n+1,l}^{rs} \right) \\
\quad \quad \quad \quad \bar{K}^{rs} = \bar{K}^{rs} \cup k \\
\quad \quad \quad \text{else } \bar{K}^{rs} = \bar{K}^{rs} - \{k\} \\
\quad \quad \quad \text{end if} \\
\quad \quad \text{end do} \\
\quad \text{end do} \\
\text{end do}
Bicycle Path Flow Estimator

• Developed based on the path-size logit (PSL) route choice model
  – Advantage: PSL can handle the route overlapping problem, account for the total route cost values, and has an economic interpretation
  – Disadvantage: Requires survey data to calibrate the parameters

• Formulated as a convex program

• Solved using an iterative balancing scheme
Case Study

- Conducted based on the real-world data in Salt Lake City, Utah

- Correlation analysis between bicycle count data and Strava data is conducted

- Bicycle O-D estimation is conducted in a sub-network
Bicycle Count Data

- Data statistics

<table>
<thead>
<tr>
<th>Statistic</th>
<th>All Counts</th>
<th>Weekday</th>
<th>Weekend</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of counts</td>
<td>95</td>
<td>57</td>
<td>38</td>
</tr>
<tr>
<td>Minimum</td>
<td>2</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Maximum</td>
<td>161</td>
<td>129</td>
<td>161</td>
</tr>
<tr>
<td>Median</td>
<td>47</td>
<td>47</td>
<td>42</td>
</tr>
<tr>
<td>Mean</td>
<td>54.8</td>
<td>54.1</td>
<td>55.9</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>35.7</td>
<td>31.3</td>
<td>41.8</td>
</tr>
</tbody>
</table>

- Count Location

Map showing count locations.
Strava Data
Correlation Analysis

Regression analysis shows a positive correlation between Strava Annual Bicycle Rides and Bicycle Counts. The coefficient of determination ($R^2$) is 0.6215, indicating that 62.15% of the variation in Strava Annual Bicycle Rides can be explained by the variation in Bicycle Counts.
Bicycle Origin-Destination Estimation

- Study area
Bicycle Origin-Destination Estimation

- Estimated zonal bicycle production and attraction flows
Bicycle Origin-Destination Estimation

- Comparison between estimated flows and Strava flows

![Graph showing comparison between estimated link flows and Strava link flows with an R2 value of 0.95.](image)
Discussion

• Bicycle count data has limited data points
• Other large-scale datasets are needed for further validation between the correlation between the count data and the Strava data
• The Strava data has the potential issue of biases
• Crowdsourced data may serve as an economical complement to official count programs
• The PFE model can effectively employ the bicycle count data and the Strava data to estimate path flow and O-D flow
Concluding Remarks

• Bicycle data that are useful for bicycle O-D estimation are explored.

• A two-stage constraint-based bicycle O-D estimation procedure is proposed. The procedure is flexible and can be adjusted to different levels of data availability and quality.

• A real-world case study was conducted in Salt Lake City, Utah to demonstrate the proposed methodology.