A GPS-based Bicycle Route Choice Model for San Francisco, California

Jeff Hood
University of California, Berkeley

Billy Charlton, Elizabeth Sall, Michael Schwartz
San Francisco County Transportation Authority

Matt Paul
MoPimp Productions

SAN FRANCISCO COUNTY TRANSPORTATION AUTHORITY
3rd Conference on Innovations in Travel Demand Modeling, Tempe, Arizona
May 10, 2010
Introduction

CycleTracks for iPhone & Android

Data processing & participants

Choice set generation

Model estimation & validation

Trip assignment & next steps
Why model bicycle route choice?
CycleTracks for iPhone & Android
Conventional GPS survey
CycleTracks smartphone app
Cycle Tracks

Powered by The San Francisco County Transportation Authority
Thanks for using CycleTracks! Please enter your user details here. It's optional, anonymous, and will really help us understand different people's biking preferences.

Age

29

Male Female

Cycling Frequency: Several times per week

Home ZIP Work ZIP School ZIP

94110 94102 School

Email address: for news and updates!!
cycletracks@sfcta.org
The primary reason for this bike trip is going to or from a social activity (e.g. at a friend's house, the park, a restaurant, the movies).
View Saved Trips

Shopping: 0m
Oct 27, 2009 10:12:56 AM
(rerecording in progress)

Social: 2806m
Oct 22, 2009 7:35:36 AM
(trip saved & uploaded)

Commute: 0m
Oct 21, 2009 9:38:14 AM
(trip saved & uploaded)

Commute: 2481m
Oct 21, 2009 7:59:27 AM
(trip saved & uploaded)
OK, it’s real pretty; 
Now, how do we get users?
SF Transportation Authority Launches iPhone App to Track Cyclists

by Matthew Roth on November 12, 2009

The San Francisco County Transportation Authority (TA), the city's congestion management agency responsible for modeling transportation and development patterns, has released its new bicycle route data application, Cycle Tracks, for iPhones and GPS-enabled iTunes players at the iTunes store. Like similar applications that give information such as speed and distance traveled, users of the TA app can map their bicycle ride, but the data they collect will be aggregated anonymously in the TA's server so that it can be applied to their SF-CHAMP modeling and travel forecasting tool.

"This app will help the cycling community help itself," TA Executive Director José Luis Moscovich said in a statement. "The data they log will contribute to better planning of bicycle facilities, and they'll also have a record of their personal cycling history. I'm sure it will be very popular."

Billy Charleton, Deputy Director for Technology Services at the TA, explained that SF-CHAMP doesn't currently have concise and reliable trip data for cyclists, but that they rely on static counts at various intersections conducted once or twice a year. Without understanding the entire length of a trip, nor the trip purpose, the agency is unable to analyze what cyclists prefer in terms of street characteristics, including..."
Tim Hickey

Wrong location spot
When I ride home at Howard and Spear, it sometimes puts my location in the bay, just north of the bay bridge. This seems to nullify my recording.

January 10 at 7:50pm · Participate

Recent activity

Jennifer Gile discussed Cycletracks on Android on the CycleTracks discussion board.

Dave Mangot discussed Cycletracks on Android on the CycleTracks discussion board.

Billy Charlton and Dave Mangot discussed Cycletracks on Android on the CycleTracks discussion board.

Dave Mangot

Cycletracks on Android
I tried recording two different trips on my Samsung Moment (cupcake) and in both cases, it said 0 miles. After that I discovered that after I hit Start Trip it says elapsed time 1 second, and that's the end. Never advances past one second. (and thus, my distance, etc...)

December 16, 2009 at 10:13am · Participate

Jeffrey Carl Faden
Data Processing & Participants
GPS post-processing

5,178 traces
497 users

\[ \exp \left( -\frac{(t - t_j)^2}{2\sigma^2} \right) \]

Gaussian smoothing

Activity & mode detection

Map matching

3,034 bike stages
366 users

(Schüssler & Axhausen 2009)
## Participants

<table>
<thead>
<tr>
<th>CycleTracks (N=366)</th>
<th>BATS (N=153)</th>
<th>z-stat</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>34</td>
<td>33</td>
<td>1.1</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>21%</td>
<td>36%</td>
<td>-3.5</td>
</tr>
<tr>
<td><strong>Cycling Frequency</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daily</td>
<td>60%</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Several times per week</td>
<td>34%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Several times per month</td>
<td>7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than once a month</td>
<td>0%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Choice Set Generation
Existing methods

Only shortest path searches work in large networks

- Link elimination
  \((k\) shortest paths)\)

- Stochastic path search

- Labeled paths

Doubly stochastic
(Bovy & Fiorenzo-Catalano 2007)
The doubly stochastic method

1. Select random coefficient vector
2. Calculate generalized cost for each link
3. Randomize link costs
4. Find shortest path
Extract unbiased priors from the network

Attribute Space

Parameter Space

Value of evaluation attribute in shortest path

Coefficient of evaluation attribute
## Significant in estimation

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Initial Interval</th>
<th>Final Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta_{i,L}$</td>
<td>$\beta_{i,H}$</td>
</tr>
<tr>
<td>Length (reference)</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Length off bike paths</td>
<td>$1.0 \times 10^{-7}$</td>
<td>1,000.</td>
</tr>
<tr>
<td>Length off bike lanes</td>
<td>$1.0 \times 10^{-7}$</td>
<td>1,000.</td>
</tr>
<tr>
<td>Length off bike routes</td>
<td>$1.0 \times 10^{-7}$</td>
<td>1,000.</td>
</tr>
<tr>
<td>Length $\times$ up-slope (ft/100 ft)</td>
<td>$1.0 \times 10^{-7}$</td>
<td>1,000.</td>
</tr>
<tr>
<td>Length wrong way</td>
<td>$1.0 \times 10^{-7}$</td>
<td>1,000.</td>
</tr>
<tr>
<td>Number of turns*</td>
<td>$1.0 \times 10^{-7}$</td>
<td>1,000.</td>
</tr>
<tr>
<td>Length $\times$ daily traffic (1,000s)</td>
<td>$1.0 \times 10^{-7}$</td>
<td>1,000.</td>
</tr>
</tbody>
</table>

*Extracted with other coefficients at their medians

## Retained for degree of freedom
96 Link Elimination Routes

96 Doubly Stochastic Routes
Cumulative Distribution of Maximum Overlap with Chosen Route

Proportion of Observations < Overlap

- Link elimination
- Doubly stochastic

Overlap

Better
Minimum Dissimilarity from Attributes of Chosen Route when Overlap < 100%

- Link elimination
- Doubly stochastic

Frequency

Minimum Normalized Euclidean Distance

Better
### Benchmarking

<table>
<thead>
<tr>
<th>Method</th>
<th>Doubly Stochastic</th>
<th>Link Elimination</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. parameters</td>
<td>32</td>
<td>–</td>
</tr>
<tr>
<td>No. link randomizations</td>
<td>3</td>
<td>–</td>
</tr>
<tr>
<td>No. unique routes</td>
<td>76 (avg.)</td>
<td>96</td>
</tr>
<tr>
<td>Search algorithm</td>
<td>Dijkstra</td>
<td>Euclidean A*</td>
</tr>
<tr>
<td>Computing time</td>
<td>3 h 46 m</td>
<td>8 h 06 m</td>
</tr>
</tbody>
</table>

2,678 observations, 4 CPUs, coded in Python
Model Estimation & Validation

<table>
<thead>
<tr>
<th>$t$-stat.</th>
<th>$p$-val.</th>
</tr>
</thead>
<tbody>
<tr>
<td>-11.80</td>
<td>0.00</td>
</tr>
<tr>
<td>-12.15</td>
<td>0.00</td>
</tr>
<tr>
<td>-19.87</td>
<td>0.00</td>
</tr>
<tr>
<td>6.17</td>
<td>0.00</td>
</tr>
</tbody>
</table>
## Estimation results

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Coef.</th>
<th>SE</th>
<th>t-stat.</th>
<th>p-val.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length (mi)</td>
<td>−1.05</td>
<td>0.09</td>
<td>−11.80</td>
<td>0.00</td>
</tr>
<tr>
<td>Turns per mile</td>
<td>−0.21</td>
<td>0.02</td>
<td>−12.15</td>
<td>0.00</td>
</tr>
<tr>
<td>Prop. wrong way</td>
<td>−13.30</td>
<td>0.67</td>
<td>−19.87</td>
<td>0.00</td>
</tr>
<tr>
<td>Prop. bike paths</td>
<td>1.89</td>
<td>0.31</td>
<td>6.17</td>
<td>0.00</td>
</tr>
<tr>
<td>Prop. bike lanes</td>
<td>2.15</td>
<td>0.12</td>
<td>17.69</td>
<td>0.00</td>
</tr>
<tr>
<td><em>Cycling freq. &lt; several per wk.</em></td>
<td>1.85</td>
<td>0.04</td>
<td>44.94</td>
<td>0.00</td>
</tr>
<tr>
<td>Prop. bike routes</td>
<td>0.35</td>
<td>0.11</td>
<td>3.14</td>
<td>0.00</td>
</tr>
<tr>
<td>Avg. up-slope (ft/100ft)</td>
<td>−0.50</td>
<td>0.08</td>
<td>−6.35</td>
<td>0.00</td>
</tr>
<tr>
<td><em>Female</em></td>
<td>−0.96</td>
<td>0.22</td>
<td>−4.34</td>
<td>0.00</td>
</tr>
<tr>
<td><em>Commute</em></td>
<td>−0.90</td>
<td>0.11</td>
<td>−8.21</td>
<td>0.00</td>
</tr>
<tr>
<td>Log(path size)</td>
<td>1.07</td>
<td>0.04</td>
<td>26.38</td>
<td>0.00</td>
</tr>
</tbody>
</table>

2,678 weighted observations, $\rho^2 = 0.28$
### Average marginal rates of substitution

<table>
<thead>
<tr>
<th>MRS of Length on street for</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length on bike paths</td>
<td>0.57</td>
<td>none</td>
</tr>
<tr>
<td>Length on bike lanes</td>
<td>0.49</td>
<td>none</td>
</tr>
<tr>
<td>Length on bike routes</td>
<td>0.92</td>
<td>none</td>
</tr>
<tr>
<td>Length wrong way</td>
<td>4.02</td>
<td>none</td>
</tr>
<tr>
<td>Turns</td>
<td>0.10</td>
<td>mi/turn</td>
</tr>
<tr>
<td>Total rise</td>
<td>1.12</td>
<td>mi/100ft</td>
</tr>
</tbody>
</table>

**User benefit of bike lanes:** $0.98 per mile per trip
Cumulative Distribution of Holdback Sample Prediction's Overlap with Chosen Route

15% exact predictions
Holdback Sample Probability-Weighted Dissimilarity from Attributes of Chosen Route

Frequency
0 50 100 150

Normalized Euclidean Distance

Better
Calibrated choice set

Best generated
Overlap: 85%
Dissimilarity: 0.23
PS-normed prob: 0.022

Predicted
Overlap: 5%
Dissimilarity: 0.83
PS-normed prob: 0.025
Trip Assignment & Next Steps
AM Assignment
Computing Time: 12 h 35 m

Bicycles per hour

- 0
- 180
- 20
- 360
Next steps for SF-CHAMP

Bike Route Choice Set Generation

Population Synthesizer

Initial Road Assignment

Roadway & Transit Skimming

Work Loc. & Destination Choice

Full Day Tour Generation

Tour & Trip Mode Choice

Final Road & Transit Assignment

Final Bicycle Assignment

Vehicle Availability

Bicycle Availability

Logsums

Logsums
Acknowledgements

Nadine Schüssler

Kay Axhausen

Contact

Jeff Hood, UC Berkeley: jeffnhood@gmail.com

Billy Charlton, SFCTA: billy.charlton@sfcta.org

Matt Paul, MoPimp Productions: www.mopimp.com
Recap

Introduction

CycleTracks for iPhone & Android

Data processing & participants

Choice set generation

Model estimation & validation

Trip assignment & next steps