



# Bus Passenger Origin- Destination Matrix Estimation

---

Application to CTA Bus Network

Alex Cui

Massachusetts Institute of Technology



# Presentation Outline

---

- Introduction and Objectives
- Review of prior work
  - Iterative proportional fitting
  - Trip-chaining method of estimating OD matrix
- Rail OD matrix estimation for CTA
- Current project – bus system OD matrix
  - Project approach
  - Preliminary results
  - Remaining Issues



# Introduction

---

- The benefit of Origin-Destination matrix:
  - Service planning
  - Operations analysis
    - Example: scheduling
  - Impact analysis—before-and after analysis
- Traditional way to obtain OD matrix: on-board survey combined with manual boarding and alighting counts
- The use of Automatically Collected Data (ACD) to get OD matrix
  - Bigger sample size
  - Cheaper, since the data are available for other purposes already
  - Faster, and can be done more frequently
  - Can be combined with more targeted surveys to get more cost effective results



# Objectives of Research

---

- Long term goal: inter-modal (bus-rail) passenger OD matrix
- Short term goal: estimate bus passenger OD matrix using available CTA data  
(Easy to apply procedures)



# Scope of the research

---

- Single bus route application
- Network level application



# Relevant prior work

- Peter Furth: Iterative Proportional Fitting (IPF) method for single bus route OD matrix estimation

Example:

Route #1 (example)		Destination				
		Stop 1	Stop 2	Stop 3	Stop 4	On count
Origin	Stop 1	0	?	?	?	40
	Stop 2	-	0	?	?	30
	Stop 3	-	-	0	?	20
	Stop 4	-	-	-	0	0
Off count		0	30	20	40	



# Relevant prior work

---

- IPF method findings
  - The quality of the seed matrix is important to the resulting OD matrix  
(Seed matrix contains the information from a small number of trips)



## Relevant prior work (2)

---

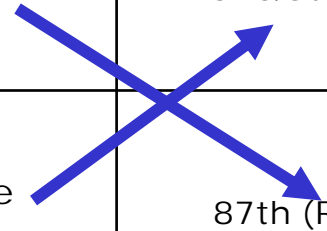
- Adam Rahbee and Jinhua Zhao:  
Trip-Chain Rail Passenger OD Matrix Inference
  - Network level
  - Based on AFC data

# Relevant prior work (2)

- Trip-Chain Rail Network OD Matrix Inference
  - Example:

Passenger  
A

Day	Time	Mode	Rail Origin Station	Inferred Destination Station
MON	8:22:15	Red Line	87th (Red)	Lake/State (Red)
MON	16:00:32	Red Line	Lake/State (Red)	87th (Red)



# Review of OD Matrix for Rail System

- Completed based on Jinhua Zhao's algorithm
- High inference rate

	Results on Sep '04 Data	
Inference Method	Number of Rail trips	Percentage Contribution
Next Train	864646	33.6%
Last Train of the Day	541809	21.1%
AVL	335254	13.0%
GIS	36695	1.4%
Multi-Swipe and other methods	53457	2.0%
Successful Inference	1831861	71.2%
Total	2571140	100.00%



# Bus vs. Rail comparison

	<b>Rail</b>	<b>Bus</b>
<b>Data available</b>	<b>AFC, Entry counts</b>	<b>APC, AFC, AVL</b>
<b>AFC coverage</b>	<b>91%* (currently almost 100%)</b>	<b>About 60%* (currently about 80% and growing)</b>
<b>APC availability</b>	<b>None</b>	<b>&gt;15%</b>
<b>Trip origin (Entry point)</b>	<b>Known</b>	<b>Need AVL data (or GIS data), 70-80% match rate</b>

\*The percentages are based on 2004 data

# Bus OD Matrix

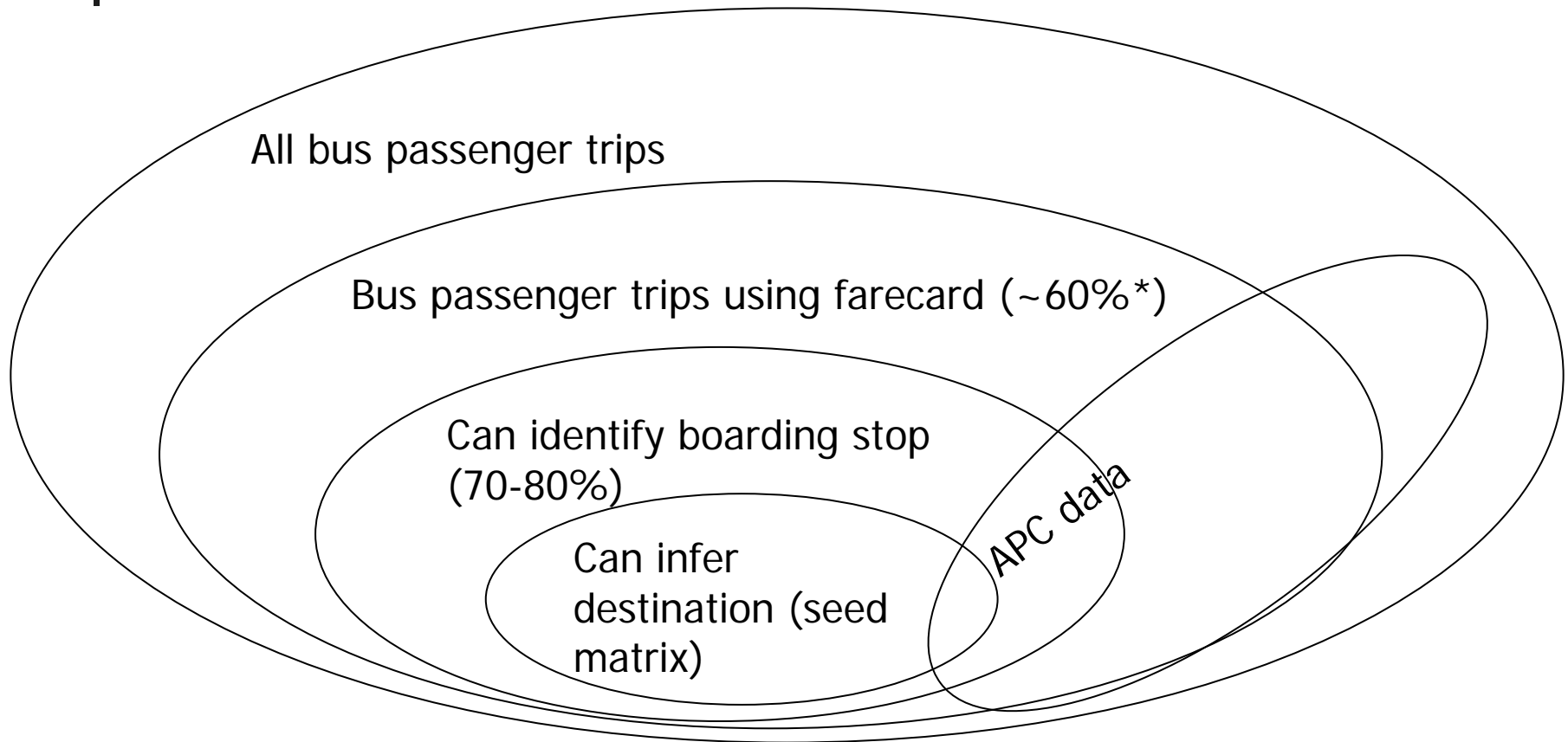
## --Data Issues

	<b>Issues</b>
<b>APC</b>	<ol style="list-style-type: none"><li><b>1. Under-count passenger boardings and alightings</b></li><li><b>2. APC coverage</b></li></ol>
<b>AFC</b>	<ol style="list-style-type: none"><li><b>1. Does not have location data</b></li><li><b>2. Only about 60%* of bus passenger trips use farecard</b></li></ol>

\*The percentages are based on 2004 data

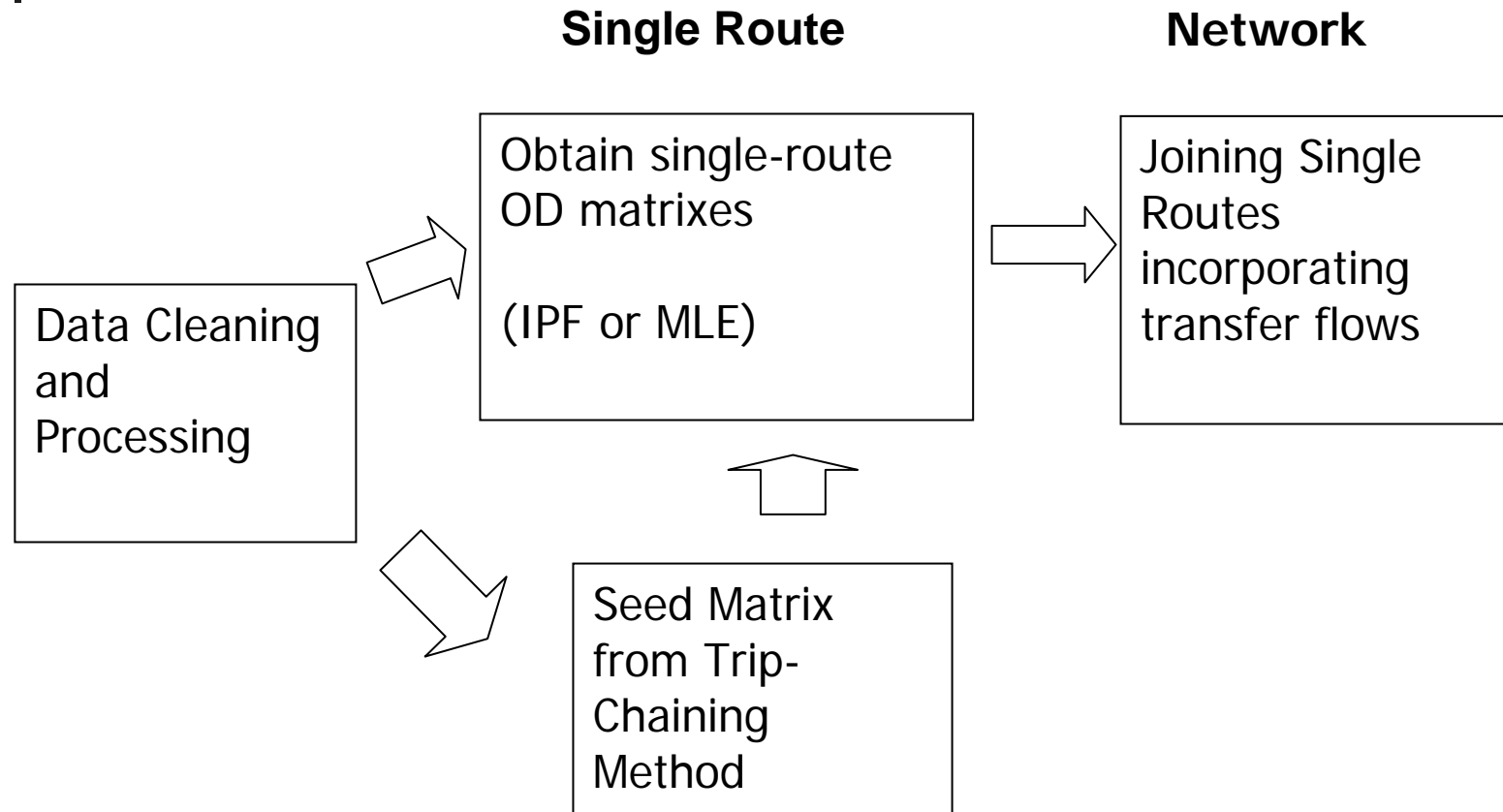
# Bus OD Matrix

## --Description of data sources



\*The percentages are based on 2004 data

# Proposed approach --Overview





# Proposed approach

## --Route-level OD matrix

---

- Seed matrix generation primarily based on AFC data
  - Obtain boarding stop
    - AVL matching
    - rail-bus transfer, GIS
    - bus-bus transfer, GIS
  - Trip-chain method
  - Seed matrix
- Iterative Proportional Fitting (IPF) or Maximum Likelihood Estimation (MLE)
  - These are two techniques to combine APC data with seed matrix to get full OD matrix

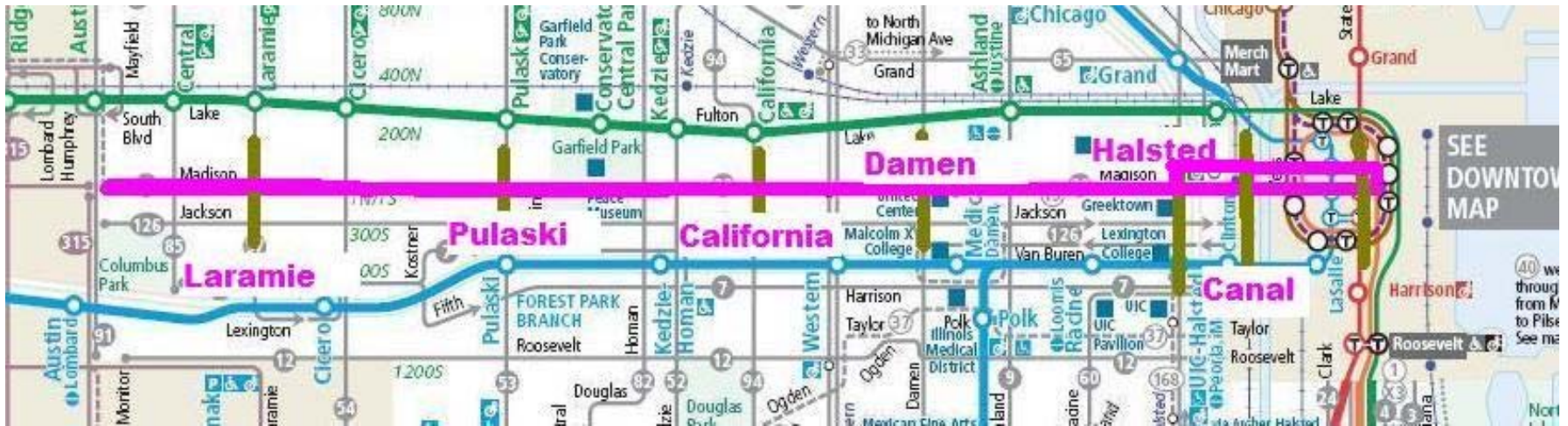


# Summary Information on the data used

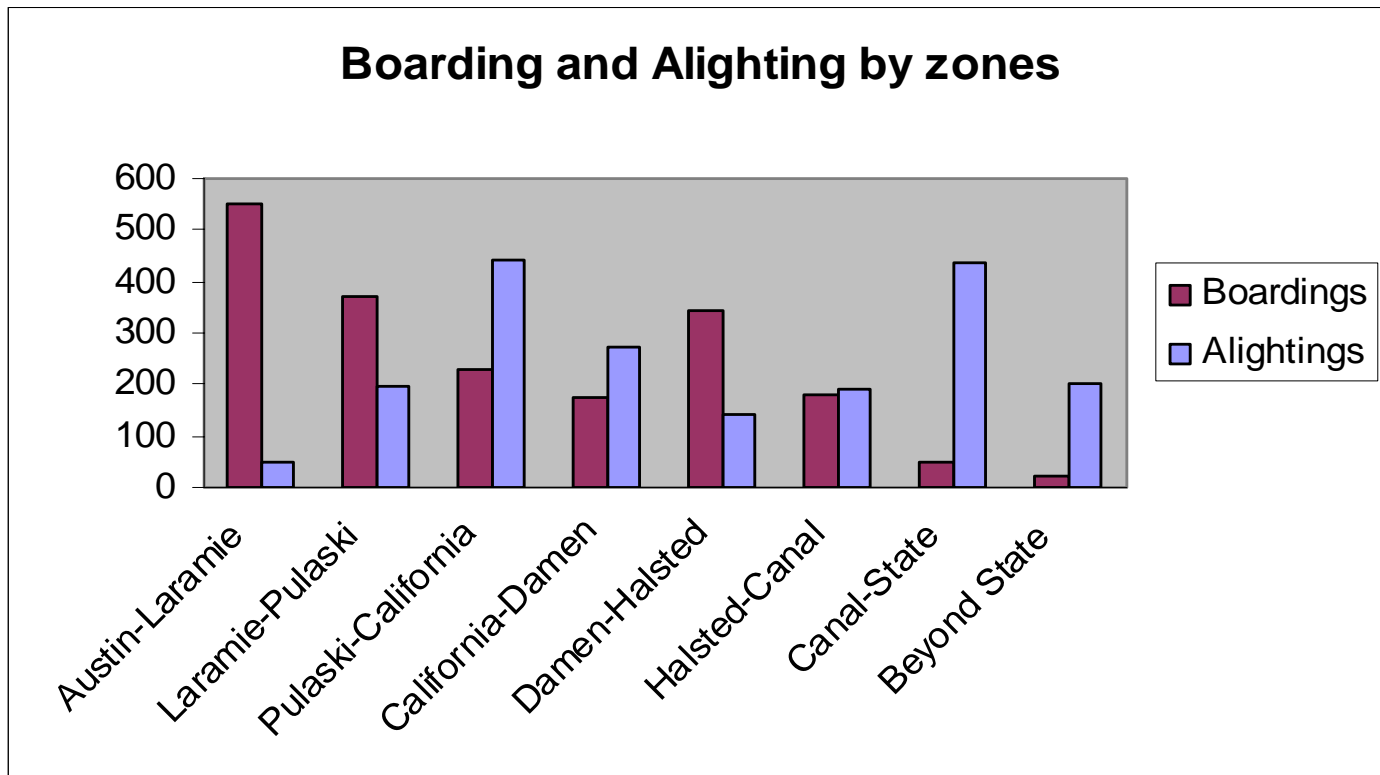
---

- Overall AFC data (Whole day, all routes):
  - 545,000 bus passenger trips using farecard
  - From these, 436,000 with boarding stop (~80% identification rate)
  - From these, 244,000 with destinations (~56% inference rate)

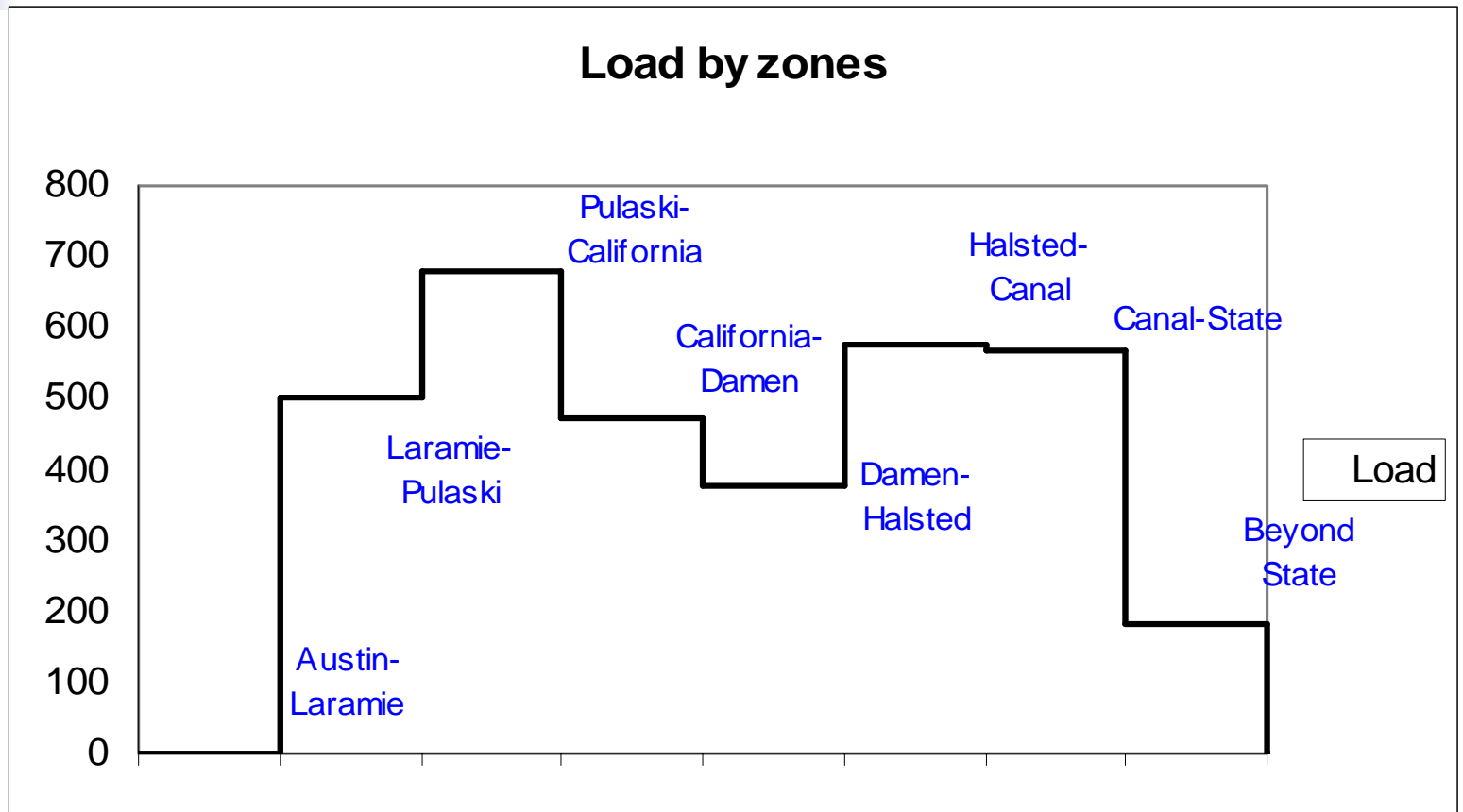
# Route 20 Map



# Route 20 Boarding and Alighting Eastbound morning peak



# Load profile for route 20 Eastbound morning peak



# Preliminary Results

## --#20 Eastbound Morning Peak

		To: → (East)							Beyond State
		Austin-Laramie (1 mile)	Laramie-Pulaski (1.5 mile)	Pulaski-California (1.5 mile)	California-Damen (1 mile)	Damen-Halsted (1.5 mile)	Halsted-Canal (0.5 mile)	Canal-State (0.5 mile)	
From	Austin - Laramie	47	<b>170</b>	<b>155</b>	65	11	30	71	0
	Laramie - Pulaski		24	<b>244</b>	72	10	6	13	4
	Pulaski - California			40	<b>113</b>	31	13	30	4
	California - Damen				24	40	10	<b>101</b>	0
	Damen - Halsted					50	<b>123</b>	74	97
	Halsted - Canal						7	<b>126</b>	46
	Canal - State							18	30
	Beyond State								21

# Preliminary Results

## --#20 Westbound Morning Peak

		To: → (Westbound)							
		Beyond State	Canal-State (0.5 mile)	Halsted-Canal (0.5 mile)	Damen-Halsted (1.5 mile)	California-Damen (1 mile)	Pulaski-California (1.5 mile)	Laramie-Pulaski (1.5 mile)	Austin-Laramie (1 mile)
From:	Beyond State	45	86	<b>149</b>	<b>180</b>	48	32	80	77
	Canal - State		47	<b>130</b>	<b>249</b>	10	7	21	37
	Halsted - Canal			2	94	4	13	12	4
	Damen - Halsted				15	40	24	43	10
	California - Damen					0	84	61	4
	Pulaski - California						28	<b>211</b>	<b>105</b>
	Laramie - Pulaski							22	87
	Austin - Laramie								35

# Compare with Stacy Schwartz's previous results

## -- #20 Eastbound 7-9:15am Nov 2003

My new result (Stacy's previous result)

OD flow as a percentage of total passenger trips

	Zones	Austin-Pulaski	Pulaski-Kedzie	Kedzie-Western	Western-Ashland	Ashland-Halsted	Halsted-Columbus	Total Boardings
Origins	Austin-Pulaski	13% (13%)	17% (14%)	5% (4%)	6% (9%)	0% (1%)	6% (6%)	48% (46%)
	Pulaski-Kedzie		2% (1%)	3% (1%)	2% (3%)	0% (1%)	1% (3%)	8% (9%)
	Kedzie-Western			1% (0%)	3% (2%)	1% (0%)	3% (3%)	7% (5%)
	Western-Ashland				1% (1%)	0% (1%)	7% (7%)	8% (9%)
	Ashland-Halsted					3% (0%)	13% (15%)	16% (16%)
	Halsted-Columbus						13% (14%)	13% (14%)
	Total Alightings	13% (13%)	19% (15%)	9% (5%)	12% (15%)	4% (4%)	43% (48%)	100% (1923 trips)
							100% (1761 trips)	



# Remaining issues (for the following semester)

---

- Formulations and algorithm of Maximum Likelihood Estimation
- Network level incorporating transfer information



---

- Thank You