New York City
Subway Ridership
Time of Day Pattern

By Jason(Xiaoqiang) Chen

Faculty advisor: Professor Cynthia Chen, CCNY
Professional advisor: Mr. James Barry, New York City Transit
Outline

• Motivations and research questions
• Datasets
• Ridership spatial distribution
• Ridership temporal distribution
• Built environment and temporal distribution
• Conclusions and future work
Motivations and research questions

Motivations
✓ Ridership projection
✓ Transit operation
✓ Subway station planning

Research questions
✓ How is ridership spatially distributed?
✓ How is ridership temporally distributed?
✓ What factors contribute to ridership temporal distribution?
Datasets

- Ridership
- Transfer ridership (bus to subway)
- Subway station and bus station location
- New York City street map files.
- Land use
- SES. data
- Employment information
- General travel cost to three zones in CBD
Ridership spatial distribution
Ridership spatial distribution

Weekday

Saturday

Sunday

Times Square
Randomly Selected Ridership Temporal Distribution

135 St (2,3)

14 St (F,V,1,2,3)/6 Av (L)

161 St – Yankee Stadium (B,D,4)

Canal St (A,C,E)
Ridership Temporal Distribution – Cluster analysis

• **Concentration of ridership**
  - Morning peak hourly ridership
  - Afternoon peak hourly ridership
  - Early morning hourly ridership
  - Midday hourly ridership

• **Transfer ridership ratio (bus to subway)**
  - Morning peak transfer ridership ratio
  - Afternoon peak transfer ridership ratio
  - Early morning transfer ridership ratio
  - Midday transfer ridership ratio

• **Position of Concentration in time**
  - Morning peak starting time
  - Morning peak duration
  - Afternoon peak starting time
  - Afternoon peak duration

• **Relative magnitude of ridership**
  - Ratio of morning peak hourly over total daily ridership
  - Ratio of afternoon peak hourly over total daily ridership
  - Ratio of early morning hourly over total daily ridership
  - Ratio of midday hourly over total daily ridership
  - Ratio of morning peak hourly over maximum hourly ridership
  - Ratio of early morning hourly over maximum hourly ridership
  - Ratio of midday hourly over maximum hourly ridership
Ridership Temporal Distribution – Cluster analysis

High Afternoon Peak Pattern

- 64 stations
- Total daily: 32,137
- Morning peak hourly riders: 2,309
- Afternoon peak hourly riders: 4,514
- Early morning hourly riders: 315
- Midday hourly riders: 1,707
- A.M. peak starting time: 8:12 am.
- A.M. peak duration: 62 minutes.
- P.M. peak starting time: 4:53 pm.
- P.M. peak duration: 77 minutes.
Ridership Temporal Distribution – Cluster analysis

High Morning Peak Pattern

- 123 stations
- Total daily: 8,196
- Morning peak hourly riders: 1,068
- Afternoon peak hourly riders: 707
- Early morning hourly riders: 143
- Midday hourly ridership: 447
- A.M. peak starting time: 7:40 am.
- A.M. peak duration: 61 minutes
- P.M. peak starting time: 4:10 pm.
- P.M. peak duration: 82 minutes
Ridership Temporal Distribution – Cluster analysis

No Morning Peak Pattern

- 12 stations
- Total daily: 24,061
- Morning peak hourly riders: no morning peak
- Afternoon peak hourly riders: 4,961
- Early morning hourly riders: 164
- Midday hourly riders: 1,461
- A.M. peak starting time: ~
- A.M. peak duration: ~
- P.M. peak starting time: 4:52pm.
- P.M. peak duration: 68 minutes
Ridership Temporal Distribution – Cluster analysis

Low Ridership Pattern

- 167 stations
- Total daily: 5,294
- Morning peak hourly riders: 1,024
- Afternoon peak hourly riders: 297
- Early morning hourly riders: 137
- Midday hourly riders: 280
- A.M. peak starting time: 7:26 am.
- A.M. peak duration: 62 minutes
- P.M. peak starting time: 3:39 pm.
- P.M. peak duration: 97 minutes
Ridership Temporal Distribution – Cluster analysis

High Transfer Pattern

• 57 stations
• Total daily: 12,100
• Morning peak hourly ridership: 2,113
• Afternoon peak hourly ridership: 724
• Early morning hourly ridership: 298
• Midday hourly ridership: 629
• A.M. peak starting time: 7:25 am.
• A.M. peak duration: 63 minutes
• P.M. peak starting time: 3:47 pm.
• P.M. peak duration: 97 minutes
• Transfer ridership (>25%)
## Ridership Temporal Distribution – Pattern Statistic

<table>
<thead>
<tr>
<th>Patterns</th>
<th>High Afternoon peak</th>
<th>High Morning peak</th>
<th>No Morning peak</th>
<th>Low Ridership</th>
<th>High Transfer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of stations</td>
<td>64</td>
<td>123</td>
<td>12</td>
<td>167</td>
<td>57</td>
</tr>
<tr>
<td><strong>Total Volume</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total daily ridership</td>
<td>32,137</td>
<td>8,196</td>
<td>24,061</td>
<td>5,294</td>
<td>12,100</td>
</tr>
<tr>
<td><strong>Concentration of the Volume</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morning peak dummy (whether there is peak or not)</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Afternoon peak dummy (whether there is peak or not)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Morning peak ridership volume (average)</td>
<td>2,331</td>
<td>1,076</td>
<td>0</td>
<td>1,058</td>
<td>2,211</td>
</tr>
<tr>
<td>Morning peak hourly ridership (average)</td>
<td>2,309</td>
<td>1,068</td>
<td>625</td>
<td>1,024</td>
<td>2,113</td>
</tr>
<tr>
<td>Ratio of morning peak hourly over total daily</td>
<td>0.06</td>
<td>0.13</td>
<td>0.03</td>
<td>0.20</td>
<td>0.18</td>
</tr>
<tr>
<td>Afternoon peak ridership volume (average)</td>
<td>5,823</td>
<td>1,019</td>
<td>5,086</td>
<td>490</td>
<td>1,284</td>
</tr>
<tr>
<td>Afternoon peak hourly ridership (average)</td>
<td>4,514</td>
<td>707</td>
<td>4,691</td>
<td>297</td>
<td>724</td>
</tr>
<tr>
<td>Early morning hourly ridership (average)</td>
<td>315</td>
<td>143</td>
<td>164</td>
<td>137</td>
<td>298</td>
</tr>
<tr>
<td>Midday hourly ridership (average)</td>
<td>1,707</td>
<td>447</td>
<td>1,461</td>
<td>280</td>
<td>629</td>
</tr>
<tr>
<td>Ratio of afternoon peak hourly over total daily</td>
<td>0.15</td>
<td>0.09</td>
<td>0.18</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>Ratio of morning peak hourly over afternoon peak hourly</td>
<td>0.49</td>
<td>1.66</td>
<td>0.18</td>
<td>3.78</td>
<td>3.00</td>
</tr>
<tr>
<td>Ratio of early morning hourly over daily hourly</td>
<td>0.21</td>
<td>0.44</td>
<td>0.17</td>
<td>0.65</td>
<td>0.59</td>
</tr>
<tr>
<td>Ratio of early morning hourly over maximum peak hourly</td>
<td>0.06</td>
<td>0.13</td>
<td>0.04</td>
<td>0.14</td>
<td>0.14</td>
</tr>
<tr>
<td>Ratio of midday hourly ridership over daily hourly</td>
<td>1.30</td>
<td>1.33</td>
<td>1.55</td>
<td>1.26</td>
<td>1.24</td>
</tr>
<tr>
<td>Ratio of midday hourly ridership over maximum peak hourly</td>
<td>0.37</td>
<td>0.41</td>
<td>0.39</td>
<td>0.27</td>
<td>0.31</td>
</tr>
</tbody>
</table>
## Ridership Temporal Distribution – Pattern Statistic

<table>
<thead>
<tr>
<th>Patterns</th>
<th>High Afternoon peak</th>
<th>High Morning peak</th>
<th>No Morning peak</th>
<th>Low Ridership</th>
<th>High Transfer</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Transfer Ridership</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ratio of morning peak transfer ridership over morning peak hourly ridership</td>
<td>0.05</td>
<td>0.07</td>
<td>0.02</td>
<td>0.04</td>
<td><strong>0.25</strong></td>
</tr>
<tr>
<td>Ratio of afternoon peak transfer ridership over afternoon peak hourly ridership</td>
<td>0.02</td>
<td>0.06</td>
<td>0.01</td>
<td>0.05</td>
<td><strong>0.24</strong></td>
</tr>
<tr>
<td>Ratio of early morning transfer ridership over early morning hourly ridership</td>
<td>0.07</td>
<td>0.10</td>
<td>0.04</td>
<td>0.05</td>
<td><strong>0.38</strong></td>
</tr>
<tr>
<td>Ratio of midday transfer ridership over midday hourly ridership</td>
<td>0.03</td>
<td>0.06</td>
<td>0.01</td>
<td>0.04</td>
<td><strong>0.25</strong></td>
</tr>
</tbody>
</table>

| **Position of the Concentration** | | | | | |
| Morning peak starting time (am.) | 8:12 | 7:40 | ~ | 7:26 | 7:25 |
| Afternoon peak starting time (pm.) | 4:53 | 4:10 | 4:52 | 3:39 | 3:47 |
| Morning peak duration (minutes) | 62 | 61 | 0.00 | 62 | 63 |
| Afternoon peak duration (minutes) | 77 | 82 | 68 | 97 | 97 |
Built environment and temporal distribution

• Discrete choice analysis
• Dependent variables: five identified patterns
• Independent variables:
  ➢ Local features (within 500 meters radius)
    ✓ Socio-economic and demographic
    ✓ Land use
    ✓ Employment
    ✓ Street density
    ✓ Bus station density
  ➢ General travel cost to three zones in CBD
### Built environment and temporal distribution:
Location Statistic associated with each pattern

<table>
<thead>
<tr>
<th>Patterns</th>
<th>High Afternoon peak</th>
<th>High Morning peak</th>
<th>No Morning peak</th>
<th>Low Ridership</th>
<th>High Transfer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Stations</td>
<td>64</td>
<td>123</td>
<td>12</td>
<td>167</td>
<td>57</td>
</tr>
<tr>
<td><strong>S.E.D.</strong></td>
<td><strong>mean</strong></td>
<td><strong>mean</strong></td>
<td><strong>mean</strong></td>
<td><strong>mean</strong></td>
<td><strong>mean</strong></td>
</tr>
<tr>
<td>Population</td>
<td>15516</td>
<td>17562</td>
<td>7512</td>
<td>16260</td>
<td>13684</td>
</tr>
<tr>
<td>Household</td>
<td>8515</td>
<td>6817</td>
<td>3936</td>
<td>5693</td>
<td>4868</td>
</tr>
<tr>
<td>White alone (%)</td>
<td>73</td>
<td>40</td>
<td>73</td>
<td>34</td>
<td>34</td>
</tr>
<tr>
<td>Black &amp; African American (%)</td>
<td>6</td>
<td>25</td>
<td>7</td>
<td>30</td>
<td>28</td>
</tr>
<tr>
<td>Asian alone (%)</td>
<td>14</td>
<td>8</td>
<td>13</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>Two races or more (%)</td>
<td>3</td>
<td>6</td>
<td>3</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Non-Hispanic (%)</td>
<td>90</td>
<td>63</td>
<td>91</td>
<td>61</td>
<td>63</td>
</tr>
<tr>
<td>Hispanic (%)</td>
<td>10</td>
<td>37</td>
<td>9</td>
<td>39</td>
<td>37</td>
</tr>
<tr>
<td>Median income</td>
<td>68359</td>
<td>34452</td>
<td>70314</td>
<td>32369</td>
<td>32872</td>
</tr>
</tbody>
</table>
## Built environment and temporal distribution:
Location Statistic associated with each pattern

<table>
<thead>
<tr>
<th>Patterns</th>
<th>High Afternoon peak</th>
<th>High Morning peak</th>
<th>No Morning peak</th>
<th>Low Ridership</th>
<th>High Transfer</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Transportation</strong></td>
<td>mean</td>
<td>mean</td>
<td>mean</td>
<td>mean</td>
<td>mean</td>
</tr>
<tr>
<td>Number of bus stops</td>
<td>43</td>
<td>27</td>
<td>49</td>
<td>21</td>
<td>26</td>
</tr>
<tr>
<td>Number of express bus stops</td>
<td>16</td>
<td>1</td>
<td>23</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Distance to downtown</td>
<td>7</td>
<td>12</td>
<td>7</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>Distance to midtown</td>
<td>7</td>
<td>11</td>
<td>7</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Distance to valley</td>
<td>7</td>
<td>12</td>
<td>7</td>
<td>13</td>
<td>14</td>
</tr>
</tbody>
</table>

### Land use and employment

<table>
<thead>
<tr>
<th></th>
<th>Total employment</th>
<th>Percentage of commercial floor area (%)</th>
<th>Percentage of residential floor area (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>71020</td>
<td>74</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>6026</td>
<td>33</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>86172</td>
<td>87</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>2553</td>
<td>17</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>3329</td>
<td>27</td>
<td>73</td>
</tr>
</tbody>
</table>
Built environment and temporal distribution

Patterns and associated features

- High Afternoon Peak Pattern – Land use mixed zones, high commercial land use percentage
- High Morning Peak Pattern – Land use mixed zones, medium commercial land use percentage
- No Morning Peak Pattern – Highly commercial zones
- Low Ridership Pattern – Highly residential zones mostly outside Manhattan
- High Transfer Pattern – Highly residential zones and subway service boundary
Influence of independent variables on Patterns

- **Population + Employment:**
  - High PM Peak -> No AM Peak -> High AM Peak -> Low Riders and High Transfer

- **Percentage of white population:**
  - No AM Peak -> High PM Peak -> High AM Peak -> Low Riders and High Transfer

- **Percentage of Asian population:**
  - No AM Peak -> Low Riders and High Transfer -> High PM Peak -> High AM Peak

- **Commercial floor area:**
  - No AM Peak -> Low Riders and High Transfer -> High PM Peak -> High AM Peak

- **Distance to midtown:**
  - Low Riders and High Transfer -> High AM Peak -> High PM Peak -> No AM Peak

- **Percentage of residential floor area:**
  - Low Riders and High Transfer -> High AM Peak -> High PM Peak -> No AM Peak
Conclusions

• Few has looked at time of day aspects, ours is one of the first
• Developed methodologies to characterize time of day pattern
• Demonstrated that:
  – Ridership time of day patterns vary by station
  – Distinctive time of day patterns can be identified
  – Local land use and station position in the transit network are both important factors in determining pattern membership
Next Steps

- Identify how the time of day attributes are related with each other
- Understand differences between weekdays and weekends
- Understand the change in patterns over time
- Linking station level analysis to metro-card level analysis
Acknowledgement

NYMTC
New York City Transit
UTRC