Near-Real Time Data for Operational Monitoring and Control

Alla Reddy
Senior Director
System Data & Research
Operations Planning

New York City Transit

APTA ITS TRANSIT BEST PRACTICES WORKSHOP
Atlanta, Georgia November 12-13, 2015
Today’s Presentation

• Overview: Big Data at New York City Transit
• Subways
  – Enhanced service monitoring with stringline charts
  – Post-hoc uses in Rail Control Center
  – Real-time monitoring of crowding at station platform
  – Dwell management with automated track occupancy data
• Buses
  – Surface line dispatcher use of automated data sources
  – Non-revenue service analyses
  – Improved planning with Intelligent Vehicle Networks
  – Customer driven schedule writing
• Summary
# Big Data at New York City Transit

<table>
<thead>
<tr>
<th></th>
<th>Bus</th>
<th>Subway</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>System Statistics</strong></td>
<td>• 2.6M weekday riders</td>
<td>• 5.5M weekday riders</td>
</tr>
<tr>
<td></td>
<td>• 300 routes</td>
<td>• 22 lines</td>
</tr>
<tr>
<td></td>
<td>• 6,000 buses</td>
<td>• 6,325 subway cars</td>
</tr>
<tr>
<td></td>
<td>• 6M daily Bus Time data records</td>
<td>• 469 stations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 100,000 daily ATS-A records</td>
</tr>
<tr>
<td><strong>Data Sources</strong></td>
<td>• MTA Bus Time: Real-time Automatic Vehicle Location (AVL) for all MTA buses</td>
<td>• Countdown Clocks (1-6, Grand Central Shuttle, L)</td>
</tr>
<tr>
<td></td>
<td>• System based on open source OneBusAway project</td>
<td>• Integrated Train Register Activity Console (I-TRAC) and Programmable Logic Controller (PLC) data (sections of 7 and lettered lines)</td>
</tr>
<tr>
<td><strong>Customer Applications</strong></td>
<td>• Live bus tracking: <a href="http://bustime.mta.info">http://bustime.mta.info</a></td>
<td>• Public Announcement/ Customer Information System (PA/CIS)</td>
</tr>
<tr>
<td><strong>Internal Performance Monitoring</strong></td>
<td>• Performance Indicators Planning and scheduling</td>
<td>• Performance Indicators Planning and scheduling</td>
</tr>
<tr>
<td></td>
<td>• Bus Ridership Model: Bus Time linked to MetroCard swipe data</td>
<td>• Subway Network Model (under development): Countdown clock data linked to MetroCard swipe data</td>
</tr>
</tbody>
</table>
Subways
Real Time Subway Movements

- A stringline chart represents each train trip by a “string” plotted with time on the horizontal axis and distance on the vertical axis.
- NYCT OP developed a tool for analysis of subway train movements.
- Interactive hover-overs support understanding of delays and trip attributes such as schedule and crew.

Delay information logged into train register.
Gap Table Example

- Gap table computes all gaps in the system in real time and sorts them by largest gap after.

<table>
<thead>
<tr>
<th>STATION</th>
<th>Train ID</th>
<th>Gap After</th>
<th>Gap Before</th>
<th>Territory</th>
</tr>
</thead>
<tbody>
<tr>
<td>174 ST</td>
<td>02 1615</td>
<td>15</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>WEST SQ</td>
<td>06 1623</td>
<td>14</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>MORRPARK</td>
<td>05 1618</td>
<td>13</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>STGA AVE</td>
<td>03 1622</td>
<td>10</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>WLCK AVE</td>
<td>06 1718+</td>
<td>9</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>176 ST</td>
<td>04 1634</td>
<td>8</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>FKLN AVE</td>
<td>03 1714</td>
<td>8</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>110 ST</td>
<td>02 1644</td>
<td>8</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>181 ST</td>
<td>01 1642</td>
<td>7</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>UTCA AVE</td>
<td>03 1720</td>
<td>7</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>138 3RD</td>
<td>06 1706</td>
<td>7</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>225 ST</td>
<td>01 1637</td>
<td>6</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>161 ST</td>
<td>04 1711+</td>
<td>6</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>176 ST</td>
<td>04 1719+</td>
<td>6</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>72 ST</td>
<td>01 1703</td>
<td>6</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>PEL PKYW</td>
<td>02 1718</td>
<td>6</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>CLRK ST</td>
<td>02 1619</td>
<td>6</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>NEVN ST</td>
<td>04 1625</td>
<td>6</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>145 ST</td>
<td>03 1729</td>
<td>6</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>145 ST</td>
<td>01 1715+</td>
<td>6</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

- Train currently near 174th Street
- 4-min gap before
- 15-min gap after
Applications in Rail Control Center

Reviewing a number of holds made to even out service on the 6 line

Evaluating wait assessment impacts from service interventions
Identifying Gaps in Service

- Bunched trains and gaps in service can be easily seen on stringline charts
- Possible causes for gaps include dropped intervals and train bunching
- Gaps identified in real time present an opportunity to take action
- Holds and skips are the simplest strategies for evening out spacing
- In this case, the train with a large gap behind it was later held at 177th Street
  - Would have been better to intervene much earlier, in Manhattan
Holding Trains

- Holds can be seen when a train slows down between stations
- More horizontal slope indicates slower speed
- Only arrival times are shown (not dwells), so the hold took place at the station before the slowdown
- Based on the stringline charts, there is no way to be certain whether a train was held intentionally, or slowed down on its own
- If a train that looks held was bunched, and the hold helped to even out spacing, it was likely intentional
Stringline Example
### Real Time Subway Platform Crowding

Three primary data sources used to estimate platform crowding levels

<table>
<thead>
<tr>
<th>Information Type</th>
<th>Data Source</th>
<th>Data Type</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passenger entrances</td>
<td>OMB AFC</td>
<td>Static (static)</td>
<td>Historic data used to forecast station entrances</td>
</tr>
<tr>
<td></td>
<td>Entry Data</td>
<td></td>
<td>6-minute AFC data uniformly distributed in 30-second intervals</td>
</tr>
<tr>
<td>Passenger direction &amp; train</td>
<td>Subway Ridership</td>
<td>Static</td>
<td>Incoming passengers assigned direction and train preferences based on 30-min period</td>
</tr>
<tr>
<td>preferences</td>
<td>Model AFC OD Table</td>
<td></td>
<td>Assignment to a preference group does not change due to gaps in service</td>
</tr>
<tr>
<td>Predicted Train Arrival &amp; Departure Time*</td>
<td>GTFS-RT Data Feed</td>
<td>Dynamic (real-time)</td>
<td>Train arrivals rounded to 30 second interval</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Passengers leave platform at arrival of needed train</td>
</tr>
</tbody>
</table>

Every 30 seconds throughout the day:
1. Add new entries to platform, grouped by train preference
2. Remove from platform when their train comes

* Development & validation done with historic train departures. Actual predictions using GTFS-RT data feed. Currently only able to tap arrival times from GTFS-RT, but looking into estimating departures.
A real time train arrival feed is combined with historic ridership patterns to estimate current crowding.
Real Time Subway Platform Crowding

Number of Passengers on Platform at Wall Street Subway Station

Train Arrivals - Northbound

Crowding Peaks within The Next 15 Mins

Train Arrivals - Southbound

Passengers by Intended Line and Direction

Last Updated: 4:57:30 PM

Under Development
## Real Time Subway Platform Crowding

Crowding data has a number of applications, each with alternative approaches

<table>
<thead>
<tr>
<th>Use crowding level to decide...</th>
<th>When prediction needed</th>
<th>Alternative approaches</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Service Control Actions by RCC</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Hold/skip/turn/reroute</td>
<td>1-15+ min prior to arrival</td>
<td>Watch stringlines for gaps</td>
<td>Some actions more likely than others, depending on station</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Service control guidelines</td>
<td></td>
</tr>
<tr>
<td><strong>Customer-Facing Actions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Limit station entries</td>
<td>Any time issue arises</td>
<td>Allow customers, staircases to naturally limit entries</td>
<td>We might not have the capacity / resources to address this in a meaningful timeframe</td>
</tr>
<tr>
<td><strong>Performance Metrics on Volume/Crowding</strong></td>
<td>Incident reports or requests of mgmt.</td>
<td>Do not currently exist (&lt;7 days out)</td>
<td>Can be provided on an as needed basis or integrated with STARS reports</td>
</tr>
<tr>
<td><strong>Answering whether service mgmt. worked</strong></td>
<td>Same day or next day</td>
<td>Stringlines</td>
<td>Ridership model output may be broader but not ready in time (at least 6 day lag)</td>
</tr>
<tr>
<td><strong>Where overcrowding is a systemic issue</strong></td>
<td>Long term</td>
<td>TransCAD models in rail planning</td>
<td>Ridership model output can likely serve this purpose more broadly at all station types</td>
</tr>
</tbody>
</table>

Under Development
Subway Dwell Management

- Dwell Time at key stations
- Understanding sources of long dwells
- Identifying stations where platform controllers may be most effective in reducing dwell times
- Monitor impact of platform controllers over time
Line Performance Summary

- Highlight key dimensions of quality service
- Internal management tool
- Wait Assessment, throughput, dwell times
Subway Passenger Metrics

- Develop metrics that reflect how passengers perceive service
- Evaluate average waiting times on subway platforms
Buses
Service Monitoring with BusTime

- GPS data available every 30 seconds
- Next-day performance review possible
- Provide Surface Line dispatchers with comprehensive overviews of route being managed for more informed dispatching strategies

Pick-level Report: Running Times, Load Profiles, boarding & alighting, etc.

Weekly Report: OTP, WA, Bunching, etc.
Non-Revenue Service Monitoring

- Expand review of BusTime data to non-revenue service
- Explore scheduling improvements of pull-outs and deadhead movements
- Can determine causes for route performance issues due to insufficient running time during non-revenue portion of schedule

Excessive Running Time Scheduled

Insufficient Running Time Scheduled
Tools for Improved Planning: IVN

- Improved granularity of IVN (Intelligent Vehicle Network) data
- Speed profiles and dwell assessment
- Planning applications (bus lanes, off-board fare payment, etc)
- Operational decisions (where not to express / short-turn, etc)
Improved Scheduling

• Schedules formerly written from daily data now based on full-month averages
• Complete system data can identify need for schedule review instead of adherence to biennial schedule
• Identify where valuable, and plan different schedules for different day types

Characteristic Ridership differences between Fridays and Midweek

Net Change in Express Trips: Friday v. Midweek, Staten Island Borough
Bus Passenger Metrics

- Develop metrics that reflect how passengers perceive service
- Evaluate waiting and in-vehicle dimensions of travel
- Evaluate impacts of introduction of bus lane on E 125th St
- Amount of excess journey time for each passenger fell considerably
Summary: Better Data – Better Service

• New York City Transit is generating and using large amounts of automatic data, for both buses and subways
  – Subway: Countdown clocks; PLCs; Modeling Ridership; Modeling Crowding; Dwell
  – Bus: BusTime; IVN; Modeling Ridership
• Data is being used in a variety of contexts
  – Live operations management and service monitoring
  – Daily reviews
  – Schedule-making
  – Planning studies
• Key lesson learned: **Improve data quality by integrating multiple sources**
  – Not all data sources available in real-time; what is available is most useful for quick decision-making and short-term analysis (80-85% correct in representing actual service)
  – Special processes and algorithms are used to connect several data streams for a more complete picture; useful for long-term analysis, planning studies (95-98% correct)
  – Long-term goal: have all data available in real-time for a comprehensive picture of service