

Outline Final Presentation on Oyster-based Performance Metrics

Michael Frumin, TfL London Rail and MIT, mfrumin@mit.edu

What, Outline?

A method for measuring performance on the London Overground (or other rail network) in a passenger-focused way using data from the Oyster smart card system. The hope is that such a measure can lead to improved management of operations, contracted or otherwise.

Presentation Structure:

- What I'm doing, why, context -- i.e. other similar systems of measurement
- Get into details, look at some data, including one digression that came out of the project
- Anticipate problems for real world use, suggest some possible solutions
- Discuss how the work fits into broader TfL contexts

Why, what advantages?

- As a direct measure of passengers origins, destinations, and end-to-end journey times, should reflect the the passenger experience, in terms of:
 - What is measured
 - How it is weighted/aggregated

Compared to:

- Network Rail's Public Performance Measure (PPM) **(SLIDE)**
 - Measures trains -- % late and delay minutes
 - Large sample sizes, mostly accurate, well assigned to responsible parties
 - No measures of demand so possibly divergent from actual passenger experience
 - Used in LOROL contract

That is, rather than measuring *punctuality and performance* of scheduled vehicles, we would hope to measure *delays and reliability* for passengers. That is, DISTRIBUTIONS.

- London Underground -- Journey Time Metric
- London Buses -- Excess Waiting Time

Data and Methodology:

- Oyster sample sizes and O/D pair classification **(SLIDE)**
- Calculation of different journey times with timetable **(SLIDE)**
- Digression into Scheduled Waiting Time, but an important one
 - Scheduled waiting time distributions **(SLIDE)**

- Shuttles to Camden and Specials to Clapham Junction were crowding relief services purchased by TfL on Silverlink, slotted in between regular services out of simplicity of arrangement.
 - Now a problem, quite unbalanced loads -- talk about even headways **(SLIDE)**
 - Note: this graph looks roughly the same when plotted just for through trips (i.e. people crossing Camden Road)
 - Oliver Bratton of LOROL -- try to get even headways in May 2009 timetable or SLC2K (Dec 2010)
-
- Validation via graphical analysis of specific days **(SLIDE)**:
 - 2008-04-03
 - Door problems on 8:06 train (18 min late at Richmond)
 - 9:30 shuttle to Camden was cancelled
 - Train from Richmond that was to depart Stratford at 9:52 short turned at Hackney Wick

Aggregate Results:

- Show some aggregate EJT plots:
 - By line, daily and weekly **(SLIDE)**
 - By timeband on NLL, weekly and period **(SLIDE)**
- Show EJT distributions by line and timeband **(SLIDE)**
 - **Distributions are important because they represent buffer time for passengers.**

Relationship Between EJT and PPM/TRUST Variables:

- Show time series and scatterplot **(SLIDE)**
 - Some correlation, but far from perfect
- Linear models **(SLIDE)**
 - No major surprises, but not very satisfying
 - Not good enough to generate EJT in the future via proxy variables

Likely problems and pitfalls:

It's all well and good to look at some nice pretty pictures, but what would it take to be able to use the Oyster data for real operations and contract management? Likely, robust (or at least believable) *assignment* of passenger delays to different parties.

This brings into stark relief the contrast between our goal in providing services and our goal in measuring performance. That is, we are trying to create as integrated a network as possible for passengers to use, but are measuring performance independently for each part of the network.

On the other hand, we can finally measure performance across networks!

Two primary steps in this assignment:

1) For each trip, which services did it use? Inherent ambiguity in looking only at entry and exit locations and times

- Transfers (**SLIDE**)
 - Transfers between LO services -- about 20% of pure LO trips involve an LO-LO transfer.
 - Even worse (or better?), behind-the-barrier transfers between LO and other (i.e. LU) services -- of trips originating at stations with LO services, about half of them terminate at stations w/out LO services.
 - Some of the biggest LO stations are LU interchanges (e.g. Stratford, H&I, Willesden, Barking, etc)
 - On the one hand, this makes it difficult for us to measure the performance of individual networks (i.e. LO vs LU). On the other hand, it is an opportunity to measure "trans-network" performance which has not previously been done.
- Path Choice (**SLIDE**)
 - This also makes it difficult to measure distinct parts of the network, but O/D pairs with multiple paths will be more robust with respect to problems in the network, which would be reflected by end-to-end measurements.
- Interavailable services (**SLIDE**) -- of trips between two LO stations, 16% are on interavailable OD pairs. Similar to above.
- Journey components -- access, wait, egress, crowding (**SLIDE**)

2) For delays to specific trips on each service, what or whom is responsible for those delays? Assignment of metrics to responsible parties is not immediate.

- Even with perfect information about the operation of the service, plenty of work is required to connect problems with the service to delays of specific passengers
- PPM and JTM systems assign delays to specific incidents, and responsible parties

Of course, typical problems with Oyster:

- Scaling up to calculate total, as opposed to average, delays
- Oyster biases relative to general riding population -- by line, station, time of day, O/D pair, etc.
- Last but not least, weird travel behavior, especially in the face of disruptions

So What? (Short Term) (SLIDE)

- Operations and Contract Management:
 - Quoth (approximately) Nick Arthurton, LO Concession Manager: "What good is it running an on-time train if it's full of air?"
 - Questions for audience:
 - How can PPM and Oyster-based metrics work along side each other?

- **What, if anything, are we sacrificing by using only PPM?**
- Apples to apples comparison with other TfL modes (i.e. JTM for LU, Excess Waiting Time for Bus)
- Oyster Extension to National Rail (OXNR):
 - TfL monitor National Rail services in London?
 - Strategic information "partnership" -- trade Oyster for LENNON

So What? (Long Term) (SLIDE)

- Simple (ish) and straightforward way to measure delays and reliability of rail journeys across London's entire rail network -- LU + DLR + LO + NR + Crossrail

Future Research:

- To deal with ambiguous path choices, can take a probabilistic approach, similar to OXNR, which should work fine for estimating Expectation[Journey Time] but I'm not quite sure about estimating distributions.
- Advanced inference for identifying different journey components -- Access/Egress/Interchange, Waiting Time, In-Vehicle Time
- BART-Like combination of passenger and line signaling data to "load" passengers onto specific services. Direct link of operational incidents to passenger delays.

Final Thoughts/Questions:

- Promising results and possibilities, but is it worth the cost and complexity?
- Possibly not just for London Overground, given the comprehensiveness and history of PPM, but possibly for trans-London, inter-network measures provided by TfL.
 - It depends on timetables and network structure for all services, but those are simple compared to a host of different signalling systems, delay measurements, and other ad-hoc performance data.