The Role of Information Technology in Improving Transit Systems

by
Nigel Wilson
MIT

OUTLINE

- **Key Automated Data Collection Systems (ADCS)**
- **Key Transit Agency/Operator Functions**
- **Impact of ADCS on Functions**
- **Traditional Relationships Between Functions**
- **State of Research/Knowledge**
- **Examples of Recent Research**
- **Emerging Research Possibilities**
- **Remaining Challenges**

Transit Agencies Are at a Critical Transition in Data Collection Technology:

Manual		**Automatic**
• low capital cost	⟹	• high capital cost
• high marginal cost		• low marginal cost
• small sample sizes		• large sample sizes
• aggregate		• more detailed, disaggregate
• unreliable		• errors and biases can be estimated and corrected
• limited spatially and temporally		• ubiquitous
• not immediately available		• available in real-time or quasi real-time

Key Automated Data Collection Systems

- **Automatic Vehicle Location Systems (AVL)**
  - **bus location based on GPS**
  - **train tracking based on track circuit occupancy**
  - **real-time availability of data**
- **Automatic Passenger Counting Systems (APC)**
  - **bus systems based on sensors in doors with channelized passenger movements**
  - **passenger boarding (alighting) counts for stops/stations with fare barriers**
  - **train weighing systems to estimate number of passengers on board**
  - **traditionally not available in real-time**
- **Automatic Fare Collection Systems (AFC)**
  - **increasingly based on contactless smart cards with unique ID**
  - **provides entry (exit) information (spatially and temporally) at the individual level**
  - **traditionally not available in real-time**

ADCS - Potential and Reality

Potential

- **Integrated ADCS database**
- **Models and software to support many agency decisions using ADCS database**
- **Providing insight into normal operations, special events, unusual weather, etc.**
- **Provide large, long-time series disaggregate panel data for better understanding of travel behavior**

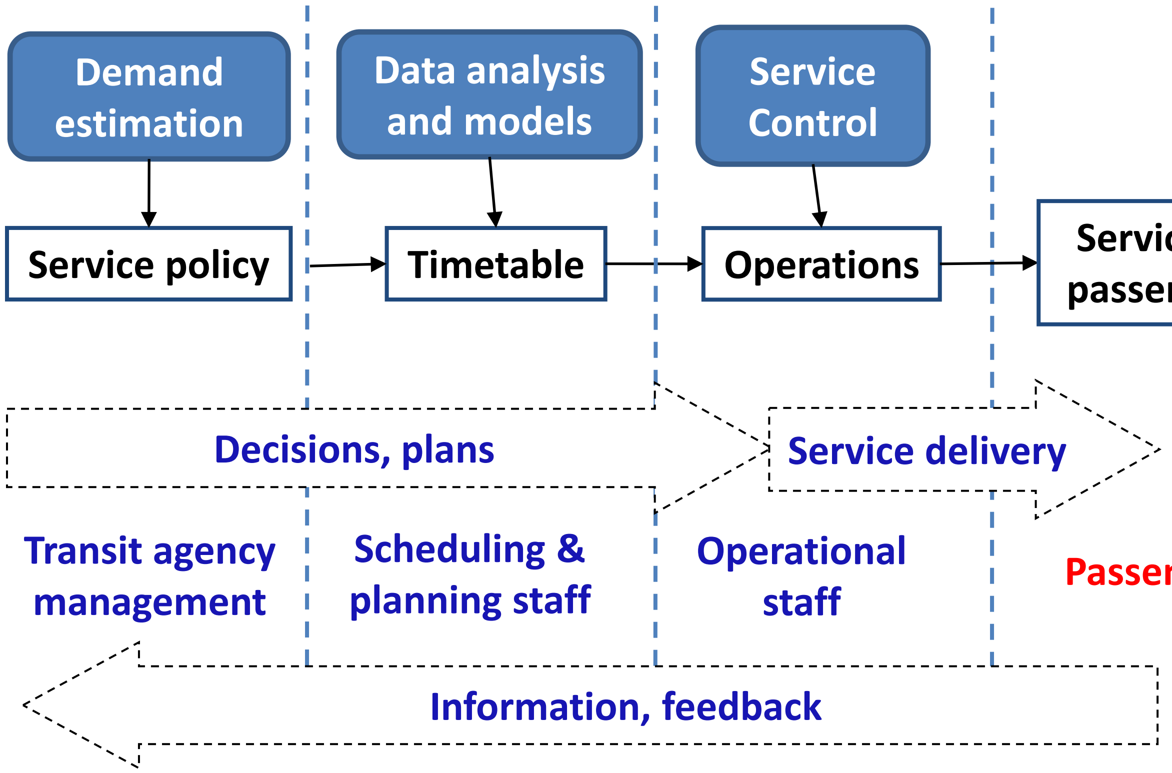
Reality

- **Most ADCS systems are implemented independently**
- **Data collection is ancillary to primary ADC function**
  - **AVL - emergency notification, stop announcements**
  - **AFC - fare collection and revenue protection**
- **Many problems to overcome:**
  - **not easy to integrate data**
  - **requires substantial resources**

Key Transit Agency/Operator Functions

- **Service and Operations Planning (SOP)**
  - **Network and route design**
  - **Frequency setting and timetable development**
  - **Vehicle and crew scheduling**
  - **Off-line, non real-time function**
- **Service and Operations Control and Management (SOCM)**
  - **Dealing with deviations from SOP, both minor and major**
  - **Dealing with unexpected changes in demand**
  - **Real-time function**

Transit Service Delivery Process*



* *Source: "Diagnosis and Assessment of Operations Control Interventions: Framework and Applications to a High Frequency Metro Line." MST Thesis, André Carrel; MIT, 2009.*

Key Transit Agency/Operator Functions (cont'd)

- **Customer Information (CI)**
  - **Information on routes, trip times, vehicle arrival times, etc.**
  - **Both static (based on SOP) and dynamic (based on SOP and SOCM)**
  - **Both pre-trip and en-route**
- **Performance Measurement and Monitoring (PMM)**
  - **Measures of operator performance against SOP**
  - **Measures of service from customer viewpoint**
  - **Traditionally an off-line function**

Impact of ADCS on Functions

IMPACT ON SOP

- **AVL:** detailed characterization of route segment running times
- **APC:** detailed characterization of stop activity (boardings, alightings, and dwell time at each stop)
- **AFC:** detailed characterization of fare transactions for individuals over time, supports better travel behavior modeling

IMPACT ON SOCM

- **AVL:** identifies current position of all vehicles, deviations from SOP

IMPACT ON CI

- **AVL:** supports dynamic CI
- **AFC:** permits characterization of normal trip-making by each individual, supports active dynamic CI function

IMPACT ON PMM

- **AVL:** supports on-time performance assessment
- **AFC:** supports passenger-oriented measures of travel time and reliability

Traditional Relationships Between Functions

- SOP serves as the basis for both SOCM and CI
- Reasonable as long as SOP is sound and deviations from it are not very large
- Input data to the SOP has improved as a result of ADCS
- Fundamentally a static model in an increasingly dynamic world

Service Planning Hierarchy

Infrequent Decisions

Service Considerations Dominate

Judgement & Manual Analysis Dominate

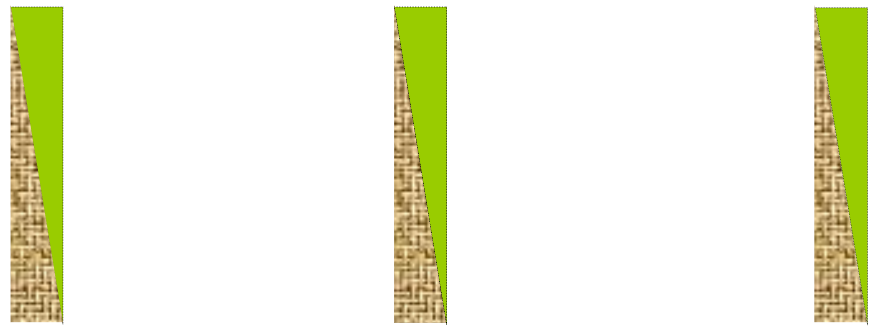
Network Design

Frequency Setting

Timetable Development

Vehicle Scheduling

Crew Scheduling



Frequent Decisions

Cost Considerations Dominate

Computer-Based Analysis Dominates

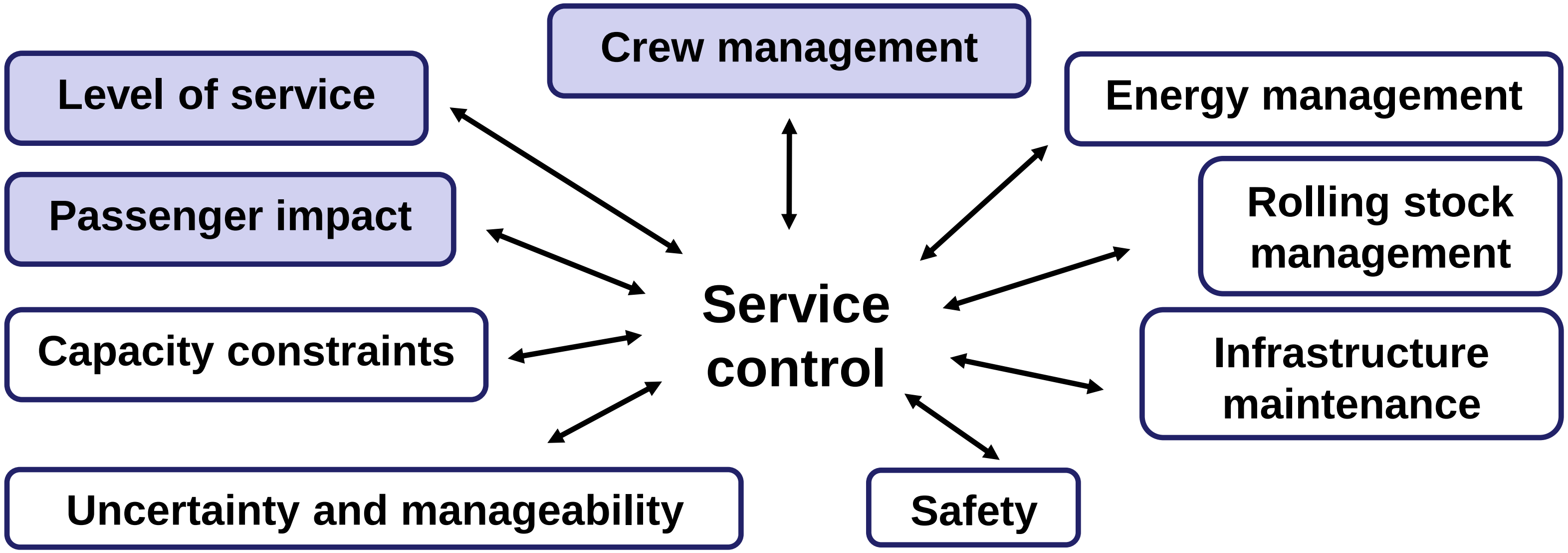
State of Research/Knowledge in SOP

- Advanced in vehicle and crew scheduling (operations planning)
- Limited in past by weak data, less of a problem now
- Limited in service planning: rules of thumb and experience still dominate
- Much research has been simplistic in terms of formulation of objectives and constraints
- Inadequate recognition of uncertainty in model formulation
- Substantial opportunities remain for better models

State of Research/Knowledge in SOCM

- Advances in train control systems help minimize impacts of small incidents
- Major disruptions still handled in individual manner based on judgement and experience of the controller
- Little effective decision support for controllers
- Models suffer from deterministic formulations of highly stochastic systems
- Simplistic view of objectives and constraints in model formulation
- Substantial opportunities remain for better models

Rail Operations Controllers Decision Factors



- These factors can trigger service control interventions or place constraints on interventions performed for other reasons
- Conflicts between objectives are frequent
- How can we best coordinate and integrate these objectives and constraints?

Source: "Diagnosis and Assessment of Operations Control Interventions: Framework and Applications to a High Frequency Metro Line." MST Thesis, André Carrel; MIT, 2009.

State of Research/Knowledge in CI

- **Next vehicle arrival times at stops/stations well developed and increasingly widely deployed**
- **Pre-trip journey planner systems widely deployed but with limited functionality in terms of recognizing individual preferences**
- **Strongly reliant on veracity of SOP**
- **Ineffective in dealing with major disruptions**

Evolution of Customer Information

- Operator view --> customer view
  - route-based --> OD-based
- Static --> dynamic
  - based on SOP --> based on SOP modified by current system state and control actions
- Pre-trip and at stop/station → en route
- Generic customer → specific customer
- Request-based systems → Anticipatory systems
- Agency/operator developed systems → "App" developers using real time data feeds from agency

State of Research/Knowledge in PMM

- **Generally takes the operator rather than customer perspective**
  - **route- or stop-based measures rather than OD pairs measures**
  - **lack of effective measures of reliability**
  - **lack of recognition of non-linear response in terms of customer satisfaction**
- **Based on achieving SOP as ultimate goal**

Examples of Recent MIT Research Based on ADCS

Service and Operations Planning

- Trip chaining to estimate OD matrix
- Travel behavior analysis

Service and Operations Control and Management

- Operations control on metro line (LUL Central Line)
- Estimation of train level passenger loads (LUL Lines)

Planning and Performance Monitoring

- Reliability metrics (LUL and LO)

Public Transport OD Matrix Estimation

Objective:

- Estimate passenger OD matrix at:
  - single route level
  - network level

Network attributes:

- multi-modal rail and bus systems
- entry-control-only or entry+exit control operations

Sources:

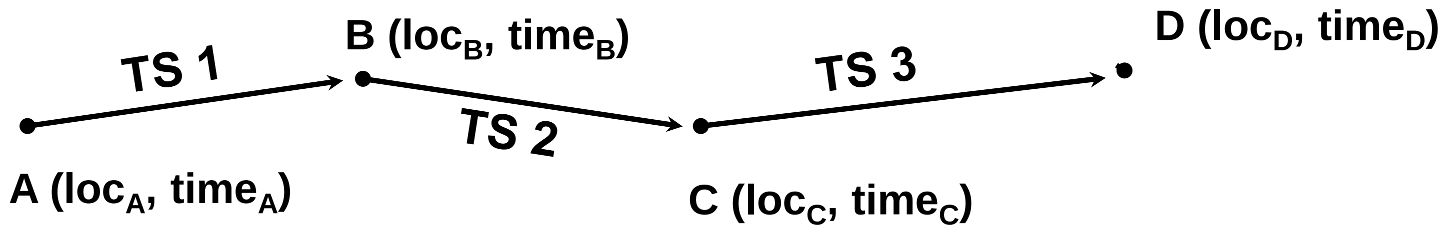
"Bus Passenger Origin-Destination Matrix Estimation Using Automated Data Collection Systems." Alex Cui, MST Thesis, MIT, June 2006

"Bus Passenger Origin-Destination Estimation and Travel Behavior Using Automated Data Collection Systems in London, UK." Wei Wang, MST Thesis, MIT, June 2010

Trip Chaining: Basic Idea

Each AFC record includes:

- AFC card ID
- transaction type
- transaction time
- transaction location: rail station or bus route (time-matching with AVL data)



The destination of many trip segments (TS) is also the origin of the following trip segment.

Trip-Chaining Method for OD Inference

Key Assumptions for Destination Inference to be correct:

- No intermediate private transportation mode trip segment
- Passengers will not walk a long distance
- Last trip of a day ends at the origin of the first trip of the day

Trip-Chaining Method for OD Inference

Steps required:

- **Infer start and end of each trip segment for individual AFC cards**
- **Link trip segments into complete (one-way) journeys for individual AFC cards**
- **Integrate individual journeys to form seed OD matrix by time period**
- **Expand to full OD matrix using available control totals**
  - **station entries and/or exits for rail**
  - **passenger entries and/or exits by stop, trip, or period for bus**

Summary Information on London Application

- **Oyster fare transactions/day:**
  - **Rail (Underground, Overground, National Rail): 6 million (entry & exit)**
  - **Bus: 6 million (entry only)**
- **For bus:**
  - **Origin inference rate: 95%**
  - **Destination inference rate: 74%**
- **Computation time for full London OD Seed Matrix:**
  - **12 mins on 2.8 GHz Intel 7 machine with 8 GB of RAM**

Travel Behavior Analysis

Objective:

- Estimate customer preferences for bus versus rail

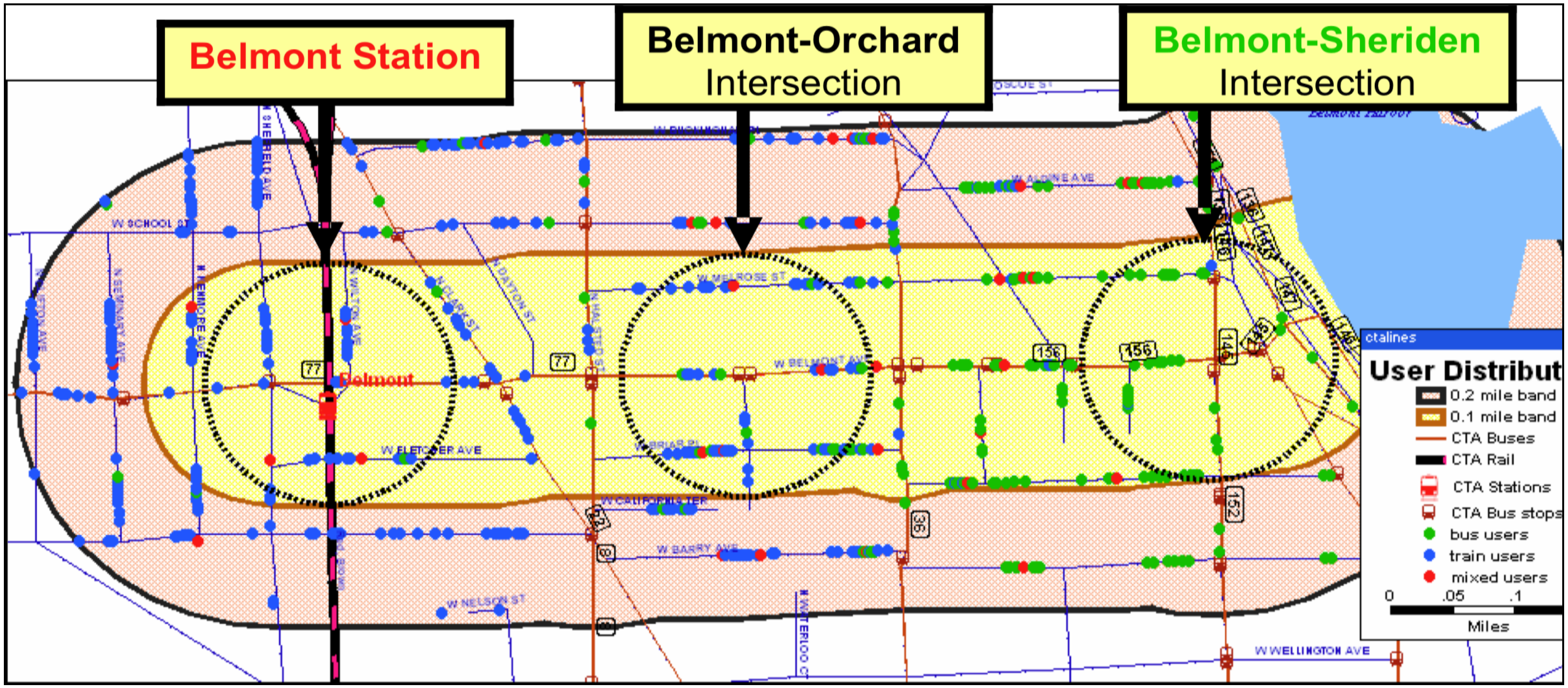
Given:

- AFC address registration
- AFC transactions

Applications:

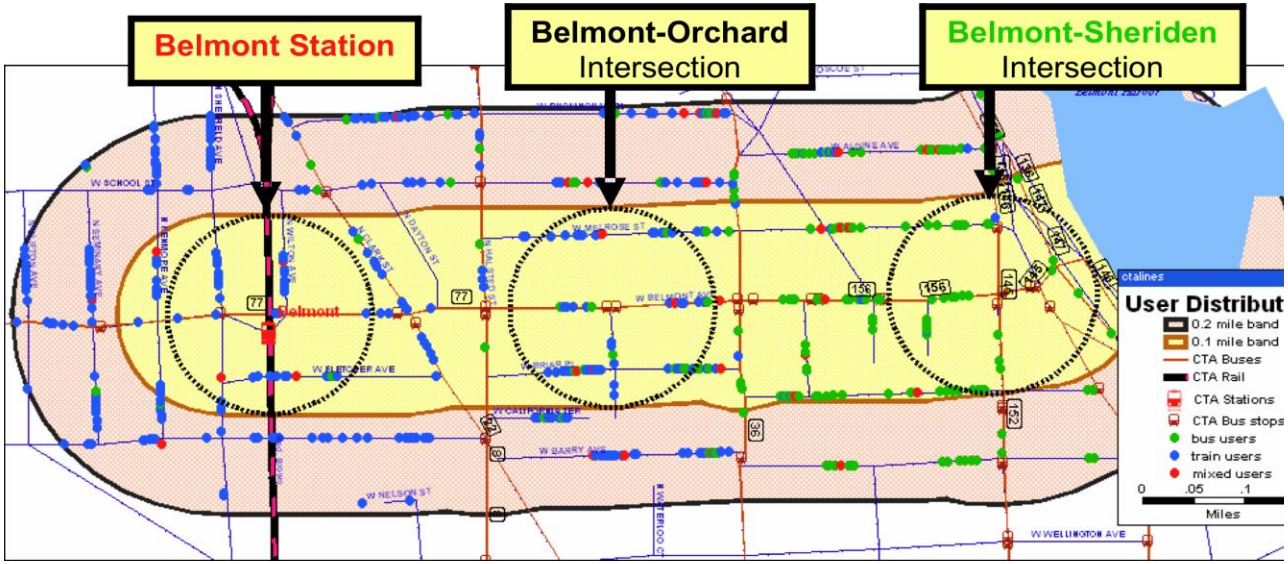
- CTA

Path Choice Analysis: Sample Users



- **Multiple rail and bus routes serving the loop (CBD)**
- **Stiff competition between express bus and rail service**

Path Choice Analysis: Access Distance



Intersection	bus	mixed	rail	Total
Belmont Station	2	4	73	79
Belmont/Orchard	20	9	83	112
Belmont/Sheriden	170	10	21	201

- **Access Distance important: Belmont, Belmont/Sheriden**
- **Why are users in Belmont/Orchard not a better mix of Bus and Rail**

Path Choice Analysis: Mode Preference

User type (first trip of day)	User type (all daily trips)			# customers
	All Bus	All rail	Mixed	
Bus	47%	0%	53%	658
Rail	0%	71%	29%	704
Mixed	0%	0%	100%	99
All users	21%	34%	44%	1461

- Sample comprise all users in 0.2 mile buffer
- 53% bus users also use rail (< 18% trips)
- 29% rail users also use bus (< 7% trips)

Reliability Metrics

Objective:

- Estimate measures of service reliability at:
  - OD level
  - Line level

Given:

- AFC transactions
- AVL data
- Timetable data

Applications:

- London Underground
- London Overground
- London Buses

Motivation

- **Unreliability is seen as a widespread problem**
  - **Passenger impacts**
    - **Longer wait times**
    - **Need for trip time reliability buffer**
    - **More perceived crowding**
  - **Agency impacts:**
    - **Increased costs**
    - **Reduced ridership and revenue**
    - **Reduced operator morale**
    - **Public and political problem**
    - **Reduced effective capacity**

Problem Complexity

- Reliability is not the only service dimension of value, also have:
  - Speed/trip time
  - Productivity
- Reliability means different things:
  - To different customers
  - On different services
- A single measure of effectiveness focused on reliability may lead to poor decisions

BUT

- We do need to measure performance *wrt* reliability

Different Service Types

A. Low Frequency Service (typically defined as headways greater than 10-15 minutes)
   - Most customers time their arrival at stops/stations based on expected service departure times (e.g. schedule)
   - On-time performance is critical, for example:
     - 1 minute early to 5 minutes late
     - 0 minutes early to 3 minutes late
     - 0 minutes early to 1 minutes late
   - Little interaction between successive vehicles

Different Service Types

B. High Frequency Service

- Most customers do not time their arrival at stops with service departures
- Expected wait time = F(mean and variance of headways)
- On-time performance not so critical
- Extensive interaction between successive vehicles:
  - Vehicle bunching
  - Long gaps

BUT

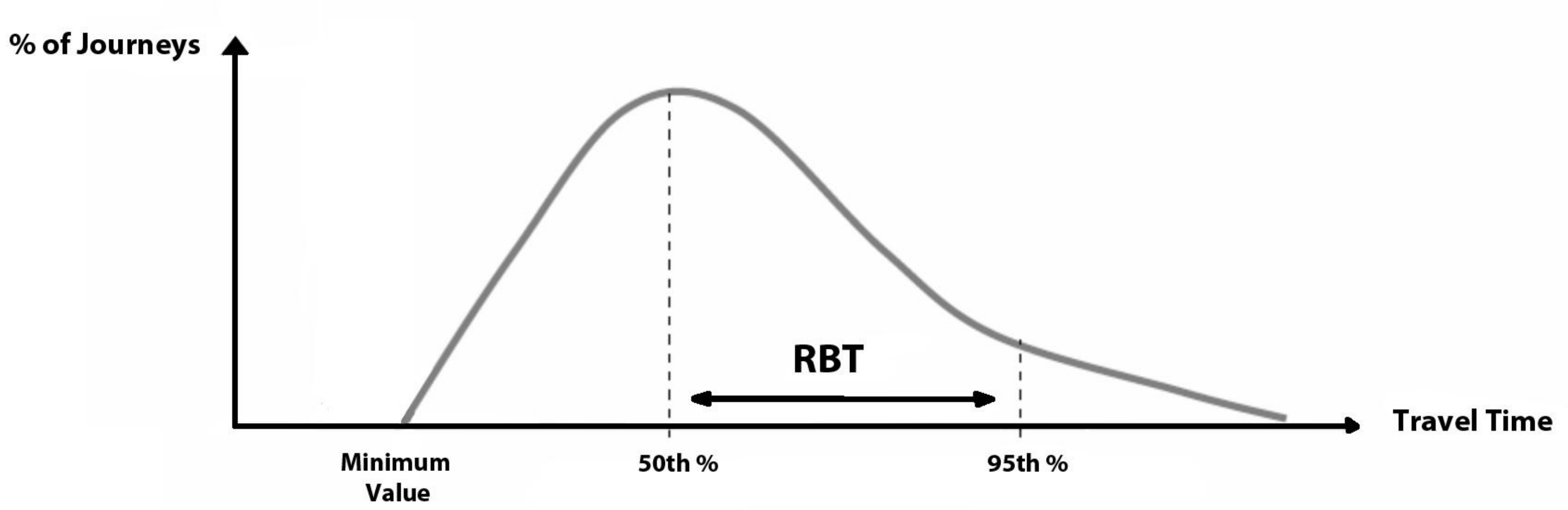
- Many high frequency routes have branches and short route variants
- So many customers may still behave like those on low frequency routes
- Schedule control is much easier than headway control…..

Reliability Metrics - Rail

High Frequency Service

- use tap-in and tap-out times to measure actual station-station journey times

s

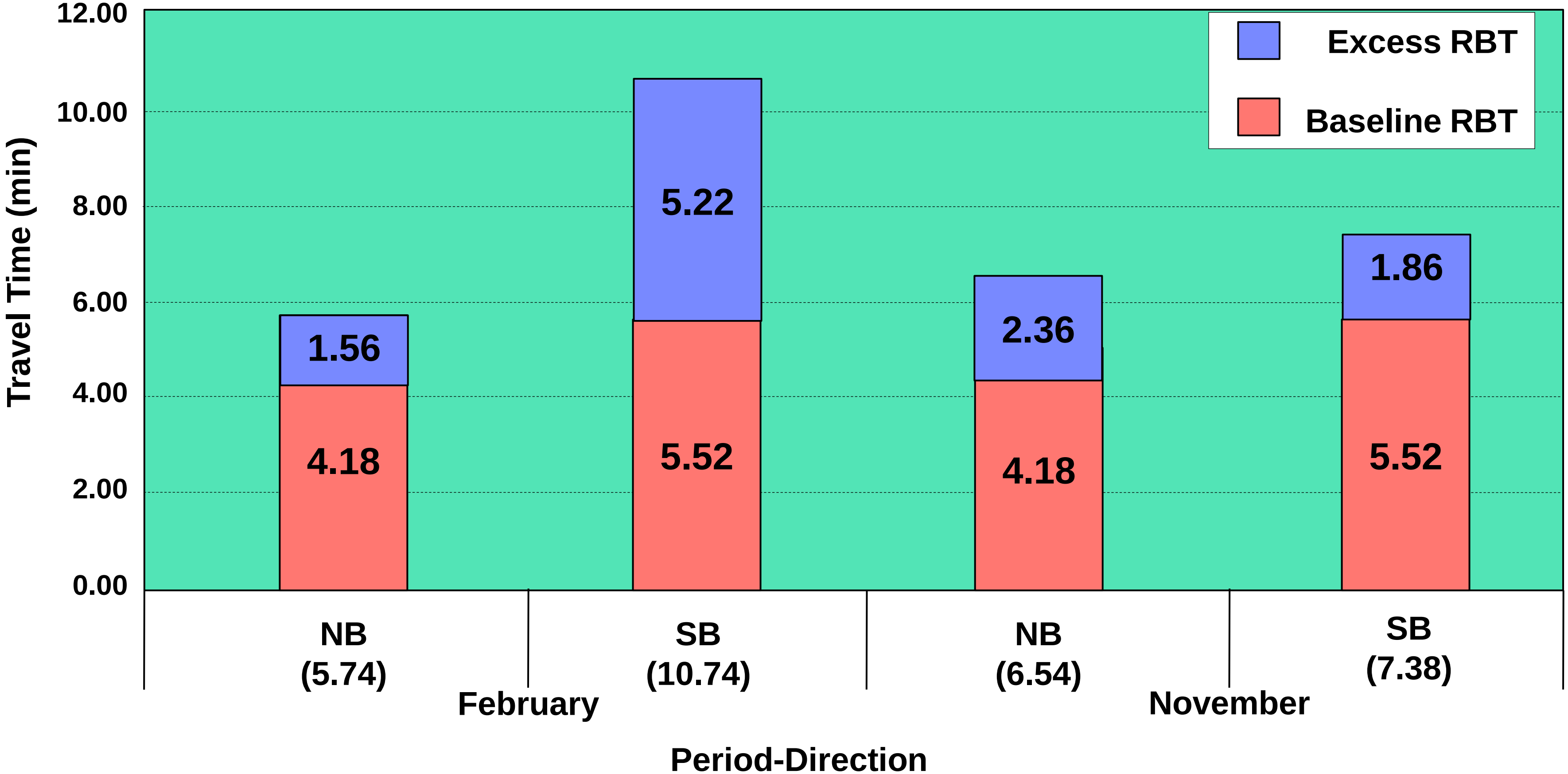


RBT = 95th percentile travel time – median travel time

The additional time a passenger must budget to arrive late no more than 5% of the time

Reliability Metrics - Rail

- Aggregate to line level by distinguishing between "normal" and "incident days"



Victoria Line, AM Peak, 2007

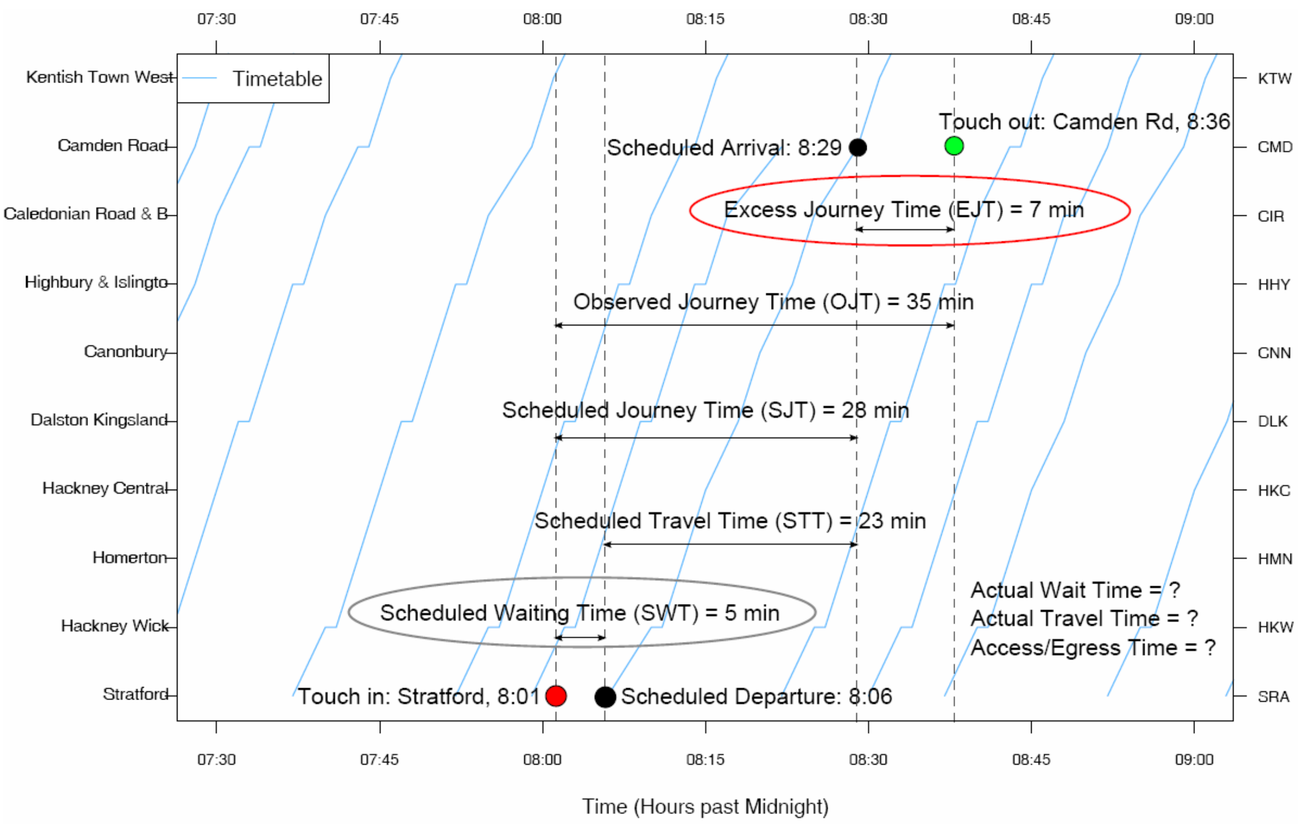
BUS RAPID TRANSIT ACROSS LATITUDES AND CULTURES

Source: David Uniman, MST thesis, MIT 2009. "Service Reliability Measurement Framework using Smart Card Data: Application to the London Underground."

Reliability Metrics - Rail

Low-Frequency Service

- compare actual journey times with scheduled times
- compare actual journey times



BUS RAPID TRANSIT
ACROSS LATITUDES AND CULTURES

Source: Michael Frumin, MST thesis, 2010 "Automatic Data for Applied Railway Management: Passenger Demand, Service Quality Measurement, and Tactical Planning on the London Overground Network."

Reliability Metrics - Bus

In contracted service delivery context, need to distinguish between:

A. Contractor performance: measure against contracted service expectations

B. Performance as seen by passenger

If service is unreliable, <u>the passenger</u> doesn't care whether the problem was caused by traffic or poor operator behavior, but <u>the authority</u> must be sure which caused the problem.

Reliability Metrics - Bus

Challenge to measure passenger journey time because:

- **(typically) no tap-off, just tap-on**
- **tap-on occurs <u>after</u> wait at stop, but wait is an important part of journey time**

Strategy to use:

- **trip-chaining to infer destination for all possible boardings**
- **AVL to estimate:**
  - **average passenger wait time (based on assumed passenger arrival process)**
  - **actual in-vehicle time**