A Framework for Identification of Systematic Service Deterioration in Urban Rail Systems

by

Kenji Chigusa

M.S., Information Science and Technology, the University of Tokyo (2011)
B.E., Electronic Engineering, the University of Tokyo (2009)

Submitted to the Department of Civil and Environmental Engineering in partial fulfillment of the requirements for the degree of

Master of Science in Transportation at the Massachusetts Institute of Technology

June 2018

© 2018 Massachusetts Institute of Technology. All rights reserved.
A Framework for Identification of Systematic Service Deterioration in Urban Rail Systems

by

Kenji Chigusa

Submitted to the Department of Civil and Environmental Engineering on May 18, 2018, in partial fulfillment of the requirements for the degree of Master of Science in Transportation

Abstract

With the increasing availability of transit information from well-developed apps, urban rail passengers are more and more aware of how long their journeys should take. This also means that passengers are becoming more sensitive to the gap between the service they expect of and what was actually provided. On the other hand, in urban rail systems operating high frequency services along with heavy ridership, trains could be frequently delayed due to less buffer time and passenger surges at platforms. In addition, passengers could experience further service deterioration resulting from overcrowding both in stations and trains. Passengers may be forced to miss several trains due to train capacity constraints. These passengers’ out-of-vehicle experience is becoming a serious concern which has not been fully captured by conventional performance measures.

This thesis develops a framework for identifying systematic service deterioration from the passengers’ perspective in urban rail systems operating near-capacity. This framework aims to support rail agencies’ problem understanding and management in the interest of service improvement. Specifically, the framework uses excess journey time as a proxy for the gap between service expectation and passenger experience. Using train-flow and passenger-flow data from automatic vehicle location and automatic fare collection systems, the framework estimates journey time for each origin and destination station pair, and time period in excess of the standard journey time used as a reference. Based on the median excess journey time, the framework identifies the time, location, frequency, degree, and (in some cases) causes of service deterioration. Guidelines for countermeasures and further analyses are also provided. The main use of the framework is for periodic problem identification, but it can also support before and after analysis regarding service changes.

The framework is tested and demonstrated by applying it to a line of Hong Kong Mass Transit Railway, which is one of the most heavily utilized rail systems in the world. The results of analysis are hotspots which identify points of significant passengers’ out-of-vehicle experience deterioration. The analysis further indicates unexpected passengers’ detouring behaviors to pursue less crowded trains. A before and after analysis focusing on the impact of a line extension shows that passengers traveling from a specific station are likely to experience greater service deterioration after the extension.
Thesis Supervisor: Nigel H.M. Wilson
Title: Professor Emeritus of Civil and Environmental Engineering

Thesis Supervisor: Haris N. Koutsopoulos
Title: Professor of Civil and Environmental Engineering, Northeastern University
Acknowledgments

The research presented in this thesis and the entire course of graduate study at MIT have been the greatest challenges in my life. The completion of these works would not have been possible without wonderful support from many people whom I would like to acknowledge here.

First and foremost, I would like to express my deep gratitude to my advisors, Professor Emeritus Nigel Wilson and Professor Haris Koutsopoulos. They have provided me an invaluable guidance and encouragement throughout these two years. Our weekly discussions have always been challenging, informative, and enjoyable for me. It was a great pleasure to work with them.

Hong Kong MTR has generously provided us valuable data and made this research possible. Felix Ng, Pandora Leung, and Henry Lo have kindly supported our research and arranged visits to Hong Kong.

I would like to thank all faculties and members of our research group. Professor Jinhua Zhao, as the research group leader and my academic adviser, has been exceptionally encouraging and supportive. Professor Rabi Mishalani has always provided me very insightful and practical advice to proceed with research and write the thesis. Saeid Saeid have given me helpful guidance and support. I am also very grateful for intelligent and delightful project members, Yiwen Zhu, Mike Ma, Abhishek Basu, Tianyou Liu, and Kerem Tuncel.

I would like to express my special gratitude to East Japan Railway Company and Professor Emeritus Joseph Sussman for giving me such a precious opportunity to expand my horizons at MIT. These two years have changed my life, literally, and I start to see things differently.

Finally, special thanks to my family, especially my wife, Yuriko. Your continuous support and encouragement were worth everything.
Contents

Chapter 1  Introduction ........................................................................................................... 13
  1.1  Motivation .................................................................................................................. 15
  1.2  Research Objective ................................................................................................. 16
  1.3  Approach .................................................................................................................. 17
  1.4  Thesis Structure ...................................................................................................... 20

Chapter 2  Literature Review ............................................................................................. 21
  2.1  Transportation Performance Measures .................................................................. 21
    2.1.1  Train-Oriented Measures .................................................................................. 21
    2.1.2  Passenger-Oriented Measures ......................................................................... 23
  2.2  Problem Identification Approaches ....................................................................... 25
    2.2.1  Train-Oriented Approaches .............................................................................. 26
    2.2.2  Passenger-Oriented approaches ....................................................................... 27
  2.3  Summary .................................................................................................................. 28

Chapter 3  A Framework for Identifying Systematic Service Deterioration ...................... 31
  3.1  Framework Design .................................................................................................. 32
    3.1.1  Service Quality Measure ................................................................................ 32
    3.1.2  Problem Identification Approach .................................................................... 35
    3.1.3  Scope of the Framework .................................................................................. 37
  3.2  Stages of Passenger Journey and Disruptors ......................................................... 39
    3.2.1  Access Time ..................................................................................................... 41
    3.2.2  Waiting Time ................................................................................................... 42
    3.2.3  In-vehicle time ............................................................................................... 43
    3.2.4  Transfer Time ................................................................................................. 44
List of Figures

Figure 1-1: Service Improvement Process ................................................................. 17
Figure 1-2: Overall Framework .................................................................................. 18
Figure 1-3: MTR System Map .................................................................................. 19
Figure 3-1: Traditional Annual Timetable Update Process in JR-EAST ................. 36
Figure 3-2: Passenger Journey Stages ..................................................................... 39
Figure 3-3: Overall Framework ................................................................................ 47
Figure 3-4: Estimation of Standard Journey Time .................................................... 50
Figure 3-5: Estimation of Standard Walking Time ..................................................... 51
Figure 3-6: Excess Journey Time Detection ............................................................... 52
Figure 3-7: Sample Heat Map of Median Excess Journey Times ........................... 53
Figure 3-8: Problem Identification ........................................................................... 54
Figure 3-9: Heat Map Patterns for in-Vehicle Experience Issues .......................... 57
Figure 3-10: Heat Map Patterns for Issues Occurring at Origin Station ............... 59
Figure 3-11: Heat Map Patterns for Issues Occurring at Destination Station .......... 62
Figure 3-12: Standard Walking Time Estimation Using Passenger-to-Train Assignment Model 65
Figure 3-13: Problem Identification Using Passenger-to-Train Assignment Model .... 66
Figure 4-1: MTR’s Tsuen Wan Line ....................................................................... 70
Figure 4-2: Median Excess Journey Times for Each OD Pair by Time Window ....... 84
Figure 4-3: Daily Probability of the Median Excess Journey Time Exceeding 300 Seconds ..... 85
Figure 4-4: Median Excess Journey Time Components by OD Pair ..................... 87
Figure 4-5: Train Running Time Distribution between Stations 1 and 2 ............... 89
Figure 4-6: Headway Distributions at Stations 2 and 3 .......................................... 91
Figure 4-7: Comparison of Journey Time Distributions from Station 2 ............... 92
Figure 4-8: Fraction of passengers possibly traveling "backwards" from Station 2 .... 93
Figure 4-9: Comparison of Service Deterioration before and after the SIL Extension .... 95
Figure 4-10: Comparison of Service Deterioration Probability before and after the Line Extension .............................................................. 96
List of Tables

Table 3.1: Problem Detection and Management in Each Journey Stage ........................................ 40
Table 3.2: Summary of Pattern Analysis .......................................................................................... 56
Table 4.1: AVL Data Record ........................................................................................................... 73
Table 4.2: AFC Data Record .......................................................................................................... 74
Table 4.3: Train Arrival Delays at Each Station during the PM Peak Period .............................. 76
Table 4.4: Scheduled in-Vehicle Times and Expected Waiting Time during the off-Peak Period ............................................................................................................................................ 79
Table 4.5: Scheduled in-Vehicle Times and Expected Waiting Time during the PM Peak Period ............................................................................................................................................ 79
Table 4.6: Train Arrival Delays at Stations during the off-Peak .................................................. 81
Table 4.7: Estimated Standard Walking Time ................................................................................ 82
Chapter 1

Introduction

Thanks to the development of information technology and the spread of smart devices, it is becoming easier for rail passengers to access the operational information provided by rail operators through variable message signs, computers, or mobile devices. Furthermore, well-developed navigation applications such as Google maps can provide rough estimates of door-to-door journey times including walking time and waiting time. As a result, more and more passengers are aware of how long their journeys should take. In this respect, it can be said that the utility of rail transportation service has been enhanced. However, due to the increased information availability, passengers have also become more sensitive to the gap between what they expect of service and what was actually provided.

In terms of transportation performance, due in part to the nature of urban rail systems in dense areas—i.e., high passenger demand and high frequency of service—rail services are prone to have insufficient buffer time and are vulnerable to disruptions. That is to say, trains are frequently delayed by prolonged alighting and boarding processes due to the large volume of passengers (or other minor incidents), leading to further delays to the trains that follow them. These delays may also cause headway irregularity and further passenger crowding. Particularly in peak periods, recurrent systematic service deterioration can include not only delayed train operations but also congestion in stations, and on trains. Train capacity constraints may even force passengers to miss one (or more) trains, referred to as a “denied boarding”. These recurrent service deteriorations have a great influence on customer satisfaction in urban rail systems and become a serious concern especially in metropolitan areas such as Tokyo and Hong Kong [1],
Improvement of service quality is critical for agencies to attract and retain passenger demand in the face of growing competition from other transportation modes. To enhance performance, it is important to measure the current performance exhaustively and precisely.

Over the last few decades, rail agencies have developed and deployed Automatic Data Collection (ADC) systems for the sake of better control and monitoring. Automatic Vehicle Location (AVL) systems have enabled agencies to record the precise times of train arrivals and departures at each station. Using these data, agencies can obtain a complete picture of train operations and hence evaluate the on-time performance of each train. Additionally, the introduction of Automatic Fare Collection (AFC) systems has made it possible for agencies to capture individual passenger transaction records at fare gates. Based on these data, passenger journey time and its reliability can also be measured. However, in urban rail systems operating near-capacity, passenger-related performance can be strongly affected by passengers’ out-of-vehicle activities, including walking time and waiting time in stations, which cannot be directly captured from those data and existing analytical methods. Hence, a practical and systematic analytical method is required from agencies to make the most of these data to identify the service deteriorations which passengers frequently experience.

This thesis develops a framework for identifying systematic service deterioration in terms of passenger journeys in urban rail systems operating near-capacity; this aims to support transportation agencies’ problem understanding and management in the interest of service improvement.
1.1 Motivation

This study’s main motivation is the need for agencies to systematically identify and manage recurrent service deterioration from both the operators’ and passengers’ perspectives. As mentioned, a performance measure relative to the expectation from passengers is becoming more important due to the higher awareness of expected service quality. And these days, agencies have plenty of automatically collected data which should be great sources of service analysis supporting all phases of management including operations control, tactical and strategic planning. However, these assets are not yet fully utilized due to the lack of a clear process for analysis and evaluation.

In the last few decades, the performance of rail systems has been measured mainly from the operators’ perspective. Train-operation-related issues such as delays and headway irregularity can be clearly identified by analyzing AVL data, which provide all train station arrival and departure times. However, passenger journeys consist of both in-vehicle and out-of-vehicle activities including walking and waiting time in stations. Setting aside train delays, there are several potential disruptors of passenger journeys such as missed transfers, prolonged walking time due to congested walkways and denied boarding due to train capacity constraints. In fact, studies have shown that there is a large discrepancy between trains’ on-time performance and passengers’ perceptions [3]. Moreover, passenger crowding can lead to further train delays.

Accordingly, it is becoming more important to consider the passengers’ perspective and their out-of-vehicle experiences. Several transportation performance metrics from the passengers’ perspective have been developed such as excess journey time [4] and reliability buffer time [5], [6]. Those metrics deal with passenger journey time and its reliability, mainly
using AFC data which contain passengers’ entry and exit times. More detailed analytical methods of estimating denied boardings [7] have also been developed, combining AVL and AFC data.

Though these analytical methods have enabled agencies to measure service performance and to analyze specific problems in detail, a systematic problem detection and identification process have not yet been developed. In other words, it is often unclear when, where, in what part of the journey passengers face service deterioration, and what causes that deterioration. Consequently, agencies need to rely on reports from station staff, requests and complaints from passengers, and costly manual surveys to diagnose problems. Even then issues with regard to passengers’ out-of-vehicle experiences are still not comprehensively captured in a cost-effective manner.

To fill this gap, a systematic detection and identification methodology for service deterioration is needed. To this end, a simple process which focuses on service deterioration and problem identification should be developed. This is the aim of this thesis.

1.2 Research Objective

The objective of this thesis is to present a simple and general framework for systematic problem identification, in terms of the service gap passengers experience, for agencies with near-capacity operations in urban areas. The framework is designed to support agencies’ problem understanding and management, leading to further service improvements.

Figure 1-1 describes the focus of this thesis in the context of the service improvement process. The main focus is on the detection and identification of recurrent problems such as train delays and over-crowding which force passengers to spend more time to complete their journey
than they expect. The outputs from this framework should contain hotspot information on service deterioration including frequency and degree. As shown in Figure 1-1, these outputs are used as the input to the problem management and improvement steps.

![Figure 1-1: Service Improvement Process](image)

More specifically, the framework aims to answer the following questions:

- When and where do passengers experience service deterioration?
- How frequently do they experience this service deterioration?
- What are the underlying causes of this service deterioration?

This kind of information could be directly used to develop countermeasures, but it is also useful as an input for further micro-scale analyses such as train operation simulation at the block level and passenger flow simulation at the station level.

1.3 Approach

In this section, the specific approach for the development and implementation of the framework is described. The structure of the entire framework is presented in Figure 1-2.
First of all, a simple metric which reflects any service deterioration passengers may experience will be developed, incorporating the timetable, and AVL and AFC data. Thereafter, using the metric, significant service deterioration in passenger journeys between a specific pair of origin-destination (OD) stations and time period, is detected as a service problem at an aggregate level. Subsequently, the type and underlying causes of service deterioration are deduced by analyzing AVL and AFC data.

The framework will be tested and demonstrated by applying it to a line in the rail system of Hong Kong Mass Transit Railway (MTR). MTR is one of the heaviest ridership rail systems in the world, serving over 5.6 million passenger journeys on a typical weekday with 10 rail lines which connect all corners of the city of Hong Kong [8]. The total route length of the MTR system is 187.4 km with 91 stations. The MTR’s system map is presented in Figure 1-3.
To cope with such high demand, MTR provides high service frequencies with around 2-minute headways during peak periods, and with 4-6 minute headways during off-peak periods. However, despite the high service frequency, train overcrowding and denied boardings have been a concern for both passengers and MTR due to the increasing demand [2].

MTR uses high-quality AVL and AFC systems, which record train arrival (departure) times and passenger station entry and exit times in seconds. These data are used in a case study in this thesis.
1.4 Thesis Structure

The rest of this thesis is organized as follows. The literature on both train-oriented and passenger-oriented performance evaluation and problem identification are reviewed in Chapter 2, to clarify what has been previously done and what issues have not been resolved. Chapter 3 presents the framework and introduces the methodologies proposed in this thesis. Chapter 4 demonstrates the application of this framework by conducting a case study using the Hong Kong MTR rail system. A before and after analysis of transportation performance with regard to a line extension is also presented. Chapter 5 summarizes the conclusions and suggests future research directions.
Chapter 2

Literature Review

The purpose of this chapter is to review the literature related to transportation performance measures and problem identification methods in urban rail systems. There are a large number of studies focusing on performance measures [9], whereas fewer works are reported with regard to systematic problem identification explicitly considering passenger out-of-vehicle activities. The rest of this chapter is organized as follows. Section 2.1 reviews how transportation performance has been defined and measured in both a train-oriented manner and passenger-oriented manner in rail systems. Section 2.2 reviews studies which aim to support problem identification by agencies. Section 2.3 summarizes the gap that this thesis is intended to fill.

2.1 Transportation Performance Measures

This section reviews transportation performance measures with both a train-oriented focus and passenger-oriented focus in rail systems.

2.1.1 Train-Oriented Measures

As mentioned, train-oriented performance measures have been widely used by rail agencies for a long time [10]. There are several reasons for this: (1) a fundamental concept in rail systems is that trains operate according to a schedule [3], (2) train-flow is simple and has been relatively easy to record with AVL systems, whereas passenger-flow had not been until the introduction of AFC systems, and thus (3) rail operators often operate under contracts, with regard to train punctuality, which impose financial penalties if designated standards are not met [11].
Train-oriented measures can be broadly categorized into two groups: schedule adherence [12] and headway adherence [13]. Details of these measures are well-documented by Frumin [4]. Schedule adherence is commonly defined as the difference between the scheduled and actual train departure/arrival times. The most popular measure of schedule adherence may be on-time performance which represents the percentage of train arrivals at designated stations or terminal station within a specified time window. The length of this time window—i.e., a threshold for delay—varies by agencies. For example, London Overground and SMRT Trains, the public transit agency in Singapore, consider that a train is “on time” if it arrives at its terminal within 5 minutes of the planned timetable [14], [15]. The other type of schedule adherence is train travel time adherence which represents the difference between the terminal-to-terminal train travel time compared to defined time window and the scheduled running time. For instance, Hong Kong MTR considers a train as delayed when its excess travel time (relative to schedule) was 2 minutes or more [6].

There have been several criticisms of these schedule adherence measures. Henderson [16], [17] points out the following problems with those measures: (1) schedule adherence measures often evaluate only the lateness at terminal or limited time points, (2) waiting time of passengers are not considered, and (3) focusing on schedule adherence may lead to headway irregularity which may negatively affect service quality for passengers. Frumin also describes the core issue that these measures treat each train as independent, and thus train headways which can greatly affect passenger waiting time are neglected.

Another train-oriented measure, headway adherence, can meet some of the above criticisms. In urban rail systems operating high-frequency service, it can be assumed that passenger may not care about the schedule itself, but do care about the frequency of service and
waiting time \cite{18}. Since these passengers generally arrive randomly at stations, headway regularity is more important than schedule adherence in terms of passenger waiting time. In this respect, some public transit agencies, including bus operators, use headway adherence measures which reflect headway regularity. The percentage of actual headways within a certain range around the scheduled headway may be the most common measure. For example, Massachusetts Bay Transportation Authority measures the percentage of headways shorter than 1.5 times the scheduled headway and sets 95\% as the schedule adherence standard \cite{19}.

Though the adoption of headway adherence measures may be related to passenger waiting time, these measures are still train-oriented and might not adequately represent the passengers’ perspective. In the next section, measures which more directly reflect the passengers’ perspective are described.

2.1.2 Passenger-Oriented Measures

As has been discussed, the passengers’ perspectives are quite important in measuring transportation performance. Traditionally, passenger surveys have been the only way to evaluate the passengers’ perspective, but recently, the introduction of AFC systems has helped agencies deal with passenger-oriented measures in a more direct manner. However, since passenger perceptions vary from individual to individual and affected by many different attributes, it is much more complex to measure than are train-oriented measures. This section focuses on measures dealing with absolute terms—i.e., travel time, waiting time, walking time, and the number of transfers.

Before the introduction of AFC systems, there were some studies which tried to measure transportation performance from the passengers’ perspective. Wilson et al. estimated the average
waiting time of passengers using actual train headway data [20]. They also estimated the expected waiting time according to the schedule and defined the difference between those two measures of waiting time as “excess waiting time”. This measure indicates how much extra time passengers had to spend on average waiting for a train due to headway irregularity relative to the ideal schedule. London Transport extended this concept to whole passenger journeys. Automatic data obtained from the AVL system was used to estimate waiting time and in-vehicle time of passengers. Access time, egress time, and transfer time were also estimated based on manual surveys in major stations. The difference between average actual journey time and expected journey time including access time, waiting time, in-vehicle time, and egress time were estimated as “excess journey time.” This measure conceptually indicates that how much extra time a passenger had to spend for an entire journey.

Chan [21] and Frumin [4] estimated excess journey time more directly using AFC data for the London Underground and Overground system respectively. As London Underground’s AFC system provides passenger tap-in and tap-out times at fare gates, Chan was able to measure actual journey time more accurately without resorting to manual surveys. Chan also examined journey time distributions for each OD pair and proposed a reliability measure as the difference between the 95th and 50th percentile of actual journey time. This measure indicates the variability of journey time for each OD pair. Since the measure is not based on schedule, journey time variation for different lines can also be evaluated and compared with this reliability measure. Uniman named this measure “reliability buffer time” and extended it, distinguishing the effect of recurrent, systematic service deterioration and that due to incidents [5]. Wood further extended the reliability buffer time concept, considering journey time variability for individual passengers [6].
More broadly, the passengers’ perspectives can be captured in a generalized travel cost function. Several studies formulated this travel cost function, mainly in the interest of optimization of timetable design and rescheduling under disruptions [22], [23]. Nielsen et al. proposed travel cost function which includes the amount of time spent in each journey stage (including access time, waiting time, in-vehicle time, and egress time), disutility associated with not having a seat, and passengers’ perception of different routes [24]. In Nielsen’s formulation, the value of time for each journey stage was also considered. The idea of applying different values of time for different stages in a passenger journey is common in formulation of travel cost functions; however, the *Transit Capacity and Quality of Service Manual* points out that the valuation of time and activities for different stages of a journey can vary across passengers, and the average valuation can also depend on the culture and characteristics of the region [25]. This underlines the difficulty of using the value of time concept in general performance measurement.

Overall, many performance measures at different levels of detail have been developed for different purposes. However, most measures evaluate train movements or passenger journeys as a whole. Since passengers traveling in urban rail systems operating near-capacity experience a variety of service deterioration in each journey stage, those measures may not be suitable for identifying specific types of service deterioration as well as their causes. Studies focus on problem detection and identification are reviewed in the next section.

### 2.2 Problem Identification Approaches

This section reviews studies focused on systematic problem detection and identification in urban rail systems. These studies can be broadly categorized into two groups: train-oriented approaches and passenger-oriented approaches.
2.2.1 Train-Oriented Approaches

Based on historical train-flow data from AVL systems, several studies have been aimed at detecting train delays and identifying their causes. Flier et al. proposed a statistical approach to detect dependencies of train delays in the interest of timetable improvement [26], mainly focusing on delays caused by intentional holding to maintain connections and resource conflicts occurring on platforms and crossings. This approach suggests to timetable planners which dependencies are significant and thus should be eliminated. Olsson and Haugland applied regression analysis to identify the most relevant factors resulting in delays in the Norwegian rail system [27]. This study indicates that controlling boarding and alighting processes plays a key role to maintain train punctuality in rail systems serving dense areas.

Some studies using pattern recognition and data mining techniques have also been developed. Andersson et al. proposed an approach to measure the timetable robustness, focusing on critical points including where trains enter a line, or where trains are overtaken [28]. They plotted delay profiles and depicted how the cluster of the delay profiles vary depending on interventions by operations controllers. The study aims to tell controllers what to pay attention to, and how to handle frequent conflicts. Goverde and Meng developed a data mining tool to identify route conflicts and delay chains due to timetable design and capacity constraints [29]. Using delay trees, the tool classifies whether delays are primary or a secondary delay resulting from delay propagation. Fabrizio et al. proposed a method to identify recurrent train delay patterns in the Danish urban rail system [30]. Fabrizio applied the K-means clustering technique to train delay data obtained from the rail signal system, thereby systematically identifying delay patterns.
All those studies, including statistical and data mining approaches, have mainly focused on train delays, and thus passengers’ out-of-vehicle activities are not explicitly considered. As mentioned, in urban rail systems operating near-capacity, these out-of-vehicle activities, such as walking and waiting time, play an important role in the passenger experience. Hence service deterioration in those activities must be taken into account to identify service issues exhaustively.

2.2.2 Passenger-Oriented approaches

Since the introduction of AFC systems, a number of studies have been developed to estimate passenger flow and analyze out-of-vehicle issues. Paul proposed a passenger-to-train assignment model, incorporating AVL and AFC data for the London Underground [31]. Paul estimated the distribution of access and egress walking times, using passenger survey data and actual walking times derived from subsets of passengers who have only one feasible train itinerary. Zhu proposed a passenger-to-train assignment method based on Bayes’ Theorem and applied it to a line in the Hong Kong MTR system [7]. The studies estimated the probability of boarding each train in a passenger’s feasible set, using egress walking speed distribution. Zhu also developed a method which estimates denied boarding rates for passengers in busy stations. The results from this method were compared with MTR’s survey data, and showed similar results.

These studies enable agencies to identify train overcrowding and denied boarding problems at specific stations; however, the studies do not focus specifically on problem detection itself and thus service deterioration at each stage of the passenger journey is not exhaustively detected and identified. However, recently several studies have been developing in this direction.

Zhang et al. proposed a passenger journey time segmentation method, by clustering passenger groups who board the same train based on their tap-out time. [32]. The specification of
the passenger who did not wait and immediately board a train plays a key role in estimating walking time. Their approach assumes that passenger access and egress walkways between fare gates and the platform are the same and invariable from individual to individual, and thus access time is equal to egress time at the same station. That is, walking time variability among passengers due to personal factors, congestion rate on walkways, and different route/fare gate choices are neglected. These assumptions may affect the accuracy of segmentation in rail systems which have large stations and high passenger demand. Similarly, Lee et al. developed a passenger journey segmentation method, using the same clustering method as Zhang et al. [33]. However, their approach deals with less frequent services and does not explicitly consider denied boardings due to extremely high passenger demand.

Singh et al. proposed a journey time segmentation method and semiparametric regression model to identify the factors affecting journey time variability in each journey stage [34]. Their approach uses the result from the passenger assignment based on Bayes’ Theorem, using only AVL and AFC data. The advantage of this approach is that additional information such as passenger walking speed and distance are not required. They use passengers who have only one feasible train itinerary (called “unambiguous passengers”) to estimate egress time distribution; however, these unambiguous passengers’ walking times can be biased because passengers who have longer egress times than train headway are excluded from these subsets of passengers.

2.3 Summary

As reviewed in this chapter, many performance measures have been developed focusing on train operations. These measures are categorized as schedule adherence or headway adherence measures. Problem detection and identification methods in terms of train delays and their causes
have also been proposed, using either a statistical or data mining approach. However, none of
these methods have explicitly considered passenger journeys, in particular, their out-of-vehicle
time experience.

Along with the development and deployment of AFC systems, a number of performance
measures focusing on passenger journeys have also been developed; however, a limited number
of studies have dealt with passengers’ out-of-vehicle time experience which plays an important
role in urban rail systems operating near-capacity. Furthermore, those studies do not
comprehensively cover the features of such urban rail systems, for example, high frequency and
high passenger demand, with reasonable assumptions. In this respect, a simple but general
systematic approach for identification of service deterioration including passenger out-of-vehicle
time experience is needed to support agencies service improvement procedure.
Chapter 3

A Framework for Identifying Systematic Service Deterioration

This chapter presents a framework for detecting and identifying systematic service deterioration in urban rail systems operating near-capacity. The goal of this framework is to comprehensively identify recurrent service problems occurring in daily rail system operations, using the timetable, AVL and AFC data. Theoretically, the framework is applicable to journeys including transfers, however, non-transfer trips are the main focus.

The remainder of this chapter is organized as follows. Section 3.1 presents the framework design, discussing desirable performance measures and requirements for the framework for problem identification. The scope of the framework is also stated. Section 3.2 describes individual stages of a passenger journey and potential disruptors for each stage. Section 3.3 clarifies the data sources required to implement this framework. Section 3.4 introduces the overall framework. Section 3.5 explains how to estimate the standard journey time as a reference. Section 3.6 presents the process of problem detection based on excess journey time. Section 3.7 demonstrates the strategy of problem identification, decomposing excess journey time into in-vehicle time and out-of-vehicle time. Section 3.8 discusses the possible extension of the framework to include a passenger-to-train assignment model.
3.1 Framework Design

3.1.1 Service Quality Measure

This section discusses what kind of performance measure is suitable for problem detection and identification to improve transportation service. As described in Chapters 1 and 2, transportation services provided on rail systems can be evaluated focusing on many aspects such as transportation capacity, duration of train trips, the frequency of operations, reliability, and energy efficiency. There is no single universal measure which can be used to indicate the quality of service and to optimize it. However, a quantitative measure is indispensable to manage and improve any service, and the passenger experience is of the major concerns of rail agencies.

Transit Capacity and Quality of Service Manual defines the quality of service in terms of how well “transit service meets the needs of its customers” [25]. It also points out the importance of striking the right balance between passengers’ desires and agencies’ affordability, given the base demand for transportation service. To discuss what kind of measure is appropriate for performance evaluation in rail systems, here we introduce three significant notions regarding service quality: service standards, service delivery, and service experience.

First of all, service standards are defined as a designated quality of train operations. The standards may take the form of train timetables or various service attributes, for example, operational frequency and train trip time. Although the form of service standards differs across regions and rail operators, these standards could be interpreted as a promise (or “contract”) between an agency and its passengers. These days, many passengers are aware of how long their journey should take in advance, by checking information at the station or on websites and mobile applications in terms of these service standards. Some rail systems further provide real-time
operation information; however, passengers generally do rely on the standards when they plan their journeys. Hence, service standards can be regarded to some extent as expectations from passengers.

Next, service delivery is defined as the quality of provided transportation service from the operators’ perspective—e.g., schedule adherence, headway consistency, and transportation capacity provided. Service delivery can be measured explicitly using AVL data. Many transportation agencies report this kind of information, either voluntarily or as a requirement, as their service quality.

Lastly, service experience is defined as the indicator of service quality from the passengers’ perspective, including total passenger journey time, journey time reliability, the number of required transfers and crowding at stations and in trains. It should be noted that the value of time for different stages of a passenger journey, including in-vehicle time, walk time, and wait time, can vary [35]. For example, because out-of-vehicle activities (walking and waiting) require passengers to make physical effort and do not allow productive use of time, the time for those activities could be regarded as more onerous than in-vehicle time. Passengers’ on-time experiences are also important factors in this indicator. Since these factors are dependent on individual journey activities and preferences, it is difficult to comprehensively evaluate those factors with a single data source such as AFC data.

As mentioned, agencies used to employ service delivery as a performance measure, but recently, the idea of performance measurement based on service experience is drawing attention. Enhancement of the service experience is essential for agencies to attract and retain passenger demand. However, there are several difficulties in service improvement based on an indicator from the passengers’ perspective. Firstly, there is no single indicator or data source which
reflects all aspects of the service experience. The valuation of time and activities for different stages of a journey can vary across passengers, and the average valuation can also depend on the culture and characteristics of the region. Secondly, agencies have to balance the passengers’ perspective and operators’ perspective—i.e., affordability and feasibility of providing better service which often requires higher costs. Consequently, decision-makers still judge these problems without clear guidelines.

One possible direction decision-makers can take is setting the goal of minimizing the gap between service standards and service experience based on a simple relative measure. Considering service standards as a commitment to passengers, it is reasonable to pursue the minimization of the gap between the service expectation and the actual experience, referred to here as the service gap. In addition, the measure should be quantitatively comparable across different service segments to judge the priority for improvement. To this end, choosing a general but representative measure is imperative. Among several aspects of service experience, excess journey time relative to the standard journey time might be the most representative and is closely related to other aspects of inconvenience such as the on-time experience of passengers. In this context, unlike the scheduled train travel time, the standard journey time must include out-of-vehicle time such as walk time and wait time in stations. Missed-transfers, over-crowding and denied boarding are also implicitly reflected in the excess journey time because these incidents will generally increase passenger journey time. Therefore, the excess journey time of passengers could be used as the proxy for the service gap they experience.

On the other hand, there is a question (or downside) to adopting excess journey time as the performance measure for some kinds of analyses; the relative measure might not be suitable for a longitudinal analysis because it is based on the setting of service standards. Uniman points
out that relative performance measures based on the service standards (e.g., timetable) could be misleading for a longitudinal analysis because an intentional degradation in service standards may show an apparent improvement in this measure [5]. It is true that an absolute measure can better clarify the pure improvement in case of timetable updates; however, minimizing the service gap is still important even if the net performance is unchanged, taking into account the heightened passenger awareness of expected service performance. In this respect, it can be said that the relative measure focus of the service gap could be useful even for longitudinal analyses.

3.1.2 Problem Identification Approach

This section describes the conventional approach to problem identification and the requirement for a framework such as proposed in this thesis. As has been described, most rail agencies do not have a systematic and cost-effective methodology to detect and identify recurrent service problems based on service experience. Instead, they diagnose their daily transportation performance based on service delivery.

For instance, Hong Kong MTR, which is well-known for its high service standards and delivery, uses train travel time adherence as one of its key performance indicators. That is, a difference between the scheduled travel time and actual travel time of a train, from one terminal to the other, exceeding 2 minutes or more is regarded as a delay, or, service deterioration. Though MTR also uses a performance indicator from the passengers’ perspective, estimating the number of passengers affected by train delays over 5 minutes, passengers’ out-of-vehicle experiences are not captured by these two indicators. Hence, MTR has conducted manual surveys to monitor out-of-vehicle experiences, especially denied boardings occurring during the peak of the peak period. Recently, a method for estimating denied boarding has been developed.
that has enabled MTR to analyze that service problem systematically [7]. However, the systematic detection and identification of all forms of service deterioration has not been realized.

Figure 3-1: Traditional Annual Timetable Update Process in JR-EAST

Figure 3-1 presents the conventional timetable update process in East Japan Railway Company (JR-EAST), which is one of the largest rail agencies in the world. In terms of problem detection and identification, JR-EAST used to rely mainly on systematic manual inspections of AVL data. Thereafter, if there is a need for further analyses, the agency may conduct a field survey to determine the underlying cause of the problem(s). Recently, a new system that
evaluates the degree of service deterioration from the passengers’ perspective was introduced by JR-EAST [36]. However, the system mainly focuses on performance evaluation during significant service disruptions, and it does not serve the function of systematic problem identification. That is, underlying causes of recurrent service deterioration are not explicitly pursued. Hence, recurrent service deterioration in daily operations still cannot be systematically captured.

To achieve the goal of this thesis, providing a systematic method for problem detection and identification, the framework needs to meet the following requirements: (1) the framework reliably detects service deterioration relative to the service expectations of passengers, (2) the framework identifies the underlying causes of service deterioration, and (3) the framework focuses the need for improvement.

3.1.3 Scope of the Framework

This section presents features of targeted rail systems and use of the framework. First of all, the applicability of the framework is described as follows.

Operating Frequency

The main targets of this framework are urban rail systems with high-frequency operations. This is because the consideration of the out-of-vehicle experience is more important for systems with heavy ridership which can lead to systematic service deterioration. Hence, the framework assumes that passenger arrivals at origin stations are random, rather than based on the schedule. However, by adopting an appropriate model of passenger incidence behavior, the framework could (theoretically) be applicable to rail systems with any frequency.
Operational Complexity

The framework assumes homogeneous train operations with a single stopping pattern because this is most common for urban rail systems serving dense areas in the city center. This assumption makes it easier to estimate the standard journey time which passengers should expect. To deal with rail systems operating heterogeneous services including local, rapid, and express trains, certain route choice models of passengers such as, the fastest-route or utility-maximizing-route should be incorporated into this framework.

The use of the framework is described as follows.

Periodic Problem Detection and Identification

The main use of the framework is for periodic problem detection and identification, for example, on a monthly basis. The average transportation performance for each OD pair and time of day across the period can be measured by this framework. An analysis period longer than one month is preferable to obtain a sufficient sample size. Once agencies identify specific systematic problems such as recurrent train delays or denied boardings, they can develop countermeasures in all phases of management including operations control, tactical planning, and strategic planning based on the degree of service deterioration and its underlying causes.

Before and after Analysis

The framework can support before and after analysis of passenger transportation performance regarding various service changes. In terms of short-term changes in operations control (e.g., passenger-flow control by station staff), the effects of the changes can be measured using the framework. When it comes to changes in tactical planning such as timetable updates, as mentioned, agencies can see how the gap between service expectations and the experience of
passengers is affected. In addition, when rail systems are extended as a result of strategic planning, for example, the framework can also show how the extension affects the existing lines’ performance.

**Incident Analysis**

Though the main focus of this framework is the identification of recurrent service deterioration, the framework also provides the basic functions for incident analysis. Agencies can pick a specific day with significant incidents, thereby evaluating the influence of the incidents on passenger journey time and estimating the degree of disturbance on each stage of the journey.

### 3.2 Stages of Passenger Journey and Disruptors

In this section, components of the passenger journey and potential disruptors for each journey stages are introduced. Figure 3-2 presents the components of a passenger journey.

![Figure 3-2: Passenger Journey Stages](image)

First of all, here we define that a passenger journey starts when the passenger taps in at a fare gate in his (her) origin station and ends when he (she) taps out at a fare gate in his (her) destination station. Accordingly, the passenger journey time is defined as the time difference between passenger tap-in and tap-out. As shown in Figure 3-2, the passenger journey is decomposed into 4 (or 5) steps: access walking time from fare gate to platform at the origin.
station; waiting time for train boarding and departing at the platform; in-vehicle time (which may include multiple line segments); transfer time (if required); and egress walking time from platform to fare gate at the destination station.

Table 3.1: Problem Detection and Management in Each Journey Stage

<table>
<thead>
<tr>
<th>Journey stage</th>
<th>Potential disruptors</th>
<th>Identification approaches</th>
<th>Countermeasures</th>
</tr>
</thead>
</table>
| Access time   | ○ Walkway congestion  
○ Incidental passenger surges (special events) | ○ Manual observation/ closed-circuit television  
○ Microscopic flow analysis                      | Short term  
○ Passenger-flow management  
ex) passenger directions  
flow control  
Long term  
○ Facility investments |
| Waiting time  | ○ Headway irregularity  
○ Denied boarding (capacity constraints) | ○ AVL data analysis  
○ Manual observation/ closed-circuit television  
○ Microscopic flow analysis |
|               |                                                          |                                                               | Short term  
○ Headway control  
(manual/systematic)  
○ Unbalanced-train-load mitigation  
Long term  
○ Transportation demand management  
○ Capacity enhancement |
| In-vehicle time| ○ Train traffic congestion  
○ Prolonged alighting/boarding  
○ Incidents/severe weather | ○ AVL data analysis  
○ Manual observation | Short term  
○ Headway control  
(manual/systematic)  
○ Boarding control  
Long term  
○ Slack time allocation  
○ System upgrades |
| Transfer time | ○ Walkway congestion  
○ Missed connection (train delays)  
○ Denied boarding (capacity constraints) | ○ AVL data analysis  
○ Manual observation | Short term  
○ Transfer maintenance  
(manual/systematic)  
Long term  
○ Transportation demand management  
○ Capacity enhancement |
| Egress time   | ○ Walkway congestion  
○ Manual observation/ closed-circuit television  
○ Microscopic flow analysis |                                                               | Short term  
○ Passenger-flow management  
ex) passenger directions  
flow control  
Long term  
○ Facility investments |

40
Table 3.1 summarizes potential disruptors, current problem detection and identification approaches, and countermeasures in each journey stage. A description of each journey stage and details of Table 3.1 are described in the following sections.

3.2.1 Access Time

Access time is defined as the walking time of a passenger from fare gate to platform at the origin station. The walking speed may differ from person to person, depending on age, sex, and other personal factors; however, in general, the access time can be regarded as a function of walking distance and facilities including elevators, escalators, and steps under free-flow conditions [7].

Once congestion occurs in walkways, passengers’ walking speeds decrease, and passengers take longer time to reach the platform. In particular, this kind of delay often occurs at bottlenecks such as the beginning of steps and escalators during peak periods. Furthermore, incidental passenger surges, in relation to train disruptions and/or special events such as sports contests and concerts, can delay passengers. These service deteriorations may even lead to safety issues, and thus certain countermeasures should be taken to alleviate the situation. These situations might be detected through manual observation or closed-circuit television, but not all of these problems are necessarily detected currently because of the lack of a systematic approach.

Given that this kind of service deterioration is detected and identified with regard to access time, there are some approaches for improvement. First of all, the root cause of the problem must be defined. To this end, either a manual field survey or microscopic passenger-flow analysis can help explain the situation. Thereafter, depending on the result of the root cause analysis, agencies should develop countermeasures from short-term and long-term perspectives.
In the short-term, passenger directions toward less congested walkways with signage and flow control by station staff may mitigate the congestion. In the long-term, facility investment such as corridor widening and additional escalators can be considered as possible resolutions.

3.2.2 Waiting Time

Waiting time is defined as the duration between the time the passenger arrives at the platform and the time the train which the passenger takes departs from the origin station. As we assume that passenger arrivals are random, the expected mean waiting time of passengers is a function of the mean and variance of operational headways. Under ideal operations with uniform headways, the expected waiting time would be half the scheduled headway.

However, headway irregularity due to the stochasticity of train operations often increases expected waiting time. Moreover, in some stations which have high passenger demand during peak periods, passengers could be forced to miss one (or more) trains due to train capacity constraints resulting in denied boarding. It should be noted that an important factor in denied boarding is often high passenger loads on incoming trains. In such situations, waiting times for those passengers can greatly increase, negatively affecting a passenger experience. Unbalanced passenger loads across the train-set and unbalanced distribution of passengers waiting at a platform may also cause such denied boarding, even though there is still some room for boarding remaining. Some instances of those headway irregularity and denied boardings are captured by agencies because they greatly affect the passenger experience and generate attention. But again, not all of them are necessarily detected currently because of the lack of a systematic approach.

Given that those kinds of service deteriorations are detected and identified with respect to waiting time, there are some possible approaches for improvement. In the short-term, in terms of
headway irregularity, better headway control may decrease the expected waiting time. Regarding denied boarding, passenger directions to reduce the imbalance of passenger loads across the train-set and that of passenger distribution waiting at the platform can be considered in the short-term. Transportation demand management strategy which shifts and reduces peak demand could also be an option [37]. In the long-term, enhancement of transportation capacity and different train-stopping strategies to balance the capacity might be considered.

3.2.3 In-vehicle time

In-vehicle time is defined as the duration between the time when a train, which a passenger boards, departs from the origin station and the time when the train arrives at the destination station. In-vehicle time consists of the train running time for each segment between adjacent stations and the dwell time at each station. All these times are scheduled in the timetable and may vary by time of day and from workdays to holidays.

Running times may increase mainly due to train congestion on the line, that is, delays of preceding trains. Incidents such as train car and signal system troubles can also greatly increase running times. In addition, severe weather conditions including heavy rain, snow, and strong wind may also prevent operation in the scheduled time. The dwell times may increase mainly due to longer passenger alighting and boarding processes. In contrast to out-of-vehicle times including walking and waiting time, excess in-vehicle time as the sum of excess dwell times and excess running times can be estimated and detected directly using AVL data.

There are many factors which affect train delays, and countermeasures can be considered for each factor. For example, enhanced headway regularity with an advanced control strategy may enable more efficient capacity utilization and decrease the time required for passenger
alighting and boarding processes. In the long-term, slack time re-allocation according to the operational reality may reduce the gap between service expectations and passenger experiences. Moreover, improvement to the signaling system, train technology, and other facilities may even reduce running time itself and generate additional time for recovery.

3.2.4 Transfer Time

Transfer time is defined as the time a passenger spends after the arrival of the train from which the passenger alights at a transfer station until the departure of the train which the passenger boards at the station. Transfer time consists of walking time from an alighting platform to the boarding platform and waiting time for the train departure. In general, agencies have a rough estimation of the minimum required walking time for each transfer segment and many consider connections between trains when they plan timetables.

Transfer times may increase for a variety of reasons. First of all, walking time may increase due to congestion on walkways in the same way as for access time. Or passengers may be denied boarding due to capacity constraints in the same manner that passengers experience at their origin station. A delayed train arrival at the transfer station could also force passengers to miss a transfer train, and a delayed departure of the transfer train could increase waiting time for passengers. Though obvious missed-connections due to train delays can be captured using AVL, it is generally difficult to detect and identify the other causes of service deterioration without an individual position tracking system like GPS.

Given that service deterioration is identified, developing operations control strategies which consider transfers could be one option. However, since the operational frequency during peak periods in urban rail systems is high, it is not generally a major concern. Rather,
countermeasures for passenger crowding including denied boarding described previously may be more important.

3.2.5 Egress time

Egress time is defined as the walking time of a passenger at his (her) destination station. More specifically, here egress time represents the duration between the time the train arrives at the platform and the time he (she) taps out at the fare gate.

Characteristics of egress walking are almost the same as access walking. However, because passenger groups get off the same train almost simultaneously, the degree of congestion on walkways is likely to be higher, leading to a higher probability of excess egress time. As already mentioned, heavy congestion may cause safety problems, and hence it is desirable for agencies to capture and comprehensively manage heavy congestion.

Given that crowding problems are detected and identified, similar approaches can be taken as for access time service deterioration.

3.3 Data Sources

This section describes data sources required to support this framework. Timetables, automatic data from AVL and closed AFC systems will be used to analyze service deterioration. These days, most rail operators have AVL data, and it is becoming common for agencies to have closed AFC data. In this respect, the framework is widely applicable to rail systems worldwide.

3.3.1 Timetable

All rail operators have timetables to manage their rail systems and train operations, even though the timetables may not be available to the public, as with Hong Kong MTR. In general, they use
two or three different timetables for weekdays, Saturdays, and Sundays, to deal with the different levels of passenger demand.

Specifically, train departure and arrival times at each station are determined to a certain level of time precision, for example, 1, 5, 15, or 30 seconds. Running and dwell times usually vary by time of day and from timetable to timetable.

3.3.2 AVL System Data

As mentioned, many rail operators have AVL systems. The data format and data fields may differ from agency to agency, but in general, operation date, train trip ID, and their scheduled (actual) arrival and departure times at each station are recorded in AVL data.

There are often some quality problems affecting AVL data. To be more specific, there are often some duplicates or missing data in AVL records. Hence, a data cleaning process is usually required before analyzing these data.

3.3.3 AFC System Data

In terms of passenger-flow data, there are three situations depending on the nature of agencies’ fare collection systems: no passenger transaction records; entry-only passenger transaction records; and both entry and exit passenger transaction records (so-called “closed” AFC data). In addition, some agencies have only aggregate flow data, whereas others have disaggregate data.

In this framework, closed AFC data which retain individual passenger transaction records are used to analyze passenger journey time. Thanks to the development of automatic fare collection systems and smart IC cards, it is becoming common for agencies to have closed AFC systems with high-quality data. For example, Hong Kong MTR, JR-EAST, and rail systems of
Transportation for London (TfL) have high quality closed AFC systems. In general, card ID, entry and exit station, and entry and exit transaction times are retained in AFC data.

3.4 Overall Framework

In this section, an overall framework for detection and identification of systematic service deterioration in urban rail systems is proposed. Figure 3-3 illustrates the overarching approach. As displayed in Figure 3-3, this framework uses the timetable, AVL, and AFC data as inputs. The framework has three stages: estimation of standard journey time, detection of excess journey time, and problem identification.

![Figure 3-3: Overall Framework](image)

In the standard journey time estimation, the standard journey time for each time of the day, as a reference for observed journey times, is estimated as the sum of scheduled in-vehicle time, expected waiting time, and standard walking time. The standard walking time for each OD pair is also estimated in this step.

Next, by comparing the standard journey time and observed passenger journey times, the excess journey time of each passenger is calculated as a proxy for the deterioration in service...
experience. These figures are aggregated for each OD pair and each time of day, thereby representing aggregate performance. Consequently, hotspot information in terms of service deterioration is detected at the OD level and for a given time period.

Finally, based on results from the previous step, underlying causes of service deterioration are deduced and specified in the problem identification step. The final outputs from the framework contain hotspot information including time of the day, location, frequency, the degree of disturbance, and proximate causes of service deterioration. Details of each stage are described in the following sections.

3.5 Standard Journey Time Estimation

This section discusses definitions of the standard journey time and presents an estimation process. Here the standard journey time is defined as an ideal passenger journey time without any disturbances or binding capacity constraints. In other words, we assume that all trains operate perfectly according to the schedule, and passengers can walk through stations under free-flow conditions. In addition, passengers can always board the first train to arrive after they reach the platform. Hence, the standard journey time consists of scheduled in-vehicle time, expected waiting time, and standard walking time for each OD pair and time of day. Since the standard journey time is based in part on the timetable, it varies by time of day.

In terms of definitions of standard journey time, there are several options to choose from. For example, in Frumin’s excess journey time metric, the standard journey time consists of only the expected waiting time and scheduled in-vehicle time, assuming walking times are negligible [4]. However, as described in this chapter, service deterioration in walking times can be
significant in urban rail systems operating near-capacity. Therefore, the walking times are included in the standard journey time in this framework.

Another option for the standard journey time is an average passenger journey time, which represents typical daily transportation performance. For instance, Tsunoda adopts the median passenger journey time as a reference for observed passenger journey time under disruptions [36]. The typical journey time is a robust standard and widely applicable to rail systems; however, it also involves systematic service deterioration effects. As the objective of this framework is to shed light on systematic service deterioration, the ideal journey time is a better reference. If we adopt the typical journey time as a reference, the framework would focus more on another type of problem. That is, the variability of transportation performance would be emphasized.

The process of standard journey time estimation is displayed in Figure 3-4. The scheduled in-vehicle time and expected waiting time are calculated from the timetable data. As mentioned, agencies generally have a rough estimate of the minimum required walking time for each transfer segment. These are used as the scheduled transfer time. The standard walking time is also estimated as the sum of standard access time and egress time for each OD pair in this step. Details of each process are presented in the following sections.
3.5.1 Scheduled in-Vehicle, Transfer, and Expected Waiting Time

The scheduled in-vehicle time for each time period is obtained directly from the timetable. As mentioned, running times and dwell times vary by time of day in general, and thus the scheduled in-vehicle time also varies. The expected waiting time for each time period is calculated from the scheduled headway. Since we assume that passenger arrivals are totally random given the high frequency of operations, the expected waiting time under the scheduled operation is simply half the scheduled headway at all stations. If a passenger journey involves a transfer, a rough estimate of scheduled transfer time should be included in the standard journey time.

3.5.2 Standard Walking Time

This section presents a process for the standard walking time estimation for each OD pair, using the timetable and AFC data for the off-peak period. Here, the standard walking time is defined as the typical time a passenger needs to walk from fare gate to a platform at his (her) origin station and to walk from a platform to fare gate at his (her) destination station, under free-flow
conditions without any disturbance or congestion. Figure 3-5 illustrates the process of the standard walking time estimation.

![Diagram of standard walking time estimation process]

Figure 3-5: Estimation of Standard Walking Time

One can assume that passengers traveling during the off-peak period walk through stations under free-flow conditions. Assuming stable train operations with trains operating at even headways, the standard walking time is estimated as follows:

\[
S_{od}^{walk} = M_{T_{od, off}} - S_{od, off}^{in-vehicle} - E_{T_{o, off}}^{wait} - S_{t}^{transfer}
\]  

(3.1)

where \(S_{od}^{walk}\) is the standard walking time, \(M_{T_{od, off}}\) is the median journey time during the off-peak period as a proxy for the standard journey time, \(S_{od, off}^{in-vehicle}\) is the scheduled in-vehicle time during the off-peak period for the OD pair, \(E_{T_{o, off}}^{wait}\) is the expected waiting time at origin station during the off-peak, and \(S_{t}^{transfer}\) represents the scheduled transfer time at a transfer station. Subsequently, the standard journey time for each OD pair and time period is obtained as the sum of the scheduled in-vehicle time, walking time (and transfer time), and expected waiting time.

51
3.6 Excess Journey Time Detection

This section presents the process of excess journey time detection. Specifically, the time, location, degree, and frequency of service deterioration passengers experience are identified in the form of higher passenger journey times. Figure 3-6 illustrates the process of excess journey time detection.

For each passenger, the excess journey time is defined and calculated as the difference between the observed journey time and the standard journey time as described in the previous section. In terms of the aggregate performance for an OD pair for a time period over a date range, the median excess journey time is determined based on the individual passenger journey times for each OD pair, operation date, and a time of day. By changing the aggregation term, for example, weekly, monthly, and annual performance measures can be derived for workdays and holidays, respectively.

These median excess journey times can be clearly visualized with a heat map as shown in Figure 3-7. In the heat map, the columns represent destination stations, the rows represent origin stations, and each grid contains the median excess journey time for each OD pair.
In addition, by setting a threshold of the median excess journey time for significant service deterioration, the frequency of the deterioration can be measured. To be more specific, the frequency of exceeding the acceptable threshold for each OD pair and each time window can be estimated.

Consequently, this step systematically provides agencies with information about when, where and how frequently service deterioration occurs. Thereafter, causes of service deterioration are deduced in the problem identification step which is described in the next section.
3.7 Problem Identification

This section presents the problem identification process. To be more specific, the process narrows the causes of service deterioration and identifies the most important cause. Figure 3-8 displays the process of problem identification.

Based on the hotspot information obtained in the previous steps, the high excess journey times can be decomposed into excess in-vehicle time and excess out-of-vehicle time. As we do not know the train which each passenger boards, the in-vehicle time of the passenger cannot be observed. By contrast, the median train travel time can be directly obtained using AVL data. Here, the train travel time between any two stations is defined as the difference between the train departure time at one station and train arrival time at the other station (for OD pairs on different
lines, the train travel time should be the sum of each leg’s travel time and estimated transfer
time). It is unlikely that the median train travel time is very different from the median in-vehicle
passenger time. Hence, the median excess train travel time is used as a proxy for the median
excess in-vehicle time for each OD pair and time period in this framework. Subsequently, an
approximation of the median out-of-vehicle time can be estimated as the difference between the
median excess journey time and the median excess train travel time.

These two figures, the median excess train travel time and estimated excess out-of-
vehicle time, are used as the key inputs to a service deterioration pattern analysis aimed at
problem identification. The goal of this analysis is to deduce and specify in which stage of the
journey passengers experience service deterioration. In the service deterioration pattern analysis,
a visual inspection is conducted to identify specific variation patterns of those two figures by OD
pair, using heat maps. Thereafter, service deterioration is divided into two categories of issues—
i.e., in-vehicle issues and out-of-vehicle issues. Table 3.2 presents a summary of pattern analysis.
Details of each issue and its specification process are given in the following sections, using
illustrative data and heat maps.
Table 3.2: Summary of Pattern Analysis

<table>
<thead>
<tr>
<th>Issue category</th>
<th>Pattern analysis 1: heat map pattern</th>
<th>Pattern analysis 2: AVL/AFC data analysis</th>
<th>Proximate cause</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-vehicle issue</td>
<td>High excess train travel time over a specific segment</td>
<td>Frequent excess dwell time</td>
<td>Prolonged dwell time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Frequent excess running time</td>
<td>Prolonged running time</td>
</tr>
<tr>
<td>Out-of-vehicle issue (origin station)</td>
<td>High excess out-of-vehicle time from a specific origin station</td>
<td>Significant excess out-of-vehicle time</td>
<td>Prolonged waiting time (denied boarding)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High volume of tap-in passengers</td>
<td>Prolonged access time (congestion on walkways)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High headway irregularity</td>
<td>Prolonged waiting time (headway irregularity)</td>
</tr>
<tr>
<td>Out-of-vehicle issue (destination station)</td>
<td>High excess out-of-vehicle time to a specific destination station</td>
<td>High volume of tap-out passengers</td>
<td>Prolonged egress time (congestion on walkways)</td>
</tr>
<tr>
<td>Out-of-vehicle issue (transfer)</td>
<td>High excess journey time over a specific interchange</td>
<td>Frequent train delays</td>
<td>Prolonged waiting time (missed-connection/headway irregularity)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Denied boarding occurrence for departing passengers</td>
<td>Prolonged waiting time (denied boarding)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Congestion occurrence for departing passengers</td>
<td>Prolonged walking time (congestion on walkways)</td>
</tr>
</tbody>
</table>

3.7.1 In-Vehicle Issues

This section presents a process to specify the in-vehicle experience issues with heat maps and AVL analysis. Figure 3-9 displays example heat maps for an in-vehicle experience related issue.
with regard to the median excess journey time, train travel time, and estimated median out-of-vehicle time.

![Heat Map Patterns](image)

<table>
<thead>
<tr>
<th>(a) Median Excess Journey Time</th>
<th>(b) Median Excess Train Travel Time</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Estimated Excess out-of-Vehicle Time" /></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 3-9: Heat Map Patterns for in-Vehicle Experience Issues**

As can be seen from Figure 3-9 (a), the median excess journey times for passengers traveling through Station 4 show a steep rise and are consistently higher after passing this station. Given that the heat map for median excess train travel time, as a proxy for median excess in-
vehicle time, also shows the same variation pattern, and the heat map for estimated out-of-
vehicle time does not, in-vehicle time obviously affects this service deterioration. In this case,
the cause of the deterioration can be narrowed to the operational segment between Stations 3 and
4, including dwell time at Station 3 and running time between Stations 3 and 4. Thereafter, the
proximate cause of service deterioration can be specified, analyzing distributions of operation
time for those two segments using AVL data.

In general, if the median excess journey time and train travel time for passengers
traveling through a specific station (either a terminal or intermediate station) show a drastic rise
and are consistently large after passing the station, the service deterioration is likely to arise from
excess in-vehicle time. The immediate cause of the excess in-vehicle time can be seen using
AVL data. For a further analysis, a root cause analysis in terms of recurrent operational delays
should be undertaken. For example, if the proximate cause is prolonged running times, a
microscopic train-flow analysis which considers the features of the railway signaling system and
operations control system is required. If prolonged dwell times are the main cause, a microscopic
passenger-flow analysis and train-flow analysis may identify the root cause of service
deterioration.

3.7.2 Out-of-Vehicle Issues

This section presents a process for specifying the out-of-vehicle experience issues using heat
maps, incorporating AVL and AFC data analysis. Out-of-vehicle experience issues are further
divided into three sub-categories: (1) service deterioration occurs at origin stations, (2) service
deterioration occurs at destination stations, and (3) service deterioration arises through transfers.
The specifications for these three types of service deterioration are described below.
**Issues at Origin Station**

Figure 3-10 displays example heat maps for issues occurring at origin stations with regard to the median excess journey time, train travel time, and estimated median out-of-vehicle time.

![Heat Maps](image)

(a) Median Excess Journey Time  
(b) Median Excess Train Travel Time  
(c) Estimated Excess out-of-Vehicle Time

Figure 3-10: Heat Map Patterns for Issues Occurring at Origin Station

As shown in Figure 3-10 (a), the median excess journey times for passengers traveling from Station 3 are consistently high regardless of the destination station. Given that the heat map
for the estimated excess out-of-vehicle time displays the same variation pattern, there must be
something affecting for passengers entering at Station 3. In this case, either excess access time or
waiting time (or both) in Station 3 can be considered as the proximate cause. However, since it is
unlikely that excess access walking time solely leads to 300-second excess out-of-vehicle time,
denied boardings are suspected to occur and affect the excess out-of-vehicle time. To further
deduce the main cause, headway variability at Station 3 must be examined to see whether it
affects the waiting time, using AVL data. Meanwhile, the volume of tap-in passengers at Station
3 during the analysis time period must also be examined to infer the effects of congestion.

In general, if the median excess journey time and estimated excess out-of-vehicle time for
passengers traveling from a specific origin station are consistently high, the service deterioration
is likely to occur at that station. It should be noted that if excess out-of-vehicle times are
significant (e.g., hundreds of seconds), the excess times are likely to arise from excess waiting
time, more specifically, denied boarding. Given that passengers are forced to miss one (or more)
trains due to denied boarding, their waiting time will increase significantly. Moreover, if denied
boardings occur at a specific station, excess out-of-vehicle times for passengers traveling from
the next station could also increase due to the overcrowded trains coming from that specific
station. On the other hand, it is unlikely that excess access time would be very large due to
congestion on walkways, except for the extreme case (for example) of escalator breakdowns.

If the excess time is not significant, train headway variability and the volume of tap-in
passengers at the station must be inspected to deduce the main cause. To specify the proximate
cause more precisely, a macroscopic passenger-flow analysis or field survey focusing on the
detected location and time period should be undertaken.
It is also noteworthy to point out that, if the service deterioration seems to arise at the origin station, and the heat map for the estimated excess out-of-vehicle times shows irregular patterns such as consistent increase of out-of-vehicle time along the travel becoming longer or sudden decrease of the out-of-vehicle time over the specific interchangeable station, unusual passenger behavior should be suspected (e.g., existence of multiple route choices). In such case, the journey time distribution of passengers must also be examined.

**Issues at Destination Station**

Figure 3-11 displays sample heat maps for issues occurring at destination stations with regard to the median excess journey time, train travel time, and estimated median out-of-vehicle time. It can be clearly seen from Figure 3-11 (a), the median excess journey times for passengers traveling to Station 4 are consistently high regardless of the origin station. Given that the heat map for the estimated excess out-of-vehicle time displays the same variation pattern, something must be affecting passengers exiting from Station 4. Since the cause of service deterioration at destination stations is likely to be excess egress time, congestion on walkways in Station 4 may be the proximate cause of service deterioration. To narrow the congestion points, the volume of the tap-out passengers for each fare gate should be examined if the gate information is available.

In general, if the median excess journey time and estimated excess out-of-vehicle time for passengers traveling to a specific destination station are consistently high, the service deterioration is likely to occur at that station. To be more specific, excess egress time due to walkway congestion is plausibly the proximate cause of the deterioration. In order to specify the point of the problem, a macroscopic passenger-flow analysis or field survey focusing on the detected location and time period should be undertaken.
Issues with Transfers

As mentioned in Section 3.2.4, there are various reasons which may lead to deterioration in passenger experience through the transfer, and it is difficult to specify the cause of the deterioration without the position tracking system or passenger-to-train assignment model which can help understand passenger behavior. However, if the median excess journey time for
passengers transferring at a specific interchange station is consistently high, the service
deterioration is likely to occur at the transfer station. In terms of missed connections and
significantly longer waiting times due to the disrupted train operations, these service
deteriorations can be detected using AVL data. Moreover, assuming that excess walking time
and waiting time due to walkway congestion and train capacity constraints in the transfer station
can be approximated based on other service deterioration information—i.e., excess access,
waiting, and egress time at the station—effects of those causes of service deterioration can also
be estimated.

3.8 Potential Extension of the Framework

This section presents a potential extension of the framework proposed in this thesis,
incorporating a passenger-to-train assignment model to infer passenger-flow in rail systems. This
extension may enable the framework to identify the causes of service deterioration more
precisely because passenger journey time can be divided into in-vehicle time and out-of-vehicle
time more rigorously. Furthermore, out-of-vehicle time can be decomposed into access and
waiting time at the origin station, transfer time, and egress time at destination station. The
limitations of this approach is that, (1) passenger-to-train assignment method may require a
considerable amount of computation to deal with passenger journeys with transfers because there
are many possible itineraries, and (2) the output from the method may include some error in the
assignment which should be carefully considered when used for detailed analysis.

    Section 3.5.2 presented estimation of the standard walking time as the difference between
the median journey time and the sum of the scheduled in-vehicle time and expected waiting time
during the off-peak period. Section 3.7 presented the process of problem identification based on

63
the rough estimation of the median out-of-vehicle time, using the median train travel time as a proxy for the median in-vehicle time. These processes are easy to implement, and they only require timetable, AVL and AFC data, which are available for many agencies. Hence, the framework is widely applicable to many rail systems.

However, to identify problems lying in each passenger journey stage in more detail, passenger-flow inference plays a significant role. To this end, a passenger-to-train-assignment model could be incorporated into the framework. Though this thesis will not develop the passenger-to-train assignment model itself, studies have shown that a probabilistic approach can infer passenger-flow in urban rail systems with satisfactory precision [7]. Details of the framework extension are presented in the following sections.

### 3.8.1 Standard Walking Time Estimation

The benefit of the standard walking time estimation using passenger-to-train assignment model is that it enables the framework to estimate the standard access time and standard egress time using the entire population of passengers accessing and egressing the station, respectively. By contrast, the estimation process proposed in Section 3.5.2 only uses a subset of passengers traveling on a specific OD pair and estimates the standard walking time for the OD pair. In consequence, the estimation process using passenger-to-train assignment model may provide less-biased results.

Figure 3-12 shows the process of the standard walking time estimation using a passenger-to-train assignment model. There are several approaches to passenger-to-train assignment, but in general, it requires AVL and AFC data as inputs, and sometimes also passengers’ walking speed distribution. Once passengers are assigned to trains, time they spent at their origin stations and destination stations can be computed, incorporating AVL data. To be more specific, the time a
passenger spent at the origin station—i.e., the sum of access time and waiting time—can be obtained as the difference between a tap-in time and train departure time. In the same way, the time the passenger spent at destination station—i.e., the egress time—can be calculated as the difference between train arrival time and tap-out time.

Figure 3-12: Standard Walking Time Estimation Using Passenger-to-Train Assignment Model

To estimate the median access time as the standard access time for each station, the time a passenger spent at the origin station must be decomposed into access time and waiting time. Here, the expected waiting time calculated from AVL data is regarded as a proxy for the median waiting time of passengers, and thus the median access time is estimated as the difference between the median time spent at the origin station and expected waiting time. Meanwhile, the standard egress time for each station can be directly estimated as the median egress time. In the end, the standard walking time for each OD pair can be estimated as the sum of the standard access time at the origin station and standard egress time at the destination station.
3.8.2 Problem Identification

A passenger-to-train assignment model would also benefit the problem identification process, by making the framework capable of analyzing passenger excess time spent on each journey stage (i.e., time spent at the origin station, in-vehicle time, transfer time, and egress time) separately. Unlike the problem identification process proposed in Section 3.7, this extension enables the identification of excess transfer time explicitly. Moreover, further analysis such as correlation analysis between the volume of tap-in passengers and excess time at origin station, and correlation analysis between the volume of tap-out passengers and excess egress time at destination station could also be done. Figure 3-13 displays the process of problem identification using passenger-to-train assignment model.

Figure 3-13: Problem Identification Using Passenger-to-Train Assignment Model

Based on hotspot information with regard to high excess journey time obtained from previous steps, passenger-flow over the hotspots is estimated by applying the passenger-to-train
assignment model. Once passengers are assigned to specific trains, by comparing the train arrival/departure times and passenger tap-in/tap-out times, their excess time spent at the origin station, excess in-vehicle time, excess transfer time, and excess egress time can be calculated directly. To distinguish the effect of walkway congestion and denied boarding at the origin station, a method which estimates denied boarding rate can be applied [7]. Subsequently, the main factor of service deterioration is identified quantitatively. Thereafter, according to the degree of service deterioration calculated, further analysis can be conducted.

Another benefit of this approach is that, since agencies can measure service deterioration at a disaggregate level, in other words, at the individual level, the agencies can easily analyze the nature of service deterioration in more detail. For example, it is possible to analyze which passenger group exiting the same fare gate group experiences the worst service deterioration. The result may indicate the most crowded walkways in the destination station. Consequently, some further analyses can be conducted without resorting to micro-scale simulations.

3.9 Summary

This chapter proposed a metric and framework for detection and identification of systematic service deterioration in urban rail systems. The framework deduces and identifies the location, time, degree, frequency, and cause of service deterioration in a systematic manner using timetable, AVL and AFC data. The framework is designed mainly for periodic problem detection and identification for transportation performance, but it can also serve to support longitudinal analysis and significant incident analysis. The outputs can be used directly to take measures, whereas they can also be used as inputs for further analysis and detailed problem identification.
The potential extension of the framework using the passenger-to-train assignment model was also presented.

The application of the framework to a line in rail systems will be demonstrated in Chapter 4, using data on Hong Kong MTR system. An example of before and after analysis regarding a line extension will also be presented.
Chapter 4

Application to MTR

In this chapter, the framework proposed in the previous chapter is applied to one of Hong Kong MTR’s busiest lines, Tsuen Wan Line (TWL) in the PM peak period on weekdays. The objective is to demonstrate how to apply the framework to an operational rail line to identify examples of systematic service deterioration. Only non-transfer passenger journeys on days with no major service disruptions are considered, in order to examine the effect of recurrent minor incidents due to system capacity constraints.

Section 4.1 presents an overview of the Tsuen Wan Line. Section 4.2 clarifies data sources including MTR’s timetable, and the AVL and AFC data used in this application. Section 4.3 describes data preprocessing steps: (1) generating single train trips from AVL data and complete passenger journeys from AFC data, and (2) removing days with disrupted train operations and passengers who spent a much longer than expected time in the system. Section 4.4 describes how to estimate standard journey time as the sum of scheduled in-vehicle time, expected waiting time and standard walking time. Section 4.5 demonstrates the detection of excess journey time, by comparing observed journey time and standard journey time. Section 4.6 demonstrates the identification of the source of the excess journey time, whether it relates to in-vehicle or out-of-vehicle service deterioration. Finally, Section 4.7 presents the summary of the application and findings.
4.1 Introduction to TWL

The (partial) MTR system map is shown in Figure 4-1 with the Tsuen Wan Line (TWL) highlighted in red with stations numbered from 1 to 16. TWL is one of the busiest lines in the MTR’s rail system with a length of 16 km and 16 stations. It takes about 30 minutes to travel between Central Station on Hong Kong island, the southern terminus, and Tsuen Wan Station in Kowloon, the northern terminus. TWL provides an important cross-harbor link and serves the heart of Kowloon. On average, there are over million weekday passenger-boardings on TWL.

Figure 4-1: MTR’s Tsuen Wan Line
Central, Tsim Sha Tsui, and Mong Kok are among the ten busiest stations in the MTR network, each used by over 120,000 weekday passengers. During the PM peak period, trains and platforms are heavily crowded in the northbound direction. In particular, many passengers are left behind at Admiralty due to fully-packed trains coming from Central. (Some northbound passengers may even travel backwards from Admiralty to avoid this congestion.)

At Central and Admiralty, TWL has interchanges with the Island Line (indicated in blue in Figure 4-1). In December 2016, the South Island Line (SIL) opened with an interchange with TWL at Admiralty. Interchanges with the Kwun Tong Line are provided at Yau Ma Tei, Mong Kok, and Prince Edward. In addition, interchanges with the Tung Chung and West Rail lines exist at Mei Foo and Lai King respectively.

To avoid considering passenger journeys that could take place on alternative paths consisting of other lines, train trips and passenger journeys within the northbound segment between Central and Lai Chi Kok are considered in this application. Mei Foo is excluded from this analysis because some passengers, traveling from Central and Admiralty to Mei Foo, seem to use the Tung Chung Line and transfer at Lai King. In addition, passengers who travel between Central and Admiralty, Yau Ma Tei and Mong Kok, Yau Ma Tei and Prince Edward, Mong Kok and Prince Edward are excluded because they have an alternative path choice because of the presence of overlapping lines.

4.2 Data Sources

4.2.1 Timetable

The MTR timetable is not available to the public. MTR, through its website, provides passengers with service standards such as average train frequency and duration of trips for each OD pair.
While the average train frequency reported varies by time of day and between workdays and holidays, the scheduled train trip time between any OD pair is reported as constant throughout the day on the website.

From the operator’s perspective, to deal with the different levels of passenger demand, service is provided based on four different operational timetables: weekdays from Monday to Thursday, Fridays, Saturdays, and Sundays. Train departure and arrival times at each station are determined in seconds, along with frequency, running times (between consecutive stations) and station dwell times for each time period. In reality, running and dwell times also vary by time of day and from timetable to timetable.

During the analysis time period, from 17:00 to 20:00, on weekdays, the frequency of service is highest where trains operate at 2-minute headways. The peak-hour frequency is almost double that of the off-peak period when operational headways are close to 4 minutes.

4.2.2 AVL System Data

A sample of MTR’s AVL data record is presented in Table 4.1. AVL data include the timestamp of a record, train ID, station ID, platform ID, and scheduled and actual arrival (departure) times in seconds. Each record also contains a train block number which is labeled “i_train_nbr”. It represents a sequence of train trips and does not correspond to a single, half-cycle train trip from one terminal to the other terminal. As no single data field identifies a single train trip, an assignment of a train trip ID is needed to distinguish and analyze individual half-cycle train trips.
Table 4.1: AVL Data Record

<table>
<thead>
<tr>
<th>Data Field</th>
<th>Record Value</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>time_stamp</td>
<td>&quot;2017-03-13 17:50:23&quot;</td>
<td>Timestamp</td>
</tr>
<tr>
<td>line_id</td>
<td>1</td>
<td>Line ID (1: Tsuen Wan Line)</td>
</tr>
<tr>
<td>i_train_nbr</td>
<td>23</td>
<td>Train block number, train block consists of a sequence of train trips</td>
</tr>
<tr>
<td>dest_code</td>
<td>A</td>
<td>Destination code</td>
</tr>
<tr>
<td>time_code</td>
<td>A</td>
<td>Time code</td>
</tr>
<tr>
<td>train_nbr</td>
<td>44</td>
<td>Train vehicle id</td>
</tr>
<tr>
<td>id_type</td>
<td>A</td>
<td>Event type, &quot;A&quot;: arrival, &quot;D&quot;: departure</td>
</tr>
<tr>
<td>station_id</td>
<td>1</td>
<td>Station ID (1: Central, 10: Lai Chi Kok )</td>
</tr>
<tr>
<td>platform_id</td>
<td>20</td>
<td>Platform ID with value &quot;10&quot; or &quot;20&quot;</td>
</tr>
<tr>
<td>sch_time</td>
<td>&quot;2017-03-13 17:50:00&quot;</td>
<td>Scheduled arrival or departure time</td>
</tr>
<tr>
<td>act_time</td>
<td>&quot;2017-03-13 17:50:24&quot;</td>
<td>Actual arrival or departure time</td>
</tr>
<tr>
<td>lead_cab</td>
<td>205</td>
<td>Lead car id</td>
</tr>
<tr>
<td>trail_cab</td>
<td>208</td>
<td>Last car id</td>
</tr>
</tbody>
</table>

4.2.3 AFC System Data

A sample of AFC data record is presented in Table 4.2. MTR has a closed AFC system, that is, passengers are required to “tap” at a fare gate both when they enter and exit from the rail system. Each AFC transaction record is retained by the fare collection system, with each record containing card ID, entry or exit station, entry or exit transaction time in seconds, hardware-type code and fare gate number used for the transaction.
Table 4.2: AFC Data Record

<table>
<thead>
<tr>
<th>Data Field</th>
<th>Record Value</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSC_PHY_ID</td>
<td>964213271</td>
<td>Anonymized smart card id</td>
</tr>
<tr>
<td>BUSINESS_DT</td>
<td>14/03/2017</td>
<td>Date of the transaction (from 5:00 to 04:59 the following day)</td>
</tr>
<tr>
<td>TXN_DT</td>
<td>14/03/2017</td>
<td>Transaction date and time</td>
</tr>
<tr>
<td></td>
<td>14:32:05</td>
<td></td>
</tr>
<tr>
<td>TXN_TYPE_CO</td>
<td>ENT</td>
<td>Transaction type, &quot;ENT&quot;: entry, &quot;USE&quot;: use with fare deduction (i.e., exit)</td>
</tr>
<tr>
<td>TXN_SUBTYPE_CO</td>
<td>ADL</td>
<td>Transaction subtype code, &quot;ADL&quot;: adult, &quot;CHD&quot;: child, etc.</td>
</tr>
<tr>
<td>TRAIN_ENTRY_STN</td>
<td>1</td>
<td>Entry station id</td>
</tr>
<tr>
<td>TXN_LOC</td>
<td>1</td>
<td>Exit station ID for a USE record</td>
</tr>
<tr>
<td>TXN_AUDIT_NO</td>
<td>3452</td>
<td>Transaction audit number, all the transactions under this card ID are</td>
</tr>
<tr>
<td></td>
<td></td>
<td>numbered sequentially</td>
</tr>
<tr>
<td>HW_TYPE_CO</td>
<td>6</td>
<td>Hardware type code, 6: gate, 7: add value machine, etc.</td>
</tr>
<tr>
<td>MACH_NO</td>
<td>G06</td>
<td>Number of the device used for this transaction</td>
</tr>
<tr>
<td>TRAIN_DIRECT_IND</td>
<td>2</td>
<td>Direction indicator (with value of 1 or 2)</td>
</tr>
<tr>
<td>TXN_VALUE</td>
<td>7</td>
<td>HK $ value deducted for this transaction</td>
</tr>
<tr>
<td>MODAL_DISC_VALUE</td>
<td>0</td>
<td>HK $ value discounted</td>
</tr>
<tr>
<td>CSC_RV_VALUE</td>
<td>7</td>
<td>Remaining amount in HK $</td>
</tr>
</tbody>
</table>

4.3 Data Preprocessing

4.3.1 Generation of Half-Cycle Train Trips and Complete Passenger Journeys

Train Trips

As mentioned in the previous section, an AVL data record does not contain an identifier for each
half-cycle train trip from one terminal station to the other. Therefore, in order to identify sequential departure and arrival events for each train trip, a trip ID must be defined.

The process of assigning a train trip ID is as follows. Firstly, the direction of each trip can be determined from the destination code. Then, by merging information on train ID, the direction with the trip and the order of the trip with that train ID and direction, a unique train trip ID is determined. The train trip ID “56_UP_2” means that this trip is operated by train No. 56 in the northbound direction for the second time that day.

**Passenger Journeys**

As each AFC data record corresponds to only one transaction (an entry or exit event of a passenger), pairs of passenger transaction records must be linked to form complete journeys. The pair of passenger transactions can be identified by matching the card ID, entry station code, and consecutive transaction numbers. Thus, by matching the two entry and exit transaction records, the complete record for each passenger journey is obtained.

**4.3.2 Outlier Removal**

**Disrupted Train Operations**

Since the framework proposed in this thesis aims to identify recurrent and systematic problems, days with major incidents should be excluded from this analysis. MTR reports the number of passengers affected by train delays over 5 minutes to the government. For this application, we classified train delays over 5 minutes (300 seconds) as major incidents.

Statistics of delays on all train arrivals at all stations relative to the schedule, from 17:00 to 20:00 for weekdays (except for Fridays) in March 2017, are shown in Table 4.3. The sample size is over 2,500 for each day. As can be seen from the table, the maximum arrival delays
exceeded 300 seconds on four days—March 2, 9, 15, and 16. Therefore, those days were excluded from this application. On all other days, there was no major incident; the maximum delays were below 300 seconds, and the standard deviations were below 60 seconds.

Table 4.3: Train Arrival Delays at Each Station during the PM Peak Period
Data source: train operations between 17:00 and 20:00, on weekdays (except Fridays) in March 2017

<table>
<thead>
<tr>
<th>Date</th>
<th>Sample size</th>
<th>Mean (sec)</th>
<th>Std. (sec)</th>
<th>Max (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2017-03-01</td>
<td>2,571</td>
<td>22.8</td>
<td>26.1</td>
<td>127</td>
</tr>
<tr>
<td>2017-03-02</td>
<td>2,586</td>
<td>108.5</td>
<td>111.3</td>
<td>444</td>
</tr>
<tr>
<td>2017-03-06</td>
<td>2,570</td>
<td>16.5</td>
<td>24.4</td>
<td>141</td>
</tr>
<tr>
<td>2017-03-07</td>
<td>2,570</td>
<td>24.1</td>
<td>26.4</td>
<td>117</td>
</tr>
<tr>
<td>2017-03-08</td>
<td>2,570</td>
<td>15.0</td>
<td>20.6</td>
<td>127</td>
</tr>
<tr>
<td>2017-03-09</td>
<td>2,570</td>
<td>59.6</td>
<td>69.2</td>
<td>319</td>
</tr>
<tr>
<td>2017-03-13</td>
<td>2,585</td>
<td>37.5</td>
<td>44.4</td>
<td>240</td>
</tr>
<tr>
<td>2017-03-14</td>
<td>2,570</td>
<td>13.7</td>
<td>16.0</td>
<td>96</td>
</tr>
<tr>
<td>2017-03-15</td>
<td>2,569</td>
<td>81.3</td>
<td>77.6</td>
<td>318</td>
</tr>
<tr>
<td>2017-03-16</td>
<td>2,569</td>
<td>121.3</td>
<td>132.6</td>
<td>430</td>
</tr>
<tr>
<td>2017-03-20</td>
<td>2,570</td>
<td>12.6</td>
<td>17.1</td>
<td>108</td>
</tr>
<tr>
<td>2017-03-21</td>
<td>2,570</td>
<td>64.8</td>
<td>55.8</td>
<td>198</td>
</tr>
<tr>
<td>2017-03-22</td>
<td>2,569</td>
<td>12.7</td>
<td>20.1</td>
<td>189</td>
</tr>
<tr>
<td>2017-03-23</td>
<td>2,571</td>
<td>12.6</td>
<td>16.5</td>
<td>175</td>
</tr>
<tr>
<td>2017-03-27</td>
<td>2,570</td>
<td>34.5</td>
<td>45.8</td>
<td>176</td>
</tr>
<tr>
<td>2017-03-28</td>
<td>2,570</td>
<td>28.4</td>
<td>36.9</td>
<td>216</td>
</tr>
<tr>
<td>2017-03-29</td>
<td>2,570</td>
<td>13.1</td>
<td>18.3</td>
<td>128</td>
</tr>
<tr>
<td>2017-03-30</td>
<td>2,571</td>
<td>9.5</td>
<td>13.6</td>
<td>179</td>
</tr>
</tbody>
</table>

**Passenger Behavior**

Regarding individual passenger behavior, those who intentionally spent a much longer time in the system than necessary should be excluded from the analysis. We are interested in passengers who walk directly between the fare gate and the platform. However, commercial areas do exist
inside the paid area, and some passengers will spend extra time there, or make phone calls, or go to a bathroom, or wait for their companions. Those passenger behaviors can be identified and excluded as outliers by examining the journey time distribution. In other words, passenger journeys which are significantly longer than others between the same OD pair in the same time window on the same day are classified as outliers.

In this pre-processing step, the interquartile range method (NIST/SEMATECH, 2013) is applied to classify (and remove) outlier passenger journeys with extremely long journey times as follows:

\[
J_{Ti}^{\text{outlier}} > J_{od}^{75} + IQR \times 1.5
\]

where \( J_{Ti}^{\text{outlier}} \) is the observed journey time of passenger journey \( i \) determined to be an outlier, and \( J_{od}^{25} \) and \( J_{od}^{75} \) are the 25 and 75 percentile journey time for each OD pair. Passenger journeys are aggregated by OD pair for each time of day.

By applying the interquartile range method for AFC data on weekdays (except Fridays) in March 2017, approximately 5% (35,723 out of 711,730) of passenger journeys within the targeted line segment were defined as outliers during the PM peak period, from 17:00 to 20:00. In the same way, about 5% (36,670 out of 771,433) of passenger journeys during the off-peak period, from 12:00 to 16:00, were defined as outliers. Those figures were relatively small, and it is reasonable to think that these fractions of passengers might spend more time, on extra activities while within the system.
4.4 Standard Journey Time Estimation

This section estimates the standard journey time as the sum of scheduled in-vehicle time, expected waiting time, and standard walking time for each OD pair and time period. The standard journey time is defined as an ideal journey time without any disturbances or binding capacity constraints. That is, all trains operate according to the schedule, and passengers can walk through a station under free-flow conditions. They can always board the first train to arrive after they reach the platform.

The standard walking time represents the sum of the standard access walking time at the origin station and standard egress walking time at the destination station for each OD pair, assumed to be the time an average passenger spends walking under free-flow conditions. While the standard in-vehicle time and expected waiting time vary across different timetables and time periods, the standard walking time is reasonably considered as a constant value for each OD pair.

Both the scheduled in-vehicle time and expected waiting times can be determined from the timetable. In order to estimate the standard walking time for each OD pair, observed passenger journey times and train trip times in the off-peak period, with moderate passenger demand and no major disruptions in train operations, are used.

4.4.1 Scheduled In-Vehicle Time and Expected Waiting Time

A sample of scheduled in-vehicle times and expected waiting time during the off-peak and PM peak periods are shown in Table 4.4 and Table 4.5. The scheduled in-vehicle time for each time period is obtained directly from the timetable. As mentioned, MTR’s trip navigation tool on the website provides constant train trip times for each OD pair. However, additional time is included in the schedule during the PM peak period, recognizing that the higher passenger demand will
increase train dwell times at stations. Consequently, the scheduled in-vehicle time during the PM period is longer than that during the off-peak period. For example, the scheduled in-vehicle time from Central to Yau Ma Tei is 504 seconds during the off-peak period, whereas that during the PM peak period is 559 seconds.

Table 4.4: Scheduled in-Vehicle Times and Expected Waiting Time during the off-Peak Period

<table>
<thead>
<tr>
<th>Origin code</th>
<th>Destination code</th>
<th>Origin name</th>
<th>Destination name</th>
<th>Scheduled in-vehicle time (sec)</th>
<th>Expected waiting time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>CEN</td>
<td>ADM</td>
<td>86</td>
<td>106.5</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>CEN</td>
<td>TST</td>
<td>282</td>
<td>106.5</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>CEN</td>
<td>JOR</td>
<td>389</td>
<td>106.5</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>CEN</td>
<td>YMT</td>
<td>504</td>
<td>106.5</td>
</tr>
</tbody>
</table>

Table 4.5: Scheduled in-Vehicle Times and Expected Waiting Time during the PM Peak Period

<table>
<thead>
<tr>
<th>Origin code</th>
<th>Destination code</th>
<th>Origin name</th>
<th>Destination name</th>
<th>Scheduled in-vehicle time (sec)</th>
<th>Expected waiting time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>CEN</td>
<td>ADM</td>
<td>86</td>
<td>60</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>CEN</td>
<td>TST</td>
<td>304</td>
<td>60</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>CEN</td>
<td>JOR</td>
<td>436</td>
<td>60</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>CEN</td>
<td>YMT</td>
<td>559</td>
<td>60</td>
</tr>
</tbody>
</table>

The expected waiting time for each time period is calculated from the scheduled headway. Since we assume that passenger arrivals are random, the expected waiting time under the scheduled operation is half the scheduled headway at all stations, 106.5 seconds during the off-peak period and 60 seconds during the PM peak period.

4.4.2 Standard Walking Time

In this section the standard walking time is estimated for each OD pair, using the timetable and AFC data for the off-peak period. The standard walking time represents the time an average
passenger takes to walk from fare gate to platform at his (her) origin station and to walk from
platform to fare gate at his (her) destination station, without any disturbances or congestion.

One can assume that passengers traveling during the off-peak period walk through
stations under free-flow conditions. Assuming stable train operations with trains operating at
even headways, the standard walking time is estimated as follows:

\[
S_{\text{walk}}^{\text{od}} \equiv M_{\text{JT}}^{\text{od, off}} - S_{\text{od, off}}^{\text{in-vehicle}} - E_{W, o, \text{off}}
\]  

where \( S_{\text{walk}}^{\text{od}} \) is the standard walking time for OD pair \( od \), \( M_{\text{JT}}^{\text{od, off}} \) is the median journey time
during the off-peak period for the OD pair as a proxy for the standard journey time, \( S_{\text{od, off}}^{\text{in-vehicle}} \)
is the scheduled in-vehicle time during the off-peak period for the OD pair, and \( E_{W, o, \text{off}} \)
represents the expected waiting time during the off-peak at the origin station \( o \).

To estimate the standard walking time using the above approximation, passenger
journeys during the off-peak period under stable train operations need to be identified, and their
journey times are used for this estimation. Statistics of delays on all train arrivals at all stations
relative to the schedule, from 12:00 to 16:00 for each weekday except for Fridays in March 2017,
were calculated and are shown in Table 4.6. The sample size is over 2,000 for each day. As can
be seen from the table, train operations during the off-peak period were stable for most days. In
particular, train operations on seven days—March 1, 2, 6, 7, 8, 9, and 13—were quite stable with
maximum arrival delays below 100 seconds, and mean delays below 10 seconds. Hence, those
days were selected for and used to estimate the standard walking time.
Table 4.6: Train Arrival Delays at Stations during the off-Peak
Data source: train operations between 12:00 and 16:00,
on weekdays (except Fridays) in March 2017

<table>
<thead>
<tr>
<th>Date</th>
<th>Sample size</th>
<th>Mean (sec)</th>
<th>Std (sec)</th>
<th>Max (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2017-03-01</td>
<td>2,028</td>
<td>3.1</td>
<td>6.0</td>
<td>60</td>
</tr>
<tr>
<td>2017-03-02</td>
<td>2,055</td>
<td>2.3</td>
<td>5.9</td>
<td>98</td>
</tr>
<tr>
<td>2017-03-06</td>
<td>2,026</td>
<td>4.3</td>
<td>8.6</td>
<td>96</td>
</tr>
<tr>
<td>2017-03-07</td>
<td>2,027</td>
<td>2.5</td>
<td>6.2</td>
<td>70</td>
</tr>
<tr>
<td>2017-03-08</td>
<td>2,055</td>
<td>2.8</td>
<td>7.1</td>
<td>70</td>
</tr>
<tr>
<td>2017-03-09</td>
<td>2,026</td>
<td>2.8</td>
<td>6.2</td>
<td>81</td>
</tr>
<tr>
<td>2017-03-13</td>
<td>2,026</td>
<td>3.4</td>
<td>6.8</td>
<td>83</td>
</tr>
<tr>
<td>2017-03-14</td>
<td>2,027</td>
<td>3.7</td>
<td>9.4</td>
<td>117</td>
</tr>
<tr>
<td>2017-03-15</td>
<td>2,026</td>
<td>4.0</td>
<td>12.8</td>
<td>156</td>
</tr>
<tr>
<td>2017-03-16</td>
<td>2,027</td>
<td>4.2</td>
<td>10.8</td>
<td>122</td>
</tr>
<tr>
<td>2017-03-20</td>
<td>2,025</td>
<td>3.0</td>
<td>7.7</td>
<td>100</td>
</tr>
<tr>
<td>2017-03-22</td>
<td>2,115</td>
<td>5.0</td>
<td>16.0</td>
<td>175</td>
</tr>
<tr>
<td>2017-03-23</td>
<td>2,115</td>
<td>3.7</td>
<td>9.2</td>
<td>119</td>
</tr>
<tr>
<td>2017-03-27</td>
<td>2,115</td>
<td>4.0</td>
<td>9.6</td>
<td>110</td>
</tr>
<tr>
<td>2017-03-28</td>
<td>2,027</td>
<td>3.7</td>
<td>9.7</td>
<td>114</td>
</tr>
<tr>
<td>2017-03-29</td>
<td>2,028</td>
<td>3.2</td>
<td>9.5</td>
<td>119</td>
</tr>
<tr>
<td>2017-03-30</td>
<td>2,027</td>
<td>3.1</td>
<td>7.9</td>
<td>114</td>
</tr>
</tbody>
</table>

During the analysis time period on those days, 279,705 passenger journeys were recorded in total, with a sample size of over 900 for each OD pair. The resulting standard walking times for a sample of OD pairs are shown in Table 4.7. As can be seen from the table, the standard walking time differs depending on the OD pair. For example, the estimated standard walking time for passengers traveling from Central to Jordan, Yau Ma Tei, and Mong Kok are greater than for those passengers traveling from other origin stations. These differences are due to the fact that Central is a large station where walking distances between most fare gates and the platform at Central are greater than those for other stations.
Table 4.7: Estimated Standard Walking Time
Data source: train operations and passenger journeys between 12:00 and 16:00, on weekdays (except Fridays) in March 2017

<table>
<thead>
<tr>
<th>Origin code</th>
<th>Destination code</th>
<th>Origin name</th>
<th>Destination name</th>
<th>Sample size</th>
<th>Standard walking time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>CEN</td>
<td>JOR</td>
<td>8,721</td>
<td>250.5</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>CEN</td>
<td>YMT</td>
<td>5,490</td>
<td>263.5</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>CEN</td>
<td>MOK</td>
<td>10,850</td>
<td>283.5</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>ADM</td>
<td>JOR</td>
<td>4,854</td>
<td>161.5</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>ADM</td>
<td>YMT</td>
<td>3,144</td>
<td>159.5</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>ADM</td>
<td>MOK</td>
<td>5,692</td>
<td>171.5</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>TST</td>
<td>JOR</td>
<td>8,272</td>
<td>132.5</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>TST</td>
<td>YMT</td>
<td>10,549</td>
<td>132.5</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>TST</td>
<td>MOK</td>
<td>23,111</td>
<td>140.5</td>
</tr>
</tbody>
</table>

4.5 Excess Journey Time Detection

This section describes the identification of when, where and how frequently passengers experience service deterioration in the form of higher passenger journey times. For each passenger, excess journey time can be calculated as the difference between the observed journey time and the standard journey time as described in the previous section. Also, the median excess journey time across all passengers associated with an OD pair, traveling on a specific operation date, and starting their trips (entry time at the origin station) within a specific 15-minute time window can be regarded as a measure of performance for that OD pair on that date and for the given time range.

To detect service deterioration recurring systematically across many days, the median excess journey time for an OD pair and 15-minute time window determined based on all passengers traveling across for weekdays (except Fridays) in March 2017 was computed to
represent the monthly performance. The results for 15-minute time periods ranging from 17:45 through 19:15 are shown in Figure 4-2. As can be seen in the figure, the median excess journey times exceeds 100 seconds for some OD pairs between 18:00 and 19:00, which implies that service deteriorates during that time period. In particular, there seemed to be appreciable service deterioration indicated by excess journey times over 300 seconds for passengers entering Station 2 (Admiralty) between 18:15 and 18:45. Because the median excess journey times from Admiralty were high regardless of the destination station, the deterioration must have occurred at Admiralty and/or on train operations between Station 2 and Station 3 (Tsim Sha Tsui).

It is also noteworthy to point out that for some time periods there are negative values of the median excess journey time. This result could occur for several reasons. Firstly, the standard walking time could be overestimated in the previous steps. Since the standard walking time is estimated based on the AFC data during the off-peak period, there might be more children and seniors on the system whose typical walking speeds may be slower than other travelers. In addition, some off-peak passengers might be unfamiliar with the line and thus spend more time in stations compared to commuters. Consequently, the standard walking time could be overestimated. Secondly, the median waiting time of passengers could be shorter than the expected waiting time. Though we assume random passenger arrivals, passengers may rush to catch a train if they see one arriving. This may decrease the median time they spend at their origin station. For these reasons, the value of the median excess journey time could be negative.
Figure 4.2: Median Excess Journey Times for Each OD Pair by Time Window

Data source: passenger journeys on weekdays (except Fridays) in March 2017
Next, by setting a threshold for the median excess journey time to indicate appreciable service deterioration, the frequency of the deteriorations can be measured. Specifically, the frequency of exceeding the acceptable threshold for each OD pair and each time window can be estimated. In this application, we set the threshold to 300 seconds based on MTR’s delay-reporting threshold. Figure 4-3 shows the estimated daily probability of appreciable service deterioration for each OD pair from 18:15 to 18:45 based on the 14 weekdays in March 2017.

As shown in the figure, passengers traveling from Station 2 to Station 5 (Yau Ma Tei) and Station 6 (Mong Kok) during the peak of the peak period were likely to experience appreciable service deterioration with around 10 to 20% probability. Meanwhile, those traveling from Station 2 to Station 7 (Prince Edward), Station 8 (Sham Shui Po), Station 9 (Cheung Sha Wan), and Station 10 (Lai Chi Kok) during that period were likely to experience appreciable service deterioration with probability in the range of 30 to 60%. By contrast, passengers traveling from stations other than Station 2 do not experience such service deterioration in daily performance during the 14 days considered. This result clearly indicates the (relatively) poor
service quality received by passengers traveling from Station 2 in the peak of the peak period.

The degree of service deterioration and its frequency for each OD pair and time window were recorded as “hotspots”. This information is used as an input to the problem identification step described in the next section.

4.6 Problem Identification

This section demonstrates a process to narrow the causes of service deterioration and specify the most important cause. In other words, this step tries to provide an answer to the question “Why did passengers experience excess journey time?” Based on the hotspot information obtained in the previous section, the high excess journey times are decomposed into journey components including in-vehicle time and out-of-vehicle time.

In the absence of a passenger-to-train assignment method, we do not know the train which each passenger boarded, and so the in-vehicle time of the passenger cannot be directly observed. However, the median train travel time could be used as a proxy for the median in-vehicle time of passengers for each OD pair and time window. Here, the train travel time for any two stations is defined as the difference between the train departure time at one station and train arrival time at the other station. Subsequently, an approximation of the median out-of-vehicle time can be obtained as the difference between the median excess journey time and the median excess train travel time. Naturally, in the absence of assigning passengers to trains, it is not feasible to distinguish between the out-of-vehicle components at the beginning and end of the passenger journeys.

The comparison among the median excess journey time, median excess train travel time as a proxy of excess in-vehicle time, and estimated median out-of-vehicle time of passengers who entered the system during the peak of the peak period, from 18:15 to 18:45, is presented in
Figure 4-4. Each heat map in the figure shows the median value aggregated by OD pair and time window across weekdays (except Fridays) in March 2017.

![Heat Map Images](image)

(a) excess journey time  
(b) excess train travel time  
(c) excess out-of-vehicle time

Figure 4-4: Median Excess Journey Time Components by OD Pair  
Data source: Passenger journeys and train operations for 18:15 and 18:45 time period on weekdays (except Fridays) in March 2017

As we can see from the figure, even during the peak of the peak period, the median excess train travel times for most OD pairs were negative except for trips departed from Station 1 (the reason of these prolonged trips is examined in the next section). This is due in part to the removal of the four days with major incidents which was described in Section 4.3.2. However,
still it should be said that MTR’s rail operation in TWL is stable in general. In addition, by
definition in this thesis, the train travel times do not include the dwell times at the origin station,
which is often prolonged due to passenger surges or train traffic congestion on the line during the
peak period. On the other hand, running time generally contains some supplemental time to
recover the schedule. Consequently, excess train travel time can be negative, unless there is an
incident.

In contrast, the value of median excess out-of-vehicle times for almost all OD pairs were
positive and showed a similar pattern to the median excess journey time. In particular, the
median excess out-of-vehicle time for passenger journeys starting at Station 2 was consistently
high regardless of the destination station. These results imply that the most significant
contribution to excess journey time lies in the journey processes in Station 2 before boarding. It
is consistent with, and corroborates the empirical knowledge of, problems known to MTR staff.

4.6.1 In-Vehicle Issues

As already mentioned, the median excess in-vehicle times were negative for most OD pairs. This
means that train delays did not systematically increase the median passenger journey time, which
represents the typical passenger performance. However, the train travel times from Station 1 to
each destination station were consistently high compared to other train travel times, slightly
increasing median excess journey times starting at Station 1. This pattern implies that the train
running times between Station 1 and 2—i.e., the difference between a train departure time from
Station 1 and the train arrival time at Station 2—are consistently longer than the scheduled time.
To clarify the cause of these excess train travel times, a detailed analysis of AVL data follows.

For this analysis, trains running between Stations 1 and 2 from 18:15 to 18:45 on
weekdays (except Fridays) in March 2017, were considered. Figure 4-5 shows the distribution of
running time from Station 1 to Station 2 during the analysis time period. It is clear that trains operating over that segment were frequently late compared with the scheduled running time of 86 seconds. Indeed, over 80% of running times exceeded scheduled running time during the analysis time period. Those delays increased the in-vehicle time for passengers traveling over this segment.

![Train Running Time Distribution between Stations 1 and 2](image)

Figure 4-5: Train Running Time Distribution between Stations 1 and 2
Data source: 203 train trips departing from Station 1 during 18:15 and 18:45 on weekdays (except Fridays) in March 2017

In terms of the underlying cause of prolonged running time between Station 1 and 2, the contention at the junction with delayed trains entering Station 1 could be one possible explanation. For a more thorough investigation, detailed analysis considering the features of the rail signaling system and operations control system would be required.

4.6.2 Out-of-Vehicle Issues

Based on the results shown in Figure 4-4, an examination of out-of-vehicle time plays a key role in identifying the main cause of service deterioration during the analysis time period. As shown in Figure 4-4 (c), the approximation of the median excess out-of-vehicle times from each origin station were consistently high regardless of the destination station. These patterns imply that
passengers were forced to spend more time at their origin stations. In particular, those who traveled from Station 2 had significantly higher excess out-of-vehicle times, in the range of 212 to 314 seconds as aggregate performance. Since it is implausible that walking times of passengers increased by hundreds of seconds due to congestion on their walking paths, their waiting times almost certainly account for these increases.

The causes of excess waiting times can be irregular train headways or denied boardings. Figure 4-6 shows the distribution of observed headway at stations 2 and 3 during the analysis time period. Though there were a few times the train headway exceeded 200 seconds, train headways were generally in the range of 100 to 150 seconds at Station 2. In addition, the headway distributions at stations 2 and 3 were similar. This implies that headway irregularity was not the cause of the inferred excess waiting time. Hence, it can be deduced that denied boardings happened frequently enough to cause the excess waiting times at Station 2. Meanwhile, the median excess journey time for passenger starting at Station 3 were not significantly affected by presumably overcrowded trains coming from Station 2. A possible explanation for this is that since many passengers alight at Station 3, denied boardings are not likely to occur at the station and do not negatively affect passenger experience traveling from the station.
Furthermore, as can be seen from Figure 4-4 (c), only the median excess out-of-vehicle time originating at Station 2 shows a clear, upward trend across destination stations. That means, the farther a passenger traveled, the longer the passengers’ out-of-vehicle time at Station 2. A possible explanation for this is that passengers traveling only a short distance might be willing to get on a very crowded train because their in-vehicle times are short, and thus the time they spent at Station 2 was shorter than others. However, it seems unlikely that passengers who traveled farther chose to miss 3 trains rather than 2 trains hoping for a less crowded train. In addition, as the platform at Station 2 is heavily crowded during the peak period, passengers on the platform have almost no choice but to obey the first come first served queueing principal. Hence, there might be some different passenger behavior occurring at Station 2.

To analyze passenger behavior at Station 2, the distributions of journey times from Station 2 to each destination station were examined. A sample of the distributions is shown in Figure 4-7. It can be seen that each distribution is bi-modal, and the mass associated with the second peak with longer journey times increases with the length of journey. Moreover, the value
of the first peak is close to the standard journey time, and the difference between those two peaks are around 360 seconds which corresponds to three times the scheduled headway, 120 seconds, during the analysis time period. This value also corresponds to the extra time required if passengers travel “backwards”, in other words, they first travel to Station 1 and from there take the less crowded train in the northbound direction. Since a passenger who travels farther might prefer to have a seat, or at least, board a less crowded train, it could be a plausible explanation for why the farther passengers travel, the more they are likely to first travel “backwards” even if this increases their journey time.

![Graphs showing journey time distributions for different destinations from Station 2.](image)

**Figure 4-7: Comparison of Journey Time Distributions from Station 2**

Data source: Passenger journeys for 18:00 and 19:00 time period on weekdays (except Fridays) in March 2017

To roughly estimate what percentage of passengers traveling “backwards” from Station 2 across different destination stations, the fraction of passengers that experienced excess journey
times exceeding 360 seconds for each destination can be calculated as a proxy for those traveling “backward”. Figure 4-8 shows the results of estimation, focusing on passenger journeys starting at Station 2 during 18:00 to 19:00 on weekdays in March 2017. As can be seen from the line chart, there is a steep rise between Stations 4 and 5 from 0.12 to 0.22, and the fraction remains relatively stable between Stations 5 and 10 around 0.20. This result implies that passengers may prefer to avoid boarding overcrowded trains and travel “backwards” when they travel farther than Station 5, in other words, when in-vehicle times exceed about 400 seconds.

![Line chart showing fraction of passengers possibly traveling backwards from Station 2](image)

**Figure 4-8: Fraction of passengers possibly traveling "backwards" from Station 2**

It should be noted that this kind of excess journey time due to unusual passenger behavior is classified as excess out-of-vehicle time, even though they spend most of the excess time on trains. However, since those excess times arise from an out-of-vehicle issue (i.e., overcrowding), it might be reasonable to identify it as out-of-vehicle service deterioration. For a more precise and accurate causal inference, a probabilistic model for passenger-to-train assignment [38], a microscopic passenger flow analysis, or manual survey in the station should be conducted.
4.7 Before and After Analysis

This section presents a before and after analysis of performance with respect to the line extension in MTR’s rail network. As mentioned in Section 4.1, the South Island Line (SIL), which opened in December 2016, has an interchange with the TWL at Admiralty [39]. SIL is a 7-km extension with 4 new stations, serving to the southern district of Hong Kong Island. On average, there are over 110,000 weekday passengers on SIL. Despite the increasing demand for TWL as a result of this extension, there has not been any major update on TWL’s timetable for weekdays during the peak period. This is because MTR was already operating TWL very close to capacity before the extension. It should be noted that, analyses in the rest of this section as well as the analyses which have been presented earlier in this chapter only consider passengers traveling within TWL. In other words, performance for passengers coming from SIL is not directly measured; however, the performance for those passengers can be approximated by that for passengers traveling within TWL.

Figure 4-9 shows the degree of service deterioration in terms of the median excess train travel time and estimated median out-of-vehicle time in March 2016 and March 2017, before and after the SIL opening. As illustrated in the heat maps of median excess train travel time, there were only slight changes in train operation performance before and after the extension. However, estimated median out-of-vehicle time for most OD pairs increased, particularly for those traveling from Station 2, in the range of 10 to 66 seconds. These results clearly indicate that the increased passenger boardings at Station 2, coming from SIL, led to further congestion, resulting in an increased rate of denied boarding or more traveling “backwards”.

94
The effect of the extension can also be measured from a different perspective—i.e., the frequency of service deterioration—using the framework proposed in this thesis. Figure 4-10 shows the daily probability of service deterioration over 180 and 300 seconds. It can be clearly seen that probability of experiencing excess journey time over 180 seconds for passengers traveling from Station 2 increased regardless of the destination station. In fact, passengers traveling from Station 2 to Stations 5, 6, 7, 8, 9, and 10 almost always experienced that level of
service deterioration after the extension. Furthermore, the estimated probability of experiencing excess journey time over 300 seconds also increased for passengers traveling from Station 2 to Stations 5-10. This service deterioration could potentially drive passengers to choose other options if they are available. Though the degree of systematic service deterioration must become worse, the implication of the results should be carefully examined because they depend on the setting of the threshold value.

(a) Excess journey time over 180 seconds (left) and over 300 seconds (right) in March 2016

(b) Excess journey time over 180 seconds (left) and over 300 seconds (right) in March 2017

Figure 4-10: Comparison of Service Deterioration Probability before and after the Line Extension
4.8 Summary

In this chapter, systematic service deterioration in the PM peak period in Tsuen Wan Line was identified and analyzed by applying the framework proposed in this thesis. The key outcomes from the application are as follows:

- **When:** During the 18:00 to 19:00 time period, MTR passengers were likely to experience service deterioration, particularly, during 18:15 to 18:45, the peak of the peak period.

- **Where:** Passenger journeys starting at Station 2 were significantly delayed in the PM peak period.

- **How frequent:** On average, passengers who traveled from Station 2 to Stations 5 and 6 during the peak of the peak period experienced excess journey time over 300 seconds with around 10 to 20% probability. What is worse, those who traveled from Station 2 to Stations 7, 8, 9, and 10 during that period experienced that service deterioration with higher probability in the range of 30 to 60%.

- **Why:** It turned out that running times between stations 1 and 2 often increased. However, train delays did not seem to explicitly cause recurrent service deterioration in terms of passenger journey times as a whole. On the other hand, the out-of-vehicle time, particularly the waiting time for boarding at a platform taking into account the possibility of being denied boarding, has great influence on the journey time from Station 2. Furthermore, a certain fraction of passengers are likely to initially travel backwards to avoid waiting for several packed trains, and to obtain a more comfortable ride.

In addition, the effect of increased passenger demand, which arose from the opening of the SIL, on passenger transportation performance was shown in the before and after analysis.
The analysis clarified that passengers traveling from Station 2 in the PM peak period are likely to experience a greater service gap with higher probability after the extension.

This application has shown that, by applying the framework, rail agencies can identify the time and location of recurrent service deterioration without relying on costly manual surveys or facing data overload. Specifically, the heat map of excess journey times provides a good indication of when and where service deterioration occurs. Further decomposing the excess journey time into in-vehicle time and out-of-vehicle time, the pattern of excess time variation by type and across OD pairs suggests the proximate cause of the service deterioration. Based on the outcomes from the framework, agencies can judge the significance of systematic service deterioration and decide whether countermeasures are needed, and if so what type, or conduct additional analyses or surveys to further pinpoint the root cause of service deterioration.
Chapter 5

Conclusion

This chapter summarizes the research presented in this thesis and suggests future research directions. Section 5.1 presents a research overview and draws conclusions about the proposed problem identification framework and its application to MTR. Section 5.2 makes recommendations for the use of this framework and discussing possible remedies for MTR. Section 5.3 presents future research directions.

5.1 Research Summary and Findings

The research aimed to support rail agencies, operating near-capacity in urban areas, to systematically identify their recurrent service problems from the passengers’ perspective. To be more specific, the research tried to identify the time, location, frequency, and infer causes of service deterioration, including passengers’ out-of-vehicle experience with the system. The ultimate goal was to enhance agencies’ service improvement process. To this end, the research developed a framework for identification of systematic service deterioration in urban rail systems, shedding light on the gap between passengers’ expectation and what has actually been provided in each stage of their journey.

The median excess journey time relative to the standard journey time, which passengers would experience under ideal rail operations and free-flow conditions in stations, was used as a performance measure. The reasons that excess journey time is desirable as the performance measure for this framework are as follows: (1) excess journey time is one of the most representative measures of the gap that passengers experience and is closely related to many
aspects of inconvenience such as passengers’ on-time experience, overcrowding, denied boardings, and missed transfers; (2) excess journey time, a relative measure based on service standards, is comparable across different time periods and lines, and thus agencies can set priorities for service improvement; and (3) the standard journey time and excess journey time can be estimated using timetable, AVL and AFC data which are readily available for many agencies operating urban rail systems.

The framework developed in this thesis consists of three steps: a) the estimation of the standard journey time; b) the detection of transport services deterioration using the high median excess journey times metric; and c) the identification of the type and cause of service deterioration.

In the first step, the standard walking time for each OD pair is estimated using historical data from the off-peak period. Subsequently, the standard journey time for each OD pair and time period is estimated as the sum of scheduled in-vehicle time, scheduled transfer time, and standard walking time. Next, hotspots of service deterioration are detected by comparing observed passenger journey times and the standard journey time. Such hotspots are characterized by high median excess journey time. Lastly, the high excess journey times are decomposed into excess in-vehicle time and excess out-of-vehicle time, and the type and cause of the deterioration are deduced by pattern analysis using heat maps. Thereafter, the direction of further analysis to identify the cause of the deterioration is presented.

The applicability of the framework was demonstrated using data from one of MTR’s busiest lines, Tsuen Wan Line, focusing on passenger journeys without transfers. Analyzing the service gap during the PM peak period on workdays in March 2017, it turned out that passenger journeys starting at Admiralty station in the northbound direction during the 18:15 to 18:45 time
period were likely to experience higher delays. Moreover, the median excess journey times from Admiralty to different destination stations exceeded 300 seconds, MTR’s threshold for a train delay, with probability in the range of 10 to 60%. The pattern analysis of service deterioration showed that the median train travel times from Admiralty during this time period were not prolonged, whereas the estimated passenger out-of-vehicle times were greatly prolonged. Hence, it was concluded that service deterioration was related to passengers’ out-of-vehicle experiences, caused by denied boardings. However, it was also observed that the farther a passenger traveled from Admiralty, the higher the excess journey time was. This may indicate that some passengers may travel “backwards” toward the terminal station so they could board less crowded trains.

The other application, a before and after analysis, focused on the impact of the SIL extension. This analysis clearly showed that passenger experience was affected by increased passenger demand due to the SIL extension. The analysis indicated that passengers traveling from Admiralty in the PM peak period are likely to experience a greater service deterioration with higher probability after the extension.

In summary, the main contributions of this thesis are the development of a framework based on readily available data that enables agencies:

- to monitor their service performance and exhaustively detect service deterioration, including passengers’ out-of-vehicle experience, occurring in their daily operations
- to identify problems and their potential causes which otherwise may require expensive manual surveys
5.2 Recommendations

5.2.1 Use of the Framework

Performance measurement from the passengers’ perspective is imperative for rail agencies in order to improve their service quality (which is critical for attracting and retaining passenger demand in the face of growing competition from other transportation modes). In addition, since passengers are becoming more sensitive to the gap between what they expect and what has been provided, minimization of this gap should be one of the main goals of service improvement. The framework proposed in this thesis serves this purpose and provides guidelines for agencies.

The main targets of this framework are urban rail systems, serving dense areas with high-frequency, homogeneous train operations. The framework sheds light on passengers’ out-of-vehicle experience which could be subject to recurrent deterioration in rail systems with heavy ridership. The framework also assumes train operations with a single stopping pattern which is common for most urban rail systems.

The main use of the framework is for periodic monitoring and problem identification of daily performance in rail systems at the OD level. The identification of the causes of service deterioration (e.g., recurrent train delays or denied boardings) facilitates the development of both short-term and long-term strategies to improve performance, based on the degree of service deterioration and their underlying causes.

The framework can also support before and after analysis with regard to a variety of service changes. If agencies deploy short-term countermeasures, such as passenger-flow control and transportation demand management to deal with crowded stations and trains, the effects of these countermeasures can be systematically measured using the proposed framework. Given
that agencies conduct timetable updates, they can examine how the gap between service expectations and the experience of passengers is affected. Moreover, when it comes to extensions of the rail system, the framework can also serve to analyze how the extension affects the existing lines’ performance.

5.2.2 Recommendations for MTR

Based on the outcomes from the applications presented in Chapter 4, the following are recommendations to MTR.

**Reinforcement of Transportation Demand Management (TDM)**

As shown, the peak passenger demands during commuting periods cause substantial service deterioration. To mitigate the deterioration, the development of transportation demand management (TDM) strategies, which shift and reduce the peak demand, is one possible approach. MTR has implemented such strategies, in the part in the form of discounts for passengers traveling before the peak of the peak. The effectiveness of TDM experiments can be evaluated using the proposed framework. The results can also inform further refinements of the TDM strategy.

**Development of Efficient and Fair Capacity Utilization Strategy**

Out-of-vehicle experience for passengers traveling from Admiralty during the PM peak period is particularly affected by insufficient train capacity. To fully utilize train capacity, it is important to balance passenger loads across the train-set. To this end, waiting passenger distribution on the platform both, at Admiralty and the previous station, Central, should be more uniform. MTR staff could guide passengers to fully utilize all available space. Furthermore, operating strategies to provide more train capacity for passengers traveling from Admiralty would be desirable. In
fact, MTR has implemented such a strategy, operating a train leaving empty from Central to allow for more capacity for those traveling from Admiralty. These proactive strategies should be further developed, and again, the effectiveness of the strategies can be measured using the framework proposed in this thesis.

**Field Survey of Passengers’ Travel “Backwards” Behavior**

The results of the case study suggest that a certain fraction of passengers traveling from Admiralty during the PM peak period are likely to travel “backwards”, in order to board the less crowded train departing from the terminal station, Central. Without individual position tracking systems, this hypothesis can be validated only by field surveys. To be more specific, MTR can record how many passengers, alighting at the TWL’s platform in Central, remain on the platform and take the train departing toward Admiralty. It is useful to note that, since the Island Line also operates between Central and Admiralty, passengers traveling “backwards” from Admiralty may use either line. However, the passengers using the Island Line and eventually boarding trains on TWL at Central cannot be easily counted with this survey.

**5.3 Future Research**

This section presents future research directions to extend the framework and enhance the service improvement process of rail agencies.
Validation of Problem Identification with Transfer Journeys

Though in theory, the framework is applicable to passenger journeys with transfers, a validation using real rail system data is needed. Since many passengers make journeys in urban rail systems requiring transfers, it is important to validate problem identification with transfers and extend or refine the framework if necessary.

Inclusion of Passenger-to-Train Assignment

As mentioned in Chapter 3, the use of a passenger-to-train assignment model in the framework might enable agencies to identify the cause of service deterioration more precisely because passenger journey time could then be divided into in-vehicle time and out-of-vehicle time more rigorously. In addition, out-of-vehicle time can be segmented into access and waiting time at the origin station, transfer time, and egress time at the destination station. To successfully incorporate passenger-to-train assignment model with this framework, the model must be robust and applicable to all OD patterns in the system.

Extension to Heterogeneous Rail Services

Despite the fact that many urban rail systems provide homogeneous train operations with a single stopping pattern on a single line, some agencies (e.g., JR-EAST) operate frequent but heterogeneous services with multiple stopping patterns on multiple lines. To apply the framework to such rail systems, care should be taken in the definition of the standard journey time because the expected journey time of passengers varies depending on the passenger route-choices.
Extension to Rail Systems with Different Data Availability

The research assumes that agencies have “closed” AFC systems which retain both passengers’ entry and exit records. However, agencies which have a flat fare policy (instead of a distance-based or zonal fare policy) often do not have such AFC systems. It is a great challenge to make the framework applicable to such rail systems. Further research is needed to conduct systematic problem detection and identification in the case of open systems. Current work on developing OD matrices for open systems, using bus and rail transactions, could be a promising direction [40].
Bibliography


