TRANSFER COORDINATION MODEL AND REAL-TIME STRATEGY FOR INTER-MODAL TRANSIT SERVICES

by

Eui-Hwan Chung

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Civil Engineering
University of Toronto

© Copyright by Eui-Hwan Chung 2009
Transfer Coordination Model and Real-Time Strategy for Inter-Modal Transit Services

Eui-Hwan Chung
Doctor of Philosophy
Department of Civil Engineering
University of Toronto
2009

Abstract

In multi-modal transit networks with several intersecting lines and modes, travel through the network typically requires one or more transfers among transit lines and modes, and as such transfer time is a significant component of transit travel time from the perspective of passengers. Accordingly, efficient transfers are very important to increase the attractiveness and productivity of transit service. This study presents two approaches for the provision of efficient transfers: schedule coordination and real-time CP (Connection Protection) control.

The coordination of transit schedules can reduce transfer time significantly. This dissertation develops an optimization model for generating transit timetables that minimize transfer-related times. The model attempts to find an optimal timetable by shifting the existing timetable and/or adding holding time to the timetable to minimize delays associated with transfers from a feeder route to a receiving route. Analytical models are developed to estimate the waiting time of the transfer passengers, and also to determine the influence of the schedule modification on the waiting times of non-transfer passengers. The developed model is evaluated through a case study, and the results show that the model reduces effectively the total transfer and waiting times through the modification of the current schedule.
However, even though timetables among intersecting lines may be properly coordinated, an operational control method is necessary to maintain coordinated transfers, which may occasionally be disrupted due to unexpected delays of transit vehicles. A promising approach is to utilize real-time CP control. It involves holding a transit unit in order to wait for another transit unit that is planned to provide a coordinated transfer but has been delayed. This study also develops a CP model to apply a holding control to a receiving run in order to protect the scheduled connection. It incorporates the probabilistic nature of transit operations in formulating a cost function, and accordingly makes more robust decisions for control. The developed model is evaluated and compared with previous models to demonstrate its ability to improve transfer efficiency and reduce the waiting times of affected passengers.
# Table of Contents

## Chapter 1  Introduction
1. Introduction ...................................................................................................................... 1
2. Scope and Objectives ....................................................................................................... 1
3. Motivations ...................................................................................................................... 3
4. Organization of Dissertation ........................................................................................ 4

## Chapter 2  Literature Review
1. Introduction ...................................................................................................................... 5
2. Transfer Coordination ...................................................................................................... 5
   2.1 Timed Transfer Approach ......................................................................................... 5
   2.2 Transfer Optimization Approach .............................................................................. 8
3. Real-time Strategy for Transfer Coordination ................................................................. 11
4. Transit Unit and Passenger Arrival Processes ................................................................ 14
   4.1 Probability Distributions of Vehicle Arrival and Travel Times ............................... 14
   4.2 Passenger Boarding and Alighting Processes ......................................................... 17
   4.3 Passenger Arrival Process at Stop ........................................................................... 19
5. Prediction of Transit Arrival Time .................................................................................. 21
6. Summary ......................................................................................................................... 23

## Chapter 3  Transfer Optimization for Inter-Modal Transit Services
1. Introduction ...................................................................................................................... 25
2. Model Formulation .......................................................................................................... 28
   2.1 PM-Peak Case ........................................................................................................... 28
   2.2 AM-Peak Case .......................................................................................................... 33
3. Estimation of Transfer Cost ........................................................................................... 38
4. Downstream Model ........................................................................................................ 46
   4.1 Model Framework .................................................................................................... 46
   4.2 Lognormal Distribution ......................................................................................... 48
Chapter 4 Real-Time Coordination Strategy

1 Introduction.................................................................................................................. 121
2 Connection Protection (CP) Model: Delay of One Feeder Run ................................. 126
   2.1 Formulation of Model............................................................................................ 126
   2.2 Analysis of Model............................................................................................... 132
      2.2.1 Testbed........................................................................................................ 132
      2.2.2 Sensitivity Analysis ...................................................................................... 135
      2.2.3 Decision of Optimal Holding Time .............................................................. 142
3 Connection Protection (CP) Model: Delay of Two Feeder Runs .............................. 143
   3.1 Formulation of Model............................................................................................ 143
   3.2 Analysis of Model............................................................................................... 151
      3.2.1 Sensitivity Analysis ...................................................................................... 151
      3.2.2 Decision of Optimal Holding Time .............................................................. 159
4 Comparison with Previous Studies ............................................................................. 162
Chapter 5  Conclusions and Recommendations

1  Introduction......................................................................................................................175
2  Transfer Optimization for Inter-modal Transfer ..............................................................175
3  Real-Time Control Strategy .............................................................................................177
4  Major Contributions .........................................................................................................178
5  Recommendations for Future Studies ..............................................................................179
List of Tables

TABLE 2-1: Probability distributions for transit travel and arrival times in past studies ......17
TABLE 2-2: Recommended Boarding and Alighting Times (Transit Capacity and Quality of Service Manual)........................................................................................................................................................................19
TABLE 3-1: Input parameters of commuter transit route-31 (PM-peak)..............................74
TABLE 3-2: Input parameters of commuter transit route-34 (PM-peak)..............................75
TABLE 3-3: Input parameters of local transit route-25 (PM-peak).......................................75
TABLE 3-4: Input parameters of local transit route-52 (PM-peak).......................................76
TABLE 3-5: Input parameters of commuter transit route-31A (AM-peak)............................78
TABLE 3-6: Input parameters of commuter transit route-31B (AM-peak)............................78
TABLE 3-7: Input parameters of commuter transit route-34 (AM-peak)............................79
TABLE 3-8: Input parameters of local transit route-25 (AM-peak).....................................79
TABLE 3-9: Input parameters of local transit route-52 (AM-peak).....................................80
TABLE 3-10: Used values of $\nu_{L,i,j,k}$ ..............................................................................82
TABLE 3-11: Solution spaces according to optimization schemes (PM-peak)...............83
TABLE 3-12: Optimized results of local transit route-25 (PM peak).................................84
TABLE 3-13: Optimized results of local transit route-52 (PM peak).................................85
TABLE 3-14: Solution spaces according to optimization schemes (AM-peak)...............96
TABLE 3-15: Optimized results of local transit route-25 (AM peak).................................97
TABLE 3-16: Optimized results of local transit route-52 (AM peak).................................98
TABLE 3-17: Tested GA parameters .................................................................................104
TABLE 3-18: Averages of three minimized costs at the last iteration (Scheme-2)..........111
TABLE 3-19: Averages of three minimized costs at the last iteration (Scheme-3)........116
TABLE 3-20: Recommended GA parameters ....................................................................117
TABLE 4-1: Input parameters for downstream stops.........................................................135
TABLE 4-2: Comparison of CP models ..........................................................................162
TABLE 4-3: Comparison of holding decisions: with early dispatching vs. without early dispatching .................................................................169
List of Figures

FIGURE 3-1: Configuration of the transit system for transfer optimization ........................................... 25
FIGURE 3-2: Shifting schemes to modify existing schedule ........................................................................... 29
FIGURE 3-3: Examples of possible transfers during AM-peak period ......................................................... 34
FIGURE 3-4: Definition of domains for Equations 3-24 and 3-25 ............................................................. 39
FIGURE 3-5: Approximation procedure for n:1-transfer ............................................................................... 45
FIGURE 3-6: Shapes of Lognormal distributions with different degrees of skewness ................................. 50
FIGURE 3-7: Possible shapes of the probability density functions of departure times ............................... 53
FIGURE 3-8: Overall framework of downstream model ................................................................................. 56
FIGURE 3-9: Downstream model vs. simulation (LN / 3.8 / without holding) ............................................. 61
FIGURE 3-10: Downstream model vs. simulation (LN / 3.8 / with holding) .............................................. 62
FIGURE 3-11: Downstream model vs. simulation (LN / 4.0 / with holding) .............................................. 64
FIGURE 3-12: Downstream model vs. simulation (N / 3.8 / without holding) ........................................... 65
FIGURE 3-13: Downstream model vs. simulation (N / 3.8 / with holding) .............................................. 66
FIGURE 3-14: Downstream model vs. simulation (N / 4.0 / without holding) ......................................... 67
FIGURE 3-15: Typical procedure of genetic algorithms .................................................................................. 68
FIGURE 3-16: Existing schedules at transfer stop (PM-peak) ..................................................................... 72
FIGURE 3-17: Existing schedule of local transit routes at downstream stops .............................................. 74
FIGURE 3-18: Existing schedules at transfer stop (AM-Peak) ................................................................. 77
FIGURE 3-19: Breakdown of transfer passengers for AM-peak case ......................................................... 81
FIGURE 3-20: Optimized schedules (L25:PM-peak) ................................................................................. 85
FIGURE 3-21: Transfer costs according to optimization schemes (L25:PM-peak) ....................................... 87
FIGURE 3-22: Optimized schedules (L52:PM-peak) ................................................................................. 88
FIGURE 3-23: Transfer costs according to optimization schemes (L52:PM-peak) ....................................... 89
FIGURE 3-24: Analysis of a transfer from feeder routes to a receiving route .............................................. 91
FIGURE 3-25: Transfer costs by Scheme-4 (L25 and L52:PM-peak) ........................................................ 95
FIGURE 3-26: Optimized schedules (L25:AM-peak) ................................................................................. 99
FIGURE 3-27: Transfer cost according to optimization schemes (L25:AM-peak) ..................................... 100
FIGURE 3-28: Optimized schedules (L52:AM-peak) ............................................................................... 101
FIGURE 3-29: Transfer cost according to optimization schemes (L52:AM-peak) .................................... 102
FIGURE 3-30: Scheme-2 of local transit route-25 (Population Size=10)...............................107
FIGURE 3-31: Scheme-2 of local transit route-25 (Population Size=30)...............................108
FIGURE 3-32: Scheme-2 of local transit route-52 (Population Size=10)...............................109
FIGURE 3-33: Scheme-2 of local transit route-52 (Population Size=30)...............................110
FIGURE 3-34: Scheme-3 of local transit route-25 (Population Size=10)...............................112
FIGURE 3-35: Scheme-3 of local transit route-25 (Population Size=50)...............................113
FIGURE 3-36: Scheme-3 of local transit route-52 (Population Size=10)...............................114
FIGURE 3-37: Scheme-3 of local transit route-52 (Population Size=50)...............................115
FIGURE 3-38: Review of solutions found by GA...................................................................118
FIGURE 4-1: Broken coordination by unexpected delay of feeder run..................................121
FIGURE 4-2: Influence of CP control on passengers’ waiting times......................................122
FIGURE 4-3: Required input for real-time CP control ...........................................................125
FIGURE 4-4: Randomness on predicted arrival time of feeder run ........................................127
FIGURE 4-5: Testbed for applying the CP model...................................................................133
FIGURE 4-6: Relationship between travel time and distance of feeder to the transfer stop...133
FIGURE 4-7: Variation of cost function against different $E(\hat{A}_{T_i})$’s ($c_v = 0.1$)..............136
FIGURE 4-8: Cost function and its cost components..............................................................137
FIGURE 4-9: Variation of cost function against different $E(\hat{A}_{T_i})$’s ($c_v = 0.2$)..............139
FIGURE 4-10: Variation of cost functions against different $P_{l,R}^{Tr}$’s..................................140
FIGURE 4-11: Variation of cost functions against different $P_{R}^{Inv}$’s.................................141
FIGURE 4-12: Variation of cost functions against different demand levels of downstream stops.......................................................................................................................141
FIGURE 4-13: Variation of cost functions against different $SDT_{R+1}$’s...............................142
FIGURE 4-14: Rule to find the optimal holding time (delay of one feeder run).....................143
FIGURE 4-15: Definition of domains for Equations 4-16, 4-17 and 4-18...............................145
FIGURE 4-16: Variation of three cost functions against different $E(\hat{A}_{T_2})$’s....................153
FIGURE 4-17: Three cost functions and their cost components............................................154
FIGURE 4-18: Variation of three cost functions against different $P_{1,R}^{Tr}$ and $P_{2,R}^{Tr}$.............156
FIGURE 4-19: Variation of three cost functions against different $P_{R}^{Inv}$’s.........................157
List of Appendices

Appendix 1: Notation (Chapter 3) ................................................................. 190
Appendix 2: Notation (Chapter 4) ................................................................. 195
Chapter 1
Introduction

1 Introduction

Inter-modal transfer time is one of the most significant components of transit travel from the perspective of passengers. In service areas where transit demand is widespread geographically, it is not cost-effective to provide direct routes connecting all origin-destination pairs. Transit networks typically consist of several intersecting lines and modes, requiring one or more transfers among transit lines and modes in order for passengers to get to their final destinations. Obviously, transfers involve a certain amount of resistance, because they interrupt travel and because they require passengers to orient themselves, walk between vehicles on different lines, and wait for the next transit unit. Therefore, efficient transfers are very important to increase the attractiveness and productivity of transit service.

The provision of efficient transfers is based on a number of comprehensive considerations (Vuchic, 2005), including: 1) the functional design of transit routes and transfer points while balancing the suppliers’ cost (transit operating cost) and the passengers’ cost (e.g., the number of transfers and transfer times); 2) the provision of convenient transfer facilities so that passengers easily connect to the next trip without confusion; 3) coordination of schedules to reduce the transfer times of passengers; 4) operational controls to protect scheduled connections; and 5) information services for passengers to access and use the transit service efficiently. The first three issues are related to the service planning stage of a transit service, while the last two are related to the operational stage. This study focuses on the third and fourth issues.

2 Scope and Objectives

This study focuses on coordinating the inter-modal transfer between a commuter transit service and a local transit service during the PM-peak and AM-peak periods. Under the scenario of the PM-peak period, the commuter transit typically serves transit demand from the Central Business
District (CBD) to a suburban area using high-speed, high-capacity modes such as trains and express buses. The local transit (typically regular buses) then provides transfer passengers with access to their final destination. In the AM-peak period, the local transit collects local residents and provides them with access to a transfer point to board commuter transit which carries them to the CBD area.

There are two alternative approaches to the coordination of transit route schedules, namely timed transfer approach or transfer optimization approach (Bookbinder and Desilets, 1992). The timed transfer approach is a comprehensive strategy to develop a transit network in which transit units arrive simultaneously at transfer stops to offer coordinated transfers in all directions (Systan, 1983; Pennsylvania University, 1981). In contrast, the transfer optimization approach does not require that transit units simultaneously meet at a transfer point. This approach schedules the departure times at a transfer point so as to minimize some objective functions that capture the overall disutility of the transfers (Bookbinder and Desilets, 1992). Although the timed transfer approach (i.e., synchronizing the timetables of the two transit services) is a candidate strategy, it is not applicable to the case considered in this study for the following reasons: 1) the two types of transit services considered in our study (i.e. commuter transit and local transit) are typically operated by two independent transit agencies with institutional barriers to synchronized scheduling; 2) the rescheduling of existing commuter transit services at a particular station is a formidable task, because of the complicated effects that this rescheduling would have on other suburban areas served by the commuter transit; and 3) the commuter rail transit has fairly constrained schedules because such operations typically share the railway with other train systems. Therefore, the transfer optimization approach is more appropriate for the application considered in this study. As such, we attempt to minimize the expected transfer time through modifying only the existing schedule of the local transit service, while fixing the schedule of the commuter transit service.

Notwithstanding timetable synchronized among intersecting lines, transit delays are inevitable due to unexpected disruption of transit operations, traffic congestion, accidents and the like. As such, transit delays cause a coordinated transfer to “break” and consequently increase transfer times. If the transfer is to an infrequent transit route, the inconvenience of the missed connection will be quite substantial. Therefore, an operational control is required to maintain coordinated transfers against unexpected delays of transit vehicles. A promising approach is the Connection
Protection (CP) control based on a holding strategy. The CP control involves holding a transit unit to wait for another transit unit that has been designated to provide a coordinated transfer but has been delayed, based on the arrival time of the delayed unit predicted in real time. Real-time control attempts to improve the quality of transfer process, and this study formulates a model for connection protection involving decisions on optimal holding times.

Accordingly, the objectives of this study are outlined as follows:

- Develop a transfer optimization model for the revision of the existing schedule of a transit route.
- Develop a CP model to protect a scheduled connection against the unexpected delay of a transit trip.
- Develop a downstream model to estimate the influences of the transfer optimization and the CP control on the waiting times of other stop passengers (non-transfer passengers).

This study explicitly incorporates the probabilistic nature of transit operations at the transfer stop as well as downstream stops in formulating the models, and consequently makes more robust decisions.

3 Motivation

The Greater Toronto and Hamilton Area (GTHA) of Ontario, Canada, one of the fastest growing areas in North America, has suffered from the negative consequences of rapid growth, such as traffic congestion, environmental pollution and depleting natural resources. The growth plan outlined for this area, called The Growth Plan for the Greater Golden Horseshoe, is a framework for implementing the vision of the government of Ontario for building prosperous communities, primarily by encouraging various transit-supportive policies, as public transit has been recognized as one of key elements to achieving smart growth (Ministry of Public Infrastructure Renewal of Ontario, 2006).
As an example of the efficiency of transit, GO Transit, a commuter transit service in the GTHA, carries the equivalent of 48 highway lanes of traffic during peak hour. A recent survey in the area shows that 79 percent of transit users and 50 percent of automobile drivers would take transit more often if transit service were improved (Metrolinx, 2008). By making transit more convenient and efficient, the public would use it more, thereby decreasing congestion in the highway system.

In the GTHA, many transit riders have to make one or more transfers between transit routes on their way to their destination. The delay caused by transfers represents a major disadvantage of the transit mode. The waiting time faced while making a transfer is weighted by transit users more heavily than the same delay while at a stop or in a vehicle (Liu et al. 1998). Coordinated transfers can be implemented in order to reduce the inconvenience related to the transfer task and attract more passengers.

4 Organization of Dissertation

This study consists of five chapters. The first chapter introduces scope, objectives and motivation of this study. Chapter 2 reviews previous studies related to this research, focusing on the schedule coordination, the characteristics of transit operations along routes, and prediction models for transit arrival times. Chapter 3 formulates a transfer optimization model. Accounting for the direction of transfers, AM-peak and PM-peak transfers are modelled differently from each other. This chapter also describes a downstream model to estimate the influence of schedule modification on the waiting times of non-transfer passengers. The developed transfer optimization model is evaluated via a case study. Chapter 4 develops and analyzes a real-time CP control model to apply the holding control to a receiving run in order to protect the scheduled connection against the delay of a feeder trip. This chapter also reviews the differences between the CP model of this study and that of the previous studies in detail. Finally, conclusions and recommendations for further research are discussed in Chapter 5.
Chapter 2
Literature Review

1 Introduction

This chapter reviews previous studies related to this study. The next section covers the literature dealing with the schedule coordination problems of transit systems. Section 3 reviews the literature on real-time control strategies to protect schedule coordination against unexpected delays. Section 4 summarizes studies related to characteristics of transit operation along routes, and Section 5 discusses the prediction models for transit arrival times. Section 6 summarizes the reviews of this chapter.

2 Transfer Coordination

Various studies have tackled the problem of planning a coordinated transit system. The major aim of such studies is to design transit schedules that reduce the inconveniences that naturally arise with transferring among different routes.

2.1 Timed Transfer Approach

A timed transfer approach is a strategy to develop a transit network in which transit units arrive simultaneously at transfer stops to offer coordinated transfers in all directions (Systan Inc., 1983; Vuchic et al., 1983). The basic idea of the timed transfer system is the pulsing of headways so that transit units from different transit routes meet at the same time. Accordingly, for this type of coordination, the pulse headway is one of the critical decision variables. Using an appropriate pulse headway common to all routes, transit units from different routes can be synchronized so that they arrive and depart at the same time. Transfer slack time is inserted into schedules in order to account for the variability of transit vehicle arrivals. There have been several studies addressing the optimization of the pulse headway and the slack time for successful timed transfer systems.
Hall developed a model for scheduling transit arrival times at a terminal (Hall, 1985). Its purpose was to find an optimal slack time, which minimized expected passenger delay, with the assumption that vehicles were delayed according to an exponential distribution. It showed that coordinating arrivals with departures was important when the headway was large relative to the average vehicle delay. The study stressed the importance of evaluating passengers’ transfer delays.

Lee and Schonfeld formulated a transfer cost function to find an optimal slack time for transfers between one bus route and one rail line (Lee and Schonfeld, 1991). Three distributions were assumed for the arrivals of transit units: empirical discrete, Gumbel and Normal distributions. The results showed the optimal slack time depended on headway, arrival variance, transfer volume, passenger time values and vehicle operating costs. This study optimized the slack time with a pre-specified headway. However, they suggested that the slack time and headway needed to be optimized jointly.

Lee developed an optimization model for a timed-transfer transit system (Lee, 1993). Cost functions were formulated to assess the effectiveness of coordinated operations. Headways and slack times were jointly optimized for all coordinated routes through a developed headway clustering algorithm. This study analyzed the next four cases in optimizing the transit system: uncoordinated operation of multiple routes with independently optimized headways; fully coordinated operation of multiple routes with a single common headway; fully coordinated operation of multiple routes with integer-ratio headways; and partially coordinated operation of multiple routes where some routes were uncoordinated while others were coordinated. The timed-transfer system was highly recommended for transit routes having a long headway, but this study emphasized that the large arrival variation of transit units undermined the benefit of the coordination.

Shih employed a design methodology for transit route networks with coordinated operation (Shih, 1994). The focus of this study was to construct a set of bus routes and determine the associated service frequencies and vehicle sizes. It explicitly incorporated the concept of the timed transfer in the bus transit network design problem. The analysis of test results showed that the coordinated design decreased the total system cost.
Chien optimized a rail transit route and its associated feeder bus services (Chien, 1995). The developed model consisted of two sub-models: inter-modal transit network optimization and route coordination. In the first sub-model, network characteristics such as route and station locations and operating headway were determined, which minimized the supplier and user costs of the network. The second model was to add slack times to the schedules in order to improve transfer reliabilities and reduce total system costs.

Ting suggested a total cost function assuming deterministic and probabilistic vehicle arrivals for a timed transfer system (Ting, 1997). The function, assessing the effectiveness of coordinated operations, consisted of user and operator costs and was applicable to three different scenarios: uncoordinated, fully coordinated and partially coordinated. Two heuristic algorithms were developed to find slack times and headways that minimized the cost function. The results showed that coordinated operation become more beneficial as the headway increased. The timed transfer approach with a single common headway was desirable, when the headways of coordinated routes were similar. However, integer-ratio headways were preferable to coordinating a transit network whose routes had quite different headways. Later, Park employed a genetic algorithm to solve this optimization model (Park, 2005).

Chowdhury and Chien applied the transfer coordination approach to an inter-modal transit system with multiple transfer stations where feeder bus routes were connected with a rail line (Chowdhury and Chien, 2002). An objective cost function was formulated in terms of operator and user costs. A four-stage procedure was employed to find optimal headways and slack times. The evaluation showed that the coordination was highly effective among routes that had long headways, low standard deviations of vehicle arrival times and large transfer demands. The large slack time was necessary for a route with a large headway and a standard deviation of vehicle arrival time. However, when the standard deviation exceeded a certain level, it was not effective to insert the slack time and coordinate the routes.

Ngamchai and Lovell considered the timed transfer approach in designing a transit network (Ngamchai and Lovell, 2003). This study optimized bus transit route configurations and service frequencies, incorporating headway coordination at transfer points and giving higher priority to pairs of routes with high transfer demands when assigning the coordination. The model of this study was formulated in a deterministic way, and found an optimal solution through a genetic
algorithm. A test with a benchmark network showed that the headway coordination could decrease the total cost of the final solution by 10%. In addition, the best result without the coordination was virtually the same as the worst result with the coordination.

Even though the above optimization models maintained a narrow scope (such as the pulse headway, the slack time and the layout of transit routes), the timed transfer approach is a comprehensive strategy requiring many considerations to plan, implement and operate it. Becker and Spielberg suggested the following considerations in implementing a timed transfer network in Norfolk, Virginia (Becker and Spielberg, 1999): the decision of a basic pulse headway; location and design of a transfer centre; direction of transfer; route design; transit unit operation (such as interlining and express service); control of transfer fees; and incident management.

2.2 Transfer Optimization Approach

Unlike the timed transfer system, the transfer optimization approach does not require that transit units simultaneously meet at a transfer point. This approach schedules the departure times at a transfer point so as to minimize some objective function that evaluate the overall disutility of the transfers.

Rapp and Gehner developed a deterministic transfer optimization tool to minimize total transfer waiting times (Rapp and Gehner, 1976). They optimized transfer delays through an automated-iterative modification of terminal departure times. The research assumed that the transit network was in a steady state over the time period, the headway of each route was predetermined, and the demands for transfers were known. A heuristic search technique was used to find terminal offset times of the routes. The developed tool was applied to a transit system in Switzerland and the results showed reductions in total transfer delays by approximately 20% with no increase in operating costs. Andreasson introduced a computer-aided transit planning system including a synchronization module similar to that developed in the study by Rapp and Gehner (Andreasson, 1977).

Salzborn proposed a rule for scheduling a bus system consisting of an inter-town route connecting interchanges and feeder routes radiating from each interchange (Salzborn, 1980). Two different scheduling procedures were discussed for the inter-town route and the feeder
routes, respectively. The main aim of the study was to generate timetables that would minimize passenger delays and the number of buses of the feeder routes.

Keudel presented two computerized tools: one to design a line network (DIANA) and another to minimize transfer times in the network (FABIAN) (Keudel, 1988). The FABIAN system applied a traffic signal synchronization algorithm to the problem of optimization.

Klemet and Stemme were the first to propose a formal integer programming approach to defining the transfer optimization problem while minimizing the total transfer waiting times (Klemet and Stemme, 1988). They solved the problem using a heuristic method based on a route-graph. Voβ reformulated the formulation of Klemet and Stemme to a quadratic assignment problem (Voβ, 1992). Later, Daduna and Voβ proposed a solution approach for the problem (Daduna and Voβ, 1995): it set an initial feasible solution using a heuristic and improved the solution using Tabu search and simulated annealing.

Most of the studies reviewed above ignored the randomness of bus arrivals at a transfer stop, even though it is an obvious cause of failed transfers. The models simply assumed the travel time as deterministic, and did not consider the influence of randomness on transfer optimization. Bookbinder and Désilets developed a transfer optimization model that incorporated the effects of randomness (Bookbinder and Désilets, 1992). They explicitly considered the bus arrival time as a random variable and estimated the inconvenience of transferring passengers as a form of the expected waiting time or waiting penalty. Each timetable was defined by a given headway and one offset value defining the departure time of the first trip. Each timetable was optimized through simulation by changing the offset value, using the heuristic search technique by Rapp and Gehner to find the optimal solution. Although transit arrival time was assumed as a shifted truncated exponential distribution, this study provided strong theoretical background for the transfer optimization problem and emphasized the importance of considering the effects of randomness. They commented that optimizing connections with the deterministic assumption could even produce a solution that was worse than the initial one.

Knopper and Muller investigated the possibilities and limitations of coordinated transfer in transit (Knopper and Muller, 1995). Their analytical model was similar to that of Bookbinder and Désilets, but they did not optimize a specific timetable. This study showed transfer waiting time could be reduced significantly by adjusting the departure time of a connecting vehicle. It was
found that coordination was worthwhile when the arrival time standard deviation on the feeder line at the transfer point was within a certain level.

Chakroborty *et al.* formulated a mathematical programming model of a bus scheduling problem at one transfer station (Chakroborty *et al.*, 1995). The decision variables were the arrival times and departure times of transit units serving the station and the objective function was to minimize the transfer time of transferring passengers and the initial waiting time of passengers starting their trip at the station. The arrivals of transit units were assumed deterministic. The authors pointed out the limitation of traditional optimization techniques in solving the given problem, and searched for the optimal schedule through a genetic algorithm. Later, Chakroborty *et al.* applied this mathematical formulation to a transit network having multiple transfer stations and selected a genetic algorithm as a solution technique (Chakroborty *et al.*, 1998).

Ceder and Tal proposed a model to maximize transit synchronization (Ceder and Tal, 2001). They defined synchronization as the simultaneous arrival of transit units at a transfer stop. This study formulated the model as a mixed integer program, and provided a heuristic algorithm to solve this problem. However, this model ignored the randomness of transit arrival times and the number of transfer passengers. Due to these weaknesses, they recommended this model only as a guideline tool for transit schedulers.

Shrivastava and Dhingra coordinated the schedules of feeder bus routes along suburban railway stations through two distinctive stages (Shrivastava and Dhingra, 2002). The first stage was to develop the feeder bus routes to the railway stations, and the second stage was the determination of optimally coordinated schedules for the feeder routes. A heuristic algorithm, based on a genetic algorithm, generated the feeder routes. Thereafter, they coordinated the schedules of the feeder routes and of the trains at the railway stations. Their objective function was the minimization of transfer time between the two services and the vehicle operating costs of the buses. Two suburban stations in Mumbai, India, were taken as a case study, and optimized timetables for the buses were developed through a genetic algorithm. Their model was basically deterministic.

Cevallos and Zhao presented a method for optimization of transfer times in a bus transit system based on a genetic algorithm (Cevallos and Zhao, 2006). They attempted to find an optimum solution for transfer optimization by shifting current timetables. As input into the model, they
utilized existing schedules and ridership data at transfer points. To incorporate the randomness of bus arrivals, they also input a schedule deviation value defined as deviation from the given schedule. This value was estimated through a Monte-Carlo simulation with a lognormal distribution fitted by real data. They applied the developed model to Broward County Transit, Florida, and reduced total transfer time approximately 10% per day.

For the provision of efficient transfers, it is also important to design a reliable transit schedules. Generally, a slack time is inserted to the time points of a transit route in order to control the randomness of transit operations. Several studies have been address what amount of the slack time should be inserted and where the time points are located. These could be determined by minimizing a total cost, including the waiting times and delays of transit passengers and the operating costs of a transit agency (Wirasinghe, 1993; Wirasinghe and Liu, 1995B; Liu and Wirasinghe, 2001).

3 Real-time Strategy for Transfer Coordination

Although transit schedules of interesting lines may be coordinated with each other, the planned coordination can be disrupted by unexpected delays of transit units. The following studies have discussed the application of transit vehicle control to maintain the coordination and to improve transfer efficiency.

Abkowitz et al. evaluated the feasibility of timed transfer through a simulation model (Abkowitz et al., 1987). They simulated four scenarios, including: (a) “Unscheduled”: do nothing, (b) “Scheduled”: buses are scheduled to meet simultaneously at a focal point, but do not wait for each other, (c) “Holding”: buses are scheduled to meet at a focal point and the bus with the longer headway waits for the next arriving bus from the higher frequency route, and (d) “Double holding”: buses are scheduled to meet at a focal point, and whichever bus arrives first waits for the next arriving bus from the other route. The simulated results revealed that the planned timed transfer method could improve transfer times significantly, but route characteristics had a direct influence on the feasibility of timed transfer and the preferred transfer strategy.

Dessouky et al. noted that new technologies such as Automatic Vehicle Location (AVL) and Automatic Passenger Counter (APC) could contribute to addressing the transfer problem
(Dessouky et al., 1999). They assessed the benefits of the devices for holding strategies at timed transfer stations through a simulation. Their simulation model was built based on empirical data collected by the Los Angeles County/Metropolitan Transit Agency. The simulation results showed that vehicle holding, based on real-time vehicle arrival information, could significantly reduce waiting times for passengers on the connecting buses without greatly increasing the number of missed connections.

Chung and Shalaby emphasized that an operational control was necessary to maintain coordinated transfers that might be disrupted due to unexpected delays of transit vehicles, and proposed a strategy to protect a scheduled coordination between two transit routes (Chung and Shalaby, 2007). The developed strategy requires a receiver trip to be held in order to wait for a delayed feeder trip. The used cost function was based on the transfer passengers’ waiting time, the in-vehicle passengers’ waiting time, and the downstream passengers’ waiting time. The simulated results showed that the optimal holding time was a trade-off between those waiting time components: the proposed strategy was able to capture adequately those trade-offs and made the appropriate decisions according to the given delay of the feeder.

Ginkel and Schböel presented a deterministic mathematical model to minimize the number of missed transit connections and transit delays (Ginkel and Schböel, 2002). It updated the schedules of all related transit units in a given future period when one train was delayed.

While showing significant improvement on the transfer time through holding control, the above studies commonly ignored the randomness of arrival times of delayed transit trips; not only are the transit arrival and departure processes stochastic, but the expected arrival times by any prediction model are also random variable having some probability distribution. The next studies formulated a cost function accounting for the randomness of transit arrival times, even though it was predicted based on real-time AVL and APC systems.

Lee studied real-time dispatch decisions in a timed-transfer transit system (Lee, 1993). The decision variable of this study was the holding times for vehicles ready to be dispatched. The holding times were optimized by minimizing the cost function, which was formulated using the delay cost of holding a vehicle that was ready to be dispatched with the passengers already on board, and the missed connection cost of late incoming passengers transferring to it from a late vehicle. The arrival time of the late vehicle was assumed to follow a Normal distribution. The
authors emphasized that the arrival of late vehicles should be accurately predicted for successful holding control.

Ting also proposed a real-time dispatching control method in a coordinated transit network (Ting, 1997). This study used a Normal distribution for the arrival time of delayed transit trips, and built a cost function, similar to that of the study by Lee. However, it had additional cost terms by assuming that there existed another transfer point on the downstream stops of the vehicle to be held. The study searched the optimal holding time through a gradient search approach, called Simultaneous Perturbation Stochastic Approximation (SPSA).

Chowdhury and Chien suggested a model to dispatch vehicles on connecting routes at a transfer station (Chowdhury and Chien, 2001). The optimization cost function consisted of the sum of the vehicle’s holding cost, the connection delay cost, and the missed connection cost. The cost function re-evaluated periodically dispatch decisions based on the updated arrival distributions. The model showed a significant improvement in transfer efficiency and in the reduction of total costs. The authors noted that the fluctuation in the predicted arrival times and transfer demand might have a negative influence on their control strategy. However, they ignored the costs of passengers who waited for the held vehicle on the downstream stops.

Those probabilistic models would make more robust decisions relative to deterministic models. However, it should be noted that the above studies defined the holding time as the extension of the scheduled departure time: a receiving vehicle is held until the pre-determined holding time even if the late feeder vehicle arrives before the expiry of that time. Thus, this definition is highly improbable and unrealistic. Hall et al. proposed early dispatching, where the receiving trip was allowed to depart immediately upon the arrival of all the feeder trips (Hall et al., 2001). They developed two analytical models to determine the optimal holding time to minimize transfer time under a stochastic condition: one was without early dispatching (i.e., the receiving tip is held until the given holding time regardless of the arrivals of feeders), and another was with early dispatching. This study showed the optimal holding time of the second model was longer than that of the first model. They mentioned that there existed some potential for improvement by allowing the early dispatching.

The Utah Transit Authority (UTA) in Salt Lake City, Utah, is one of the few documented examples of real-world applications of real-time transfer control (Battelle Memorial Institute,
It operates a connection protection system for passengers transferring from light rail to bus transit. This system makes use of Global Positioning System (GPS) data to predict the arrival times of trains. When a train is late, it decides whether the bus should wait for the train by comparing the predicted time to the maximum holding time of the bus. The maximum holding time is exogenously defined according to frequency, the time of day, and the bus market type. The evaluation of the system showed improvements in the probability of successful connections.

4 Transit Unit and Passenger Arrival Processes

The transit operation along routes is characterized by a number of processes and patterns such as travel and arrival times of transit units and their underlying probability distributions, passenger arrival patterns and transit service time. These are essential factors in coordinating transit schedules and controlling a transit unit for protecting transfers. The literature on various line characteristics is reviewed and summarized below.

4.1 Probability Distributions of Vehicle Arrival and Travel Times

Turnquist suggested the Lognormal distribution for the bus arrival time at a stop, based on observed data from four bus services (Turnquist, 1978). This study discussed important characteristics of the bus arrival times at a stop: (1) The left tail of a bus arrival time distribution could be truncated at some point, because the bus arrival time could have the definite earliest time that was dictated by the distance between two stops and the speed limit in effect; (2) The right tail of the distribution would be relatively long in that there was a finite possibility of the bus being late or even cancelled; (3) Some initial delay at a stop brought longer delays to its downstream stops, since late arrival caused larger boarding volumes than expected. This study concluded that the Lognormal distribution was consistent with those characteristics of the bus arrival time.

Andersson et al. developed a simulation program as a tool to tackle undesirable situations such as transit delays and the over-saturation of transit units (Andersson et al., 1978; Andersson and
Scalia-Tomba, 1981). Based on a statistical analysis of an urban bus route in Stockholm, Sweden, they observed that the statistical distribution of link travel times should be skewed to the right rather than symmetric due to speed regulations and bus performance. They also recommended the Lognormal distribution for the link travel time of transit units.

Turnquist and Bowman explored the effect of network structure on transit service reliability through a simulation (Turnquist and Bowman, 1980). They suggested the Gamma distribution to simulate link travel times via analysis of bus travel times in Chicago, Illinois. They pointed out that it was important to reduce the variability of link travel times for a reliable transit service and the provision of efficient transfers also depended on a reliable transit operation.

Abkowitz and Engelstein analyzed factors influencing running times of transit routes (Abkowitz and Engelstein, 1983). An empirical study, based on data collected on transit routes in Cincinnati, Ohio, showed that the mean transit running times were strongly affected by trip distance, the number of boarding and alighting passengers, and signalized intersections. Initial running time deviations on a route propagated to the downstream stops.

Talley and Becker inspected distributions of late and early time intervals of transit units (Talley and Becker, 1987). The data of the time intervals were collected at transit stops in southeastern Virginia. They built two separate models for the early and late cases for each route, instead of constructing one distribution for all arrivals. The sample data of both cases were reasonably fitted with exponential distributions. Even though they did not present the shapes of overall arrivals, the resulting tables showed that the total counts and average adherences of the late arrival cases were larger than those of the early cases. That implied that the overall shapes could not be symmetric based on scheduled arrival times.

Guenthner and Hamat analyzed the data of bus arrivals sampled from a transit system in Milwaukee County, Wisconsin, in order to evaluate the on-time performance of the system (Guenthner and Hamat, 1988A). 792 bus arrivals were sampled on four selected routes. Unlike Talley and Becker, they did not divide the data into two groups. Their observations revealed that the distribution of the differences between the actual arrival time and the scheduled arrival time was with a long right tail and a short left tail. A statistical test concluded that the data would rather follow a Gamma distribution than a Lognormal distribution. However, they commented that more studies were needed to confirm that a Gamma distribution fitted bus arrival times more
adequately, and that a Lognormal distribution could be considered due to the similarities between the two distributions.

Seneviratne analyzed transit travel time along one route (Seneviratne, 1988). The travel time was decomposed into natural running time, acceleration and deceleration delays and dwell time. Based on the data of a transit system in Halifax, Nova Scotia, those three elements were fitted as Normal, Gamma, and Normal distributions, respectively.

Strathman and Hopper empirically studied the effect of various scheduling, route, driver and operating characteristics on schedule adherence (Strathman and Hopper, 1993). A total of 1552 bus arrival times were observed at time points encountered in 200 bus trips on 59 routes in Portland, Oregon. About 88% of the total fell with the on-time range of one minute early to five minutes late; 6.8% of the observations arrived early and 5.5% arrived late. They concluded that the bus arrival times were approximately distributed according to a Lognormal distribution, even though they did not statistically test the goodness-of-fit of the distribution to the collected data.

Hall et al. analyzed transit travel times of infrequent services using the data collected by the Los Angeles County/Metropolitan Transportation Agency (Hall et al., 1997). They found that a Normal distribution was better than other distributions (such as Lognormal and Gamma) in fitting the travel times. For long headway routes, the travel times were negatively correlated with lateness at stops. That is, the effect of initial lateness was not significantly propagated to the downstream stops but partially recovered due to built-in slack in its schedule.

Cevallos modeled the distribution of bus arrival times to simulate schedule deviations (Cevallos, 2006). This study tested four probability distributions (Lognormal, Gamma, Normal and Weibull) using data sampled at 12 stations. The Lognormal was selected as the one that best fit the data for the points through visual inspection as well as statistical tests.

In summary, past studies suggested different probability distributions for transit travel and arrival times, as seen in TABLE 2-1. Although there does not exist a unique rule to determine which specific distribution is best, positively skewed distributions, such as Lognormal and Gamma, have been more preferable than other ones.
### TABLE 2-1 Probability distributions for transit travel and arrival times in past studies

<table>
<thead>
<tr>
<th>Authors</th>
<th>Location of collected data</th>
<th>Variable</th>
<th>Probability distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turnquist (1978)</td>
<td>Not Specified</td>
<td>Arrival Time</td>
<td>Lognormal</td>
</tr>
<tr>
<td>Andersson et al. (1979)</td>
<td>Stockholm, Sweden</td>
<td>Travel Time</td>
<td>Lognormal</td>
</tr>
<tr>
<td>Turnquist and Bowman (1980)</td>
<td>Chicago, Illinois</td>
<td>Travel Time</td>
<td>Gamma</td>
</tr>
<tr>
<td>Guenthner and Hamat (1988)</td>
<td>Milwaukee County, Wisconsin</td>
<td>Arrival Time</td>
<td>Gamma</td>
</tr>
<tr>
<td>Seneviratne (1988)</td>
<td>Halifax, Nova Scotia</td>
<td>Travel Speed (= Travel Time)</td>
<td>Normal</td>
</tr>
<tr>
<td>Strathman and Hopper (1993)</td>
<td>Portland, Oregon</td>
<td>Arrival Time</td>
<td>Lognormal</td>
</tr>
<tr>
<td>Hall et al. (1997)</td>
<td>Los Angeles, California</td>
<td>Travel Time</td>
<td>Normal</td>
</tr>
<tr>
<td>Cevallos (2006)</td>
<td>Broward County, Florida</td>
<td>Arrival Time</td>
<td>Lognormal</td>
</tr>
</tbody>
</table>

### 4.2 Passenger Boarding and Alighting Processes

One of the initial studies about boarding and alighting time is a survey by Kulash, cited in the study by Koffman (Koffman, 1978). Kulash estimated boarding time and alighting time per each passenger as 4.3 sec., and 2 sec. based on bus route data collected in Cambridge, Massachusetts.

Kraft and Deutschman investigated passenger service time of bus transit in Montreal, San Diego and New Brunswick through a photographic study (Kraft and Deutschman, 1977). The distributions of the service times were fitted with Erlang functions. Average boarding times and alighting times were about 1.5 sec. and 2.0 sec. for two door buses, while the average values and their variance increased significantly if one door was shared for both alighting and boarding.
Guenthner and Sinha analyzed the distribution of passengers boarding and alighting at stops along transit lines in Milwaukee, Wisconsin, and Lafayette, Indiana (Guenthner and Sinha, 1983). The Poisson distribution appropriately explained the service times for routes with low ridership, but Negative binomial distribution was a better descriptor regardless of ridership levels.

Zografos and Levinson examined passenger service time for a no-fare bus system. Two sec. per boarding passengers were commonly accepted for un-crowded buses (Zografos and Levinson, 1986). However, they pointed out that boarding times increased linearly, as bus crowding levels and the number of boarding passengers increased.

Guenthner and Hamat investigated the effect of a complex fare structure on passenger dwell time, based on data collected from a transit system in Detroit, Michigan (Guenthner and Hamat, 1988B). They found that the different fare types and payment methods did not have a significant influence on dwell times. The boarding and alighting times were described as a function of the number of passengers and were lognormally distributed.

Adamski developed an analytical model to explain passenger service processes at a bus stop (Adamski, 1992). Accounting for the randomness of the numbers of boarding and alighting passengers and their boarding and alighting times, this study derived analytical formulas for probability distribution functions of transit dwell time. It also showed the probability density function of the boarding times and the alighting times could be approximated by Normal or Gamma distributions.

Dueker et al. analyzed a bus dwell time based on AVL and APC data of TriMet, the transit service for Portland, Oregon (Dueker et al., 2004). The results showed that each boarding and alighting passenger required the service times of 3.48 sec. and 1.70 sec., respectively. The overall dwell times of the morning peak period was lower than that of other periods, due to the higher ratio of regular riders during the morning peak. Also, low floor buses contributed to reducing dwell times. They pointed out some irregular events, such as the lift operation for a wheel chair, increased the dwell time significantly (62.07 sec./wheel chair).
The Transit Capacity and Quality of Service Manual by the TRB (Transportation Research Board) provide the following recommendations for boarding times according to payment type, and alighting times according to door type (Kittelson & Associates et al., 2003).

**TABLE 2-2 Recommended Boarding and Alighting Times (Transit Capacity and Quality of Service Manual)**

<table>
<thead>
<tr>
<th>Boarding Time (sec./pass)</th>
<th>Payment Type</th>
<th>Observed Range</th>
<th>Recommended Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-payment including no-fee</td>
<td>2.25-2.75</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>Single ticket or token</td>
<td>3.4-3.6</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td>Exact change</td>
<td>3.6-4.3</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>Swipe or dip card</td>
<td>4.2</td>
<td>4.2</td>
</tr>
<tr>
<td></td>
<td>Smart card</td>
<td>3.0-3.7</td>
<td>3.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Alighting Time (sec./pass)</th>
<th>Door Type</th>
<th>Observed Range</th>
<th>Recommended Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Front Door</td>
<td>2.6-3.7</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td>Rear Door</td>
<td>1.4-2.7</td>
<td>2.1</td>
</tr>
</tbody>
</table>

### 4.3 Passenger Arrival Process at Stop

The passenger arrival pattern at a stop is highly related to the characteristics of a provided transit service. Many previous studies suggested two groups of passengers: 1) non-aware passengers who do not refer to transit schedules, and 2) aware passengers who do refer to schedules to reduce their waiting time at a stop. The former group arrives at stops randomly, independent of the schedule, and thus their waiting times are characterized by the mean and variance of service headway. In contrast, the second group schedules arrival times according to transit timetables, and experience shorter waiting times at a stop. Passengers tend to refer to timetables to minimize the inconvenience of infrequent transit services. Schedule reliability of a transit service has an influence on this passenger’s behaviour as well: poor schedule adherence decreases the share of aware passengers. The next several studies discussed these properties through empirical studies.
O’Flaherty and ManGan examined the relationship between bus service frequency and average passenger waiting times under peak and off-peak conditions through a survey in Leeds and London (O’Flaherty and ManGan, 1970). They found that the actual average waiting time from the survey was less than the theoretical expected waiting time, assuming random passenger arrivals. This suggested that bus passengers did not arrive at random but tended to adjust their arrival to the arrival times of the required bus.

Bowman and Turnquist investigated the existence of aware passengers (Bowman and Turnquist, 1981). The properties of the aware passenger were empirically examined through the data collection of seven locations for the morning peak period over five to ten days. Virtually all the passengers were aware of the transit schedules in nearly all cases, even though the scheduled headways were from 5 to 20 min. These high shares resulted from the data collected during morning peak period: most trips were either work trips or other trips repeated daily in this period, and most passengers were expected to be aware of the available transit services. This study speculated that mid-day periods would have a different pattern. They also analyzed the effect of service reliability on passenger arrivals, and showed that reliable transit service results in a timed arrival pattern of passengers.

Csikos and Currie quantitatively explored the features of passenger arrivals through extensive data culled from an automated fare collection system for heavy rail services in Melbourne, Australia (Csikos and Currie, 2008). Their analysis showed that passengers were aware of provided timetables and arrived at stations earlier than scheduled departures by a fair safety margin. This awareness was more apparent during peak periods as opposed to inter-peak periods; on infrequent routes as opposed to frequent routes; and on weekdays as opposed to the weekend. It was also observed that unreliable service made passengers arrive more randomly, but the degree to which this randomness occurred was lower than the study of Bowman and Turnquist. They found that passengers choose arrival times based not only on published schedules, but also their experienced knowledge of vehicle arrivals.

Luethi et al. suggested probability distribution functions for passenger arrivals using data collected at 28 bus, tram and commuter rail stations around Zurich, Switzerland (Luethi et al., 2007). Uniform and Johnson SB distributions were chosen for timetable-independent and timetable-dependent passengers, respectively. The share of time-dependent passengers started to
increase at headway of five minutes in this study. Time-of-day and transit reliability influenced passenger arrivals in the ways described in the previous studies.

Hall et al. found an interesting property of the long headway route (Hall et al., 1997). For a frequent bus line, it is commonly known that if a bus falls behind schedule, it would carry additional passengers, forcing it to fall further behind schedule. However, they pointed out that, in a long headway service, the delay of a bus did not cause further delays at downstream stops because passengers likely arrived at each stop according to the bus schedule. Furthermore, Hall et al. observed that the long headway service, indeed, allowed the bus to recover its delay as it proceeded downstream when slack time was inserted into the schedule.

5 Prediction of Transit Arrival Time

The accurate prediction of a transit arrival time is an essential element to implement a real-time control strategy for protecting schedule coordination. There have been several studies on this topic.

Lin and Zeng developed four algorithms to predict the arrival times of buses based on GPS (Global Positioning System) data (Lin and Zeng, 1999). Whereas the simplest algorithm used only the GPS-based bus location data, the extent to which other information was used by the other algorithms varied. The algorithm showing the best performance utilized bus schedule information, the difference between scheduled and actual arrival time and waiting time at time-check stops, as well as the GPS data.

Dailey et al. presented an algorithm to predict transit vehicle arrival times up to one hour in advance (Dailey et al., 2001). This study developed a Kalman filtering framework using historical statistics, and utilized time series data consisting of time and location pairs that were transmitted from an AVL system. The filter continuously predicted the arrival time as new data came through the AVL system. The model relied on real-time location data and historical statistics of the remaining time to arrival. It is assumed that other variables possibly influencing the arrival time were implicitly included in the statistics. The algorithm was implemented as an application to disseminate the predicted arrival times of transit in Seattle, Washington.
Chen et al. developed a methodology for bus arrival time prediction using data collected by an APC system (Chen et al., 2004). The model consisted of two components: 1) an ANN (Artificial Neural Network) model to predict bus travel time between time points, and 2) a Kalman filter-based dynamic algorithm to adjust the predicted arrival time using bus location information. The ANN model was trained with four input variables, day-of-week, time-of-day, weather, and segment, producing a baseline estimate of the travel time. The dynamic algorithm then combined the most recent information on bus location with the baseline estimate to predict arrival times at downstream time points. The algorithm not only explicitly considered variables influencing the travel time, but also updated it using the real-time APC data.

Bin et al. presented Support Vector Machine (SVM), a new Neural Network algorithm, to predict the bus arrival time (Bin et al., 2006). Unlike the traditional ANN, SVM is not amenable to the overfitting problem, and it could be trained through a linear optimization process. This study predicted the arrival time based on the travel time through the current segment and the latest travel time of the next segment. The authors built separate models according to the time-of-day and weather conditions. The developed model was tested using off-line data of a transit route and exhibited advantages over an ANN-based model.

Shalaby and Farhan developed a bus travel time prediction model using AVL and APC dynamic data (Shalaby and Farhan, 2004). It was based on two Kalman filtering algorithms for separate predictions of running time and dwell time. It used the historical AVL data of the actual running times of the previous bus trips on the same day, and of the last three days, to predict the running time along downstream route segments of the current trip. Similarly, it estimated the dwell time at downstream stops based on the APC data of the passenger arrival rates of the previous bus trip on the same day and of the last three days. The separation of the two elements enhanced the model’s ability to capture the effect of lateness or earliness of bus arrivals on the bus dwell time.

Unlike the above-mentioned studies, some research tried to estimate arrival time by observing how the delay at upstream sections was propagated downstream. Lin and Bertini formulated a Markov chain model for bus arrival time prediction that explicitly captured this delay propagation (Lin and Bertini, 2004). This study pointed out that the delay at a stop was correlated with that at upstream stops.
Rajbhandari developed stochastic time series and delay propagation models for arrival time prediction (Rajbhandari, 2005). This research also adopted Markov models to predict the propagation of bus delays to downstream stops. An AutoRegressive (AR) model, using a time series modelling approach, captured the temporal variations of bus travel times. Each of the two models showed satisfactory performance for each modelling purpose. The combination of the two models reduced the relative error when predictions were made at time points with some distance from the origin time point.

6 Summary

As reviewed, schedule coordination could be seen as a very promising method to reducing the transfer times of transit users. It commonly resulted in improving the level of transit services and attracting more passengers to the services. In the timed-transfer approach, which synchronizes schedules, pulse headways and slack times are the main decision variables of interest. Instead of striving for strict synchronization, however, the transfer optimization approach schedules arrival and/or departure times of transit trips at transfer points, so that the trips arrive and depart at fairly close times, reducing the overall transfer cost. This approach has been utilized to improving an existing schedule by shifting it, or creating a new schedule with either a predetermined headway or variable headway within given bounds. Transfer time was usually considered as the transfer cost. Some studies emphasized the influence of transit arrival randomness on the transfer time, but many models still ignored it. Although the idea of transfer optimization seems very straightforward, it is a problem that is hardly solvable via traditional optimization techniques. Accordingly, most of the previous studies used heuristic searches as solution techniques, while genetic algorithms have been widely used more recently.

Protecting planned coordination against unexpected delays of transit units, referred to as Connection Protection (CP) control, also improved transfer efficiency. Previous models mainly focused on the waiting times of affected passengers, and some models included vehicle operating costs. This review shows the importance for the CP model to incorporate the current operating conditions as realistically as possible, since its benefits can vary with route characteristics and preferred transfer strategies. It is important to capture the operational state of the transit system through AVL and APC systems and predict the input values of the model, such as the arrival
time of transit units, as well. There have been several studies on predicting arrival times, using various approaches, such as Kalman filter, ANN, SVM and time-series models.

Due to the probabilistic nature of transit operations, arrival times of a transit unit at a stop are a random variable. Even though there is no agreement on which distribution is the best to fit arrival times, a skewed distribution appears to be more suitable for the arrival and travel times in that the bus travel time has a finite minimal time based on speed regulations and bus performance. The boarding and alighting times are elements of a transit operation, which are influenced by several factors, such as the types of payments, the occupancy of a transit unit, the features of a transit unit, the time of a day and so on. The arrival patterns of passengers at a stop are affected by the service characteristics of a transit route. Passengers likely refer to the timetables to reduce their waiting time for an infrequent transit service, but unreliable services decrease the proportion of aware passengers.

This study has focused on the transfer management problem through two approaches: the transfer optimization approach for the planning-stage, and the real-time CP control approach for the operation-stage. As reviewed earlier, the timed transfer approach can be considered to build a coordinated transit schedule. The institutional barrier between the two transit agencies, and the complexity of simultaneous rescheduling of commuter transit and local transit services, renders the timed transfer approach a poor candidate for inter-modal transfer coordination. Accordingly, the transfer optimization is more appropriate for the application considered in this study. Unlike previous studies that focused on revising schedules by shifting whole timetables, this study additionally considers various schedule modification schemes: the changing of existing headways and the additional dispatching policy that further reduce transfer-related costs. The real-time CP control is a backup strategy for the operation-stage in order to maintain coordinated transfers against unexpected delays. Like previous studies, this study makes a holding decision based on the related passengers’ waiting times. However, this study incorporates the early dispatching of a receiving trip, and estimates the downstream passengers’ cost through the downstream model, explaining the downstream operation of the local transit.
Chapter 3  
Transfer Optimization for Inter-Modal Transit Service

1 Introduction

This chapter presents a transfer optimization model for transfer coordination among inter-modal services, such as those shown schematically in FIGURE 3-1. Two optimization models will be presented for the PM-peak and AM-peak periods. At the transfer point of FIGURE 3-1, the commuter transit service typically serves the transit demand from the Central Business District (CBD) to a suburban area during the PM-peak. The local transit service then provides transfer passengers with access to their final destinations. In contrast, the flow of transfer passengers of the AM-peak period is opposite to that of the PM-peak case. Local transit collects the transit demand of the suburban area traveling to the CBD area, and a commuter transit service provides them with long-haul access to their final destinations or other transfer points.

![FIGURE 3-1 Configuration of the transit system for transfer optimization](image)

In order to formulate a cost function for the model and evaluate the coordinated operation, the following assumptions are made:
• The total demand of transfer passengers from the commuter transit route to the local transit route (PM-Peak period), or from the local transit route to the commuter transit route (AM-Peak period) does not change with revising the existing schedule.

• The total demand of non-transfer passengers of the local transit route also does not change with revising the existing schedule.

• The transfer passengers take the first available run on their destined route (called the primary receiving run), and they always wait for the next run (called the secondary receiving run) if they miss the first one (i.e. they do not take alternate means of transport). We also assume that the passengers never miss the secondary receiving run. Note that for the PM-Peak period, the feeder route and the receiving route represent the commuter transit service and the local transit service, respectively; and vice-versa for the AM-Peak period.

• The arrival times of transit runs at stops are assumed to be random variables that follow the Lognormal distribution.

• The existing schedule of each transit route provides transit users with the scheduled arrival and departure times of each transit run at the transfer point, and each route is reliably operated within the existing schedule. Generally, transit schedules provide departure times. Therefore, if scheduled arrival times are not specified, they can be replaced by mean arrival times collected by surveys. The transit users know the mean arrival times through their experienced knowledge of transit arrivals.

• Reverse commuters (who travel from the suburban area to the CBD area during the PM-peak, or from the CBD area to the suburban area during the AM-peak) and local transfer passengers are ignored.

• The schedule modification does not cause additional costs to the transit agency.

The first two assumptions imply that the demand levels are insensitive to the quality of transit services. In reality, it is expected that revisions to the existing schedule would have some
influence on the level of the service, which may cause some auto drivers to switch to transit, or transit users to switch to other routes/modes; however, these effects are not expected to be major ones and accordingly they are ignored in this research. Regarding the third assumption, some of the transfer passengers may not wait for the secondary connecting run, but instead try to find an alternative mode or route to continue their trip if they miss the primary receiving run. This assumption may be relaxed by introducing an additional model that estimates how they choose the alternative modes/routes and what the relevant costs are, but that is beyond the scope of this study. It is worth noting that the assumption that passengers would not miss the secondary connecting run may be invalid if the headway of a receiving route is too short. However, transfer coordination is not of much relevance to connecting routes with short headways.

The Lognormal distribution is assumed to describe the probabilistic behaviour of the arrival times of transit runs at the transfer point. Although there is no concrete evidence as to which distribution is the best for the arrival times, it has been pointed out that a positively skewed distribution, such as Gamma or Lognormal distributions, fit field data better than a symmetric distribution, because transit runs need some minimum travel times to get to the next stops (Turnquist, 1978; Andersson et al., 1979; Turnquist and Bowman, 1980; Guenthner and Hamat, 1988A; Strathman and Hopper, 1993; Cevallos, 2006). As such, the Lognormal distribution has been used by other researchers and is used in this research as well.

The fifth assumption specifies that the existing schedule of each route individually is built such that the transit service operates with reasonable schedule adherence, without significant delays. The problem of schedule deviation and its effect on the transfer process is dealt with in this thesis via real-time operational control, which is the focus of Chapter 4.

The proposed model considers only optimizing inter-modal transfers between commuter transit services and local transit services, as assumed. Since there exist the directions of transfers at the transfer point during the PM-peak and the AM-peak (the majority of passengers travel from the CBD to the suburban area for the PM-Peak period, while moving from the CBD to the suburban area for the AM-Peak period), this study ignores reverse commuters and local transfer passengers. It also excludes the costs to transit operators, since it builds new schedules by modifying the existing schedules without additional transit units. However, those costs can be
added to the objective function quite easily with some conversion factors to unify the difference in attributes between the system costs and the waiting times of passengers.

The next section formulates the transfer optimization model. Accounting for the direction of transfers, AM-peak and PM-peak transfers are modelled differently from each other. Section 3 defines a function to estimate the expected transfer time of transfer passengers, and Section 4 describes a “downstream” model that specifies the arrival and departure process of the local transit runs at the downstream stops in order to estimate the influence of modifying its schedule on the waiting times of non-transfer passengers. Section 5 explains genetic algorithms, which is the solution approach used for the optimized coordination of timetables. Section 6 evaluates the developed model via a case study. Finally, the conclusions are presented and discussed in Section 7.

2 Model Formulation

2.1 PM-Peak Case

The schedule coordination for the PM-Peak period focuses on optimizing inter-modal transfers from the commuter transit services (feeder routes) to one local transit service (receiving route). More specifically, it implements the transfer optimization by modifying the existing schedule of the local transit service, while keeping the commuter transit schedules unchanged.

This model modifies the existing timetable by shifting local transit runs using two types of shifts, namely Shift-α and Shift-β. As indicated in FIGURE 3-2, Shift-α moves the entire schedule of the local transit route by an offset value, hence no change in headways occur. Shift-β moves each set of transit runs run by the same transit unit independently, and therefore, headways among consecutive transit units change. For example, in FIGURE 3-2, the first, third, and fifth departures from the transfer stop are one set, since the same transit unit serves those runs. The second and fourth comprise another set. Two different offset values will be applied to the two sets respectively for Shift-β.
Schedule-based dispatching has been commonly employed as a dispatching policy for receiving routes in order to facilitate transfers. In effect, this policy requires a transit unit not to depart before its scheduled departure time, but allows it to depart after that time regardless of whether its feeder run has arrived. In this model, an additional dispatching policy for the receiving run is considered. It allows the transit unit to extend its departure beyond the scheduled departure time up to a given holding time so as to wait for the feeder run, but to be dispatched immediately if the transfer is successfully completed before the maximum holding time has expired. This dispatching rule has been separately implemented through a real-time holding control that is the focus of Chapter 4. It is to hold a transit unit to wait for another transit unit that is planned to provide a coordinated transfer but has been delayed using real-time information from AVL (Automatic Vehicle Location) and APC (Automatic Passenger Counter) systems. This chapter focuses on schedule coordination taking operational control into consideration, while the next chapter focuses on real-time operational control. In contrast to other transfer optimization models only focusing on producing the optimal schedules, this model not only optimizes the schedule but also tackles the holding control simultaneously at the scheduling stage. The optimal holding time will be determined in sync with the optimal schedule.

The modification of the local transit schedule and the added dispatching policy are obviously beneficial to the transfer passengers. However, these actions affect the waiting times of non-transfer passengers, namely passengers already onboard the transit units and passengers starting their trips at the transfer stop or at downstream stops. Therefore, this model takes into account the waiting time as well as the transfer time. The downstream operations of the receiving routes
are explicitly modelled to estimate the influences of the two-level shifting and the holding time on the waiting times of other stop (non-transfer) passengers.

With the above considerations, the total cost function of the PM-peak case, Equation 3-1, is the sum of the total transfer cost and the total stop waiting cost. Since this transfer process is probabilistic, the transfer cost—the expected transfer time—consists of the successful connection cost and the missed connection cost (Equation 3-2). The former represents the expected time spent by transfer passengers to transfer to their intended receiving run (the primary receiving run) successfully. As assumed, they always wait for the run following the primary run (the secondary receiving run) if they miss it. Hence, the latter corresponds to the expected time spent transferring to the secondary receiving run. The number of the transfer passengers, $P_{i,j,L}^{tr}$, is considered for each arrival of the commuter runs. Section 3 describes how the transfer cost is estimated in detail. The stop waiting cost, the expected waiting time of the non-transfer passengers, is estimated through the downstream model, presented in Section 4. It models the arrival and departure process of the local transit run along its downstream stops and estimates the numbers of the stop passengers and their waiting times at the stops. Note that the arrival of a local transit run at the last stop corresponds to returning to the transfer stop as in FIGURE 3-2, and it will trigger the next local transit run. For example, in FIGURE 3-2, the arrival time of the third, fourth and fifth local transit runs at the transfer point is not given, but is provided by the downstream model.

Objective Function:

Minimize $\sum_{i=1}^{I} \sum_{j=1}^{C} \sum_{n=1}^{M} \delta_{i,j,L,n} \times P_{i,j,L}^{tr} \times E(T_{i,j,L}^{tr})$

$+ \sum_{n=1}^{M} \sum_{k=0}^{K_{i}} P_{L,n,k}^{st} \times E(T_{L,n,k}^{st})$  \hspace{1cm} (EQ. 3-1)

Constraint Conditions:

$E(T_{i,j,L}^{tr}) = E(T_{i,j,L}^{trS}) + E(T_{i,j,L}^{trM})$  \hspace{1cm} $\forall i, j$  \hspace{1cm} (EQ. 3-2)
\[ SDT_{L,n,k} = SDT_{L,n,k}^O + \alpha_L + \beta_{L,n} \quad \forall n, k \]  
(EQ. 3-3)

\[ \alpha_L^{\text{min}} \leq \alpha_L \leq \alpha_L^{\text{max}} \]  
(EQ. 3-4)

\[ \beta_{L,n}^{\text{min}} \leq \beta_{L,n} \leq \beta_{L,n}^{\text{max}} \quad \forall n \]  
(EQ. 3-5)

\[ \beta_{L,n+k} = \beta_{L,n} \quad \forall n \]  
(EQ. 3-6)

\[ 0 \leq HT_{L,n} \leq HT_{L,n}^{\text{max}} \quad \forall n \]  
(EQ. 3-7)

\[ \delta_{i,j,L,n} = \begin{cases} 1 & \text{if } SDT_{L,n-1,0} < SAT_{i,j,0} \leq SDT_{L,n,0} \\ 0 & \text{Otherwise} \end{cases} \quad \forall i, j, n \]  
(EQ. 3-8)

\[ \sum_{n=1}^{M_L} \delta_{i,j,L,n} = 1 \quad \forall i, j \]  
(EQ. 3-9)

Where:

- \( I \) : Number of commuter transit routes
- \( i \) : Index of commuter transit routes (\( i = 1,2,3,\ldots, I \))
- \( C_i \) : Number of runs on commuter transit route \( i \)
- \( j \) : Index of runs on a commuter transit route (\( j = 1,2,3,\ldots,C_i \))
- \( L \) : Index of a local transit route (\( n(L) = 1 \))
- \( M_L \) : Number of runs on local transit route \( L \)
- \( n \) : Index of runs on a local transit route (\( n = 1,2,3,\ldots,M_L \))
- \( K_L \) : Number of downstream stops on local transit route \( L \)
- \( k \) : Index of stops on a local transit route (\( k = 0,1,2,\ldots,K_L \) : transfer point, if \( k = 0 \); downstream stops, if \( k \geq 1 \))
$B_L$ : Number of transit units serving local transit route $L$

$P_{i,j,L}^{Tr}$ : Number of transfer passengers from run $j$ on commuter transit route $i$ to local transit route $L$

$P_{L,n,k}^{St}$ : Number of stop passengers for run $n$ on local transit route $L$ at stop $k$

$T_{i,j,L}^{Tr}$ : Transfer cost of a transfer passenger from run $j$ on commuter transit route $i$ to local transit route $L$

$T_{i,j,L}^{TrS}$ : Successful connection cost of a transfer passenger from run $j$ on commuter transit route $i$ to local transit route $L$

$T_{i,j,L}^{TrM}$ : Missed connection cost of a transfer passenger from run $j$ on commuter transit route $i$ to local transit route $L$

$T_{L,n,k}^{St}$ : Waiting cost of a stop passenger who start trip at stop $k$ through run $n$ on local transit route $L$

$\alpha_L$ : Offset value (in minutes) to shift all runs on local transit route $L$, (integer value)

$\alpha_{L}^{\min}, \alpha_{L}^{\max}$ : Minimum and maximum of $\alpha_L$

$\beta_{L,n}$ : Offset value (in minutes) to shift run $n$ on local transit route $L$, (integer value)

$\beta_{L,n}^{\min}, \beta_{L,n}^{\max}$ : Minimum and maximum of $\beta_{L,n}$

$HT_{L,n}$ : Holding time (in minutes) for run $n$ on local transit route $L$, (integer value)

$HT_{L,n}^{\max}$ : Maximum of $HT_{L,n}$

$\delta_{i,j,L,n}$ : Primary connection flag (binary value): 1, if run $n$ on local transit route $L$ is the primary connected run of run $j$ on route $i$
$SDT_{L,n,k}^O$ : Original scheduled departure time of run $n$ on local transit route $L$ at stop $k$

$SDT_{L,n,k}$ : Shifted scheduled departure time of run $n$ on local transit route $L$ at stop $k$

$SAT_{i,j,0}$ : Scheduled arrival time of run $j$ on commuter transit route $i$ at transfer point

$E(X)$ : Expected value (mean) of $X$

Equations 3-3 to 3-6 are the constraints related to Shift-$\alpha$ and Shift-$\beta$, as depicted in FIGURE 3-2. Equation 3-7 constrains the holding time to a maximum boundary. Equation 3-8 sets the primary connection flag, $\delta_{i,j,L,n}$. The primary receiving run of a commuter transit run is the first departing run on its receiving route whose scheduled departure time is later than the scheduled arrival time of the commuter transit run. Equation 3-9 forces a feeder run to have one primary receiving run.

The decision variables of this problem are $\alpha_L$, $\beta_{L,n}$ and $HT_{L,n}$, and this study finds the solution using a genetic algorithm. As previously reviewed, heuristic searches (e.g., tabu search, simulated annealing and so on.) have been found to be an efficient solution approach for the transfer optimization problem. More recently, genetic algorithms have been widely used by many researchers (Chakroborty et al., 1995; Chakroborty et al., 1998; Shrivastave and Dhingra, 2002; Cevallos and Zhao, 2006).

2.2 AM-Peak Case

Similar to the PM-peak case, the transfers in the AM-peak case are optimized by shifting the existing schedule of the local transit service while fixing the schedule of the commuter transit service. However, the demand of the transfer passengers at the transfer point is defined in a different way from that of the previous case. Due to the relatively low frequency of commuter transit on most lines even during the AM-peak period, the suburban commuters of the AM-peak usually target a specific commuter transit run, and many plan their access run on local transit in
order to catch the targeted run. Therefore, the transfer demand is defined for each departure of the commuter runs. It is also assumed that the passengers choose a local transit run whose arrival time at the transfer point is earlier than the departure time of their desired commuter run. FIGURE 3-3 illustrates two transfer examples with this assumption.

![Figure 3-3: Examples of possible transfers during the AM-peak period](image)

**FIGURE 3-3 Examples of possible transfers during the AM-peak period**

FIGURE 3-3 (a) is a simple case showing that the passenger demand targeting Run-C1 and Run-C2 take Run-L4 and Run-L5, respectively. In the case shown in FIGURE 3-3 (b), the passenger demand targeting Run-C2 has to take Run-L4, but Run-C1 is also available for them. For this situation, it is assumed that such passengers board Run-C1 instead of Run-C2 (which corresponds to the third assumption of Section 1). These passengers have a hidden waiting cost which is the difference between the scheduled departure times of Run-C1 and Run-C2, as depicted in the figure. This cost represents the waiting time of a transit user arriving at the destination earlier than the targeted time due to the given transit schedule. The formulation of the AM-peak case is:
Objective Function:

Minimize \( \sum_{n=1}^{M_L} \sum_{i=1}^{I} \sum_{j=1}^{C_i} \delta_{L,n,i,j} \times P_{L,n,j}^{Tr} \times E(T_{L,n,j}^{Tr}) \)

\[ + \sum_{n=1}^{M_L} \sum_{k=0}^{K_n} P_{L,n,k}^{St} \times E(T_{L,n,k}^{St}) \]

\[ + \sum_{i=1}^{I} \sum_{j=1}^{C_i} P_{L,i,j}^{Tr} \times T_{L,i,j}^H \]  \( \text{(EQ. 3-10)} \)

Constraint Conditions:

\[ E(T_{L,n,i}^{Tr}) = E(T_{L,n,i}^{Tr}) + E(T_{L,n,i}^{TrM}) \quad \forall \ n, \ i \]  \( \text{(EQ. 3-11)} \)

\[ SAT_{L,n,k} = SAT_{L,n,k}^O + \alpha_{L,n} \quad \forall \ n, \ k \]  \( \text{(EQ. 3-12)} \)

\[ \alpha_{L,n}^{\min} \leq \alpha_{L,n} \leq \alpha_{L,n}^{\max} \]  \( \text{(EQ. 3-13)} \)

\[ \beta_{L,n}^{\min} \leq \beta_{L,n} \leq \beta_{L,n}^{\max} \quad \forall \ n \]  \( \text{(EQ. 3-14)} \)

\[ \beta_{L,n+B_L} = \beta_{L,n} \quad \forall \ n \]  \( \text{(EQ. 3-15)} \)

\[ \delta_{L,n,i,j} = \begin{cases} 1 & \text{if } SDT_{i,j-1,0} < SAT_{L,n,K_L} \leq SDT_{i,j,0} \text{, } \forall \ n, \ i, \ j \\ 0 & \text{Otherwise} \end{cases} \]

\( \text{(EQ. 3-16)} \)

\[ \sum_{j=1}^{C_i} \delta_{L,n,i,j} = 1, \quad \forall \ n, \ i \]  \( \text{(EQ. 3-17)} \)

\[ \eta_{L,n,i,j} = \begin{cases} 1 & \text{if } SAT_{L,n,K_L} \leq SDT_{i,j,0} < SAT_{L,n+1,K_L} \text{, } \forall \ n, \ i, \ j \\ 0 & \text{Otherwise} \end{cases} \]

\( \text{(EQ. 3-18)} \)

\[ \sum_{n=1}^{M_L} \eta_{L,n,i,j} = 1 \quad \forall \ i, \ j \]  \( \text{(EQ. 3-19)} \)
\[ P_{L,n,i}^{Tr} = \sum_{j=1}^{C_i} \eta_{L,n,i,j} \times P_{L,i,j}^{Tr*} \quad \forall \ n, \ i \]  

(EQ. 3-20)

\[ \varphi_{L,i,j}^* = SDT_{i,j,0} \quad \forall \ i, \ j, \]  

(EQ. 3-21)

\[ \varphi_{L,i,j} = \sum_{n=1}^{M_i} \left\{ \eta_{L,n,i,j} \times \sum_{j'=1}^{C_i} \left( \delta_{L,n,i,j',j} \times SDT_{i,j',0} \right) \right\} \quad \forall \ i, \ j, \]  

(EQ. 3-22)

\[ T_{L,i,j}^H = (\varphi_{L,i,j}^* - \varphi_{L,i,j}) \quad \forall \ i, \ j, \]  

(EQ. 3-23)

Where:

- \( P_{L,n,i}^{Tr} \): Number of transfer passengers who aim for transferring from local transit route \( L \) to run \( j \) on commuter transit route \( i \) (i.e., Real demand for run \( j \) on route \( i \) from route \( L \))

- \( P_{L,n,i}^{Tr} \): Number of transfer passengers from run \( n \) on local transit route \( L \) to commuter transit route \( i \)

- \( T_{L,n,i}^{Tr} \): Transfer cost of a transfer passenger from run \( n \) on local transit route \( L \) to commuter transit route \( i \)

- \( T_{L,n,i}^{TrS} \): Successful connection cost of a transfer passenger from run \( n \) on local transit route \( L \) to commuter transit route \( i \)

- \( T_{L,n,i}^{TrM} \): Missed connection cost of a transfer passenger from run \( n \) on local transit route \( L \) to commuter transit route \( i \)

- \( T_{L,i,j}^H \): Hidden waiting cost of transfer passengers who aim to transfer from local transit route \( L \) to run \( j \) on commuter transit route \( i \)

- \( \delta_{L,n,i,j} \): Primary connecting flag (binary value); 1, if run \( j \) on commuter transit route \( i \) is the primary receiver of feeder run \( n \) on local transit route \( L \)
\( \eta_{L,n,i,j} \): Local transit selection flag (binary value): 1, if transfer passengers destined to run \( j \) on commuter transit route \( i \) take run \( n \) on local transit route \( L \)

\( \varphi^*_i \): Scheduled departure time of the commuter transit run aimed for by the transfer passengers who transfer from local transit route \( L \) to run \( j \) on commuter transit route \( i \) (i.e., \( \varphi^*_i = SDT_{i,j,0} \))

\( \varphi_{L,i,j} \): Scheduled departure time of the commuter transit run actually taken by the transfer passengers who transfer from local transit route \( L \) to run \( j \) on commuter transit route \( i \) (i.e., \( \varphi_{L,i,j} \leq \varphi^*_i \))

\( SAT^O_{L,n,k} \): Original scheduled arrival time of run \( n \) on local transit route \( L \) at stop \( k \)

\( SAT_{L,n,k} \): Shifted scheduled arrival time of run \( n \) on local transit route \( L \) at stop \( k \)

\( SDT_{i,j,0} \): Scheduled departure time of run \( j \) on commuter transit route \( i \) at transfer point

The total cost function of the AM-peak case, Equation 3-10, has an extra cost, the hidden waiting cost. Equations 3-11 to 3-17 define the total transfer cost, the shifting of the existing schedule of the local transit route and the primary connection flag as in the PM-peak case. However, the additional dispatching rule, holding a commuter transit run, is not applied here.

The transfer demands are given in \( P_{L,n,i,j}^{Tr} \), the number of transfer passengers who aim to transfer from local transit route \( L \) to run \( j \) on commuter transit route \( i \): this is the real demand for run \( j \) on route \( i \) from route \( L \). \( \eta_{L,n,i,j} \), set by Equation 3-18, is a binary matrix, representing which run of route \( L \) the transfer passengers destined to run \( j \) on commuter route \( i \) take. The equation implies that they choose the latest run of route \( L \) that is scheduled to arrive earlier than run \( j \) on route \( i \). Equation 3-19 implies that the transfer passengers are assigned to one local transit run. Therefore, the number of transfer passengers from run \( n \) on local transit route \( L \) to commuter transit route \( i \), \( P_{L,n,i}^{Tr} \), is obtained from Equation 3-20.
Equations 3-21, 3-22 and 3-23 define the hidden waiting cost. Because the passengers transfer to the first available commuter transit run, some of them may experience the hidden waiting time (see FIGURE 3-3 (b)). $\Phi_{L,i,j}^*$ represents the scheduled departure time of the commuter transit run sought by the transfer passengers who transfer from local transit route $L$ to run $j$ on commuter transit route $i$. Hence, $\Phi_{L,i,j}^*$ is simply $SDT_{i,j,0}$ (Equation 3-21). $\Phi_{L,i,j}$ is the scheduled departure time of the commuter transit run actually taken by the transfer passengers (Equation 3-22). That is, they will take the commuter transit run they originally desired, if $\Phi_{L,i,j}^* = \Phi_{L,i,j}$, and they will take the earlier run than the run they were originally destined for (the hidden waiting cost occurs), if $\Phi_{L,i,j} \leq \Phi_{L,i,j}^*$. Accordingly, the hidden waiting cost is given as Equation 3-23.

### 3 Estimation of Transfer Cost

A function to estimate the expected transfer time of transfer passengers is defined in this section. Assume that passengers transfer from run $j$ on commuter route $i$ to run $n$ on local transit route $L$ at a transfer point, stop $k = 0$, according to the transfer scenario of the PM-peak case. Denote the arrival time of feeder run $j$ as $AT_{i,j,0}$, with $g_{AT_{i,j,0}}$ as its probability density function. $AT_{L,n,0}$ and $g_{AT_{L,n,0}}$ are those of receiving run $n$. Let the arrivals of the two runs be independent of one another. The scheduled departure time of run $n$ and its additional holding time to protect this transfer are $SDT_{L,n,0}$ and $HT_{L,n}$, respectively. According to the dispatching policy previously specified, the connecting run must stay at the stop until $SDT_{L,n,0}$, but it can be extended by $HT_{L,n}$ if run $j$ is late. That is, it will be maximally held by $SDT_{L,n,0} + HT_{L,n}$, but will be dispatched immediately if it meets run $j$ before that time. $\Delta_{L,n,0}^T$ is the total boarding time of the transfer passengers to run $n$. Similarly, $\Delta_{L,n,0}^S$ is that of the stop passengers. Those two are estimated by multiplying a predefined boarding time per passenger by the numbers of transfer passengers and stop passengers, respectively. The transfer time of the transfer passengers and the departure time of run $j$ are dependent on the variables discussed above, and are defined as:
FIGURE 3-4 Definition of domains for Equations 3-24 and 3-25

\[
T_{i,j,L}^{Tr} = \begin{cases} 
SDT_{L,n,0} - AT_{i,j,0} & \text{for } R^{Case1} \\
AT_{L,n,0} + (\Delta^{St}_{L,n,0} + \Delta^{Tr}_{L,n,0}) - AT_{i,j,0} & \text{for } R^{Case2} \\
\Delta^{Tr}_{L,n,0} & \text{for } R^{Case3} \\
E(DT_{L,n+1,0}) - AT_{i,j,0} & \text{for } R^{Case4} \\
E(DT_{L,n+1,0}) - AT_{i,j,0} & \text{for } R^{Case5} 
\end{cases} 
\]  
(EQ. 3-24)

\[
DT_{L,n,0} = \begin{cases} 
SDT_{L,n,0} & \text{for } R^{Case1} \\
AT_{L,n,0} + (\Delta^{St}_{L,n,0} + \Delta^{Tr}_{L,n,0}) & \text{for } R^{Case2} \\
AT_{i,j,0} + \Delta^{Tr}_{L,n,0} & \text{for } R^{Case3} \\
SDT_{L,n,0} + HT_{L,n} & \text{for } R^{Case4} \\
AT_{L,n,0} + \Delta^{St}_{L,n,0} & \text{for } R^{Case5} 
\end{cases} 
\]  
(EQ. 3-25)

Where:

\(AT_{i,j,0}\) : Arrival time of run \(j\) on commuter transit route \(i\) at transfer point
$AT_{L,n,0}$ : Arrival time of run $n$ on local transit route at transfer point

$DT_{L,n,0}$ : Departure time of run $n$ on local transit route at transfer point

$\Delta_{L,n,0}^{Tr}$ : Boarding time of transfer passengers to run $n$ at transfer point

$\Delta_{L,n,0}^{St}$ : Boarding time of stop passengers to run $n$ at transfer point

$R^{Case}$ : Domain of Case defined in FIGURE 3-4

As shown in FIGURE 3-4, five domains are defined according to the combination of the related variables. The X-axis and Y-axis represent the arrival times of the commuter and the local transit runs, respectively. In this formulation, it is assumed that a transit unit is an ordinary two-door bus that simultaneously handles boarding and alighting passengers and admits boarding through the front door. The alighting time is ignored in the formulation. The arrival time of the connected run is shifted by a determined walking time between the two routes.

Case 1 represents situations where the local transit run completes the transfer before the scheduled departure time, and departs at its given schedule. $SDT_{L,n,0} - (\Delta_{L,n,0}^{St} + \Delta_{L,n,0}^{Tr})$, the crossing point of the X-axis and Y-axis of FIGURE 3-4, represents the critical arrival time of the local transit run, the time after which it will have to depart at a time later than $SDT_{L,n,0}$. When the two runs come before the critical time, the local transit run will be dispatched at $SDT_{L,n,0}$ (region 1 of the figure). There is another chance, fortunately, at region 2, at which the local run can depart at $SDT_{L,n,0}$. When it comes before the critical time, it will complete the boarding of stop passengers by $SDT_{L,n,0} - \Delta_{L,n,0}^{Tr}$ at the latest. Departure at $SDT_{L,n,0}$ is possible, however, as long as the commuter transit run contacts it in advance of $SDT_{L,n,0} - \Delta_{L,n,0}^{Tr}$. The local transit run cannot depart at its scheduled time except in Case 1. The transfer time for this case is $SDT_{L,n,0} - AT_{i,j,0}$. Note that the transfer time represents the time a passenger spends in transferring from a connected run to a connecting run. This formulation does not decompose it into platform waiting time and in-vehicle waiting time.
Case 2 implies that the transfer will always be successful if the commuter transit run is ahead of the local transit run (regions 3 and 4). However, a transfer is also possible even if the commuter transit run arrives later than the local transit run, such as in cases represented by region 5, where the commuter transit run comes to a stop while the local transit run is processing the stop passengers (i.e. non-transfer passengers). In Case 2, the local transit run is immediately dispatched after handling all passengers, since the departure time is always behind the scheduled departure time. Consequently, the departure time and the transfer time will be

\[ AT_{L,n,0} + (\Delta_{L,n,0}^{S_i} + \Delta_{L,n,0}^{Tr}) \] and \[ AT_{L,n,0} + (\Delta_{L,n,0}^{S_i} + \Delta_{L,n,0}^{Tr}) - AT_{i,j,0} \], respectively.

In Case 3, the commuter transit run is later than the local transit run, but the transfer is saved by the additional holding time. The local transit run is completely ready to leave the transfer stop but waits for the commuter transit run until a maximum time of \( SDT_{L,n,0} + HT_{L,n} \). Note that it does not stay once the transfer is complete. Therefore, its departure time is \( AT_{i,j,0} + \Delta_{L,n,0}^{Tr} \), which is directly dependent on the commuter run’s arrival. The transfer time is \( \Delta_{L,n,0}^{Tr} \).

Case 4 and Case 5 represent the situation where the transfer passengers miss their intended local transit run. Because it is assumed they wait for the next run, run \( n+1 \) on route \( L \), the transfer time is \( E(DT_{L,n+1,0}) - AT_{i,j,0} \). The departure time of the local transit run is \( SDT_{L,n,0} + HT_{L,n} \) for Case 4. However, that of Case 5 is \( AT_{L,n,0} + \Delta_{L,n,0}^{S_i} \), since the local transit run is too late to depart even at \( SDT_{L,n,0} + HT_{L,n} \).

Using the above definitions, the expected transfer time for the successful transfer, and the missed transfer are written as Equations 3-26 and 3-27, respectively. In the same way, Equations 3-28 and 3-29 will be the expected departure time of the connecting run and its variance. The solutions of the equations can be found using numerical methods.

\[
E(T_{i,j,L}^{TrS}) = \int \int_{R^{2}\times R^{2}} [T_{i,j,L}^{Tr} \times g_{AT_{i,j,0}}(x) \times g_{AT_{L,n,0}}(y)]dxdy \quad (EQ. 3-26)
\]

\[
E(T_{i,j,L}^{TrM}) = \int \int_{R^{2}\times R^{2}} [T_{i,j,L}^{Tr} \times g_{AT_{i,j,0}}(x) \times g_{AT_{L,n,0}}(y)]dxdy \quad (EQ. 3-27)
\]
\[
E(DT_{L,n,0}) = \iiint_{R^{CASE^{S} \cup CASE^{M}}} [DT_{L,n,0} \times g_{AT_{i,j,0}}(x) \times g_{AT_{L,n,0}}(y)] \, dx \, dy \]  \hspace{1cm} (EQ. 3-28)
\]

\[
Var(DT_{L,n,0}) = \iiint_{R^{CASE^{S} \cup CASE^{M}}} [DT_{L,n,0}^2 \times g_{AT_{i,j,0}}(x) \times g_{AT_{L,n,0}}(y)] \, dx \, dy - \{E(DT_{L,n,0})\}^2 \]  \hspace{1cm} (EQ. 3-29)
\]

Where:

\( g_X(X) \) : Probability density function (Lognormal distribution) of random variable \( X \)

\( Var(X) \) : Variance of \( X \)

\( CASE^{S} \) : Set of successful transfer cases, \{Case1,Case2,Case3\}

\( CASE^{M} \) : Set of missed transfer cases, \{Case4,Case5\}

The expected number of transfer passengers to the primary and secondary receivers can be known using the following three equations. The expected missed passengers, Equation 3-32, are passed to the secondary receiver and are considered as stop passengers.

\[
\delta_{i,j,L,n}^{S} = \iiint_{R^{CASE^{S}}} [g_{AT_{i,j,0}}(x) \times g_{AT_{L,n,0}}(y)] \, dx \, dy \]  \hspace{1cm} (EQ. 3-30)
\]

\[
E(P_{i,j,L}^{TS}) = \delta_{i,j,L,n}^{S} \times P_{i,j,L}^{TR} \]  \hspace{1cm} (EQ. 3-31)
\]

\[
E(P_{i,j,L}^{TM}) = (1 - \delta_{i,j,L,n}^{S}) \times P_{i,j,L}^{TR} \]  \hspace{1cm} (EQ. 3-32)
\]

Where:

\( \delta_{i,j,L,n}^{S} \) : Probability that transfer from run \( j \) on commuter transit route \( i \) to run \( n \) on local transit route \( L \) is successful
\[ E(P_{i,j,L}^{TrS}) : \text{Expected number of transfer passengers from run } j \text{ on commuter transit route } i \text{ to the primary receiver on local transit route } L \]

\[ E(P_{i,j,L}^{TrM}) : \text{Expected number of transfer passengers from run } j \text{ on commuter transit route } i \text{ to the secondary receiver on local transit route } L \]

The expected departure time and the expected total number of boarding passengers on run \( n \) are passed to the downstream model explaining the arrival and departure processes of the local transit run along its downstream stops and estimating the numbers of stop passengers and their waiting times at the stops. Note that the probability density function of \( DT_{L,n,0} \) is unknown—only its mean and variance can be estimated by Equations 3-28 and 3-29. The downstream model estimates the operation of the receiver run (the arrival and departure times) throughout the downstream stops and the expected waiting times of the passengers (non-transfer passengers) who use the run at the downstream stops. The detailed procedure is explained in Section 4. The arrival at the last stop corresponds with returning to the transfer stop, and it will trigger the next receiver run as:

\[ AT_{L,n,K_L} = AT_{L,n+b_L,0} \quad \text{(EQ. 3-33)} \]

The proposed model is formulated assuming transfer from one feeder to one receiver, i.e., 1:1-transfer. Transfers from multiple feeders, n:1-transfer, results in a \((n+1)\)th order joint probability function, which is difficult to manage. As one reasonable alternative to such a complicated formulation of the n:1-transfer, this study individually applies the 1:1-transfer model to each connection with the following considerations:

- In estimating transfer time from one feeder, transfer passengers from other feeders are counted as stop passengers.
The expected departure time of the receiver is estimated in the same manner. However, since it must be a unique value, it is approximated as the latest one among the individual expected departure times.

The expected transfer times of the connections, whose corresponding expected departure time is earlier than the latest one, get increased by the gap between their expected departure time and the selected one.

The above approximation can underestimate the expected departure time of a receiver when feeders arrive at approximately the same time, because it is estimated through the individual application of the 1:1-transfer model. However, such a case occurs infrequently at a suburban transfer point where commuter transit services (feeders) have relatively long headways. The detailed procedure is presented in FIGURE 3-5.

The transfer model of this section is also applicable for the AM-peak case in the same way: the receiving run will be the run on commuter transit service, while the feeder will be that of the local transit service.
Given conditions are:

For receiver run \( n \) on route \( L \):

\[ g_{AT_{L,n}} \cdot \Delta_{L,n,0}^S \cdot SDT_{L,n,0} \text{ and } HT_{L,n} \text{ are given.} \]

For a set of feeder runs, \( \Theta = \{ \text{Run } j \text{ on Route } i | \delta_{i,j,L,n} = 1 \} \):

\[ g_{AT_{i,0}} \text{ and } P_{i,j,L}^T \text{ are given.} \]

Step-1) Compute initial successful transfer passengers, \( E'(P_{i,j,L}^{TS}) \), from each feeder run.

For all run \( j \) on route \( i \in \Theta \{ \)

\[ E'(P_{i,j,L}^{TS}) \leftarrow P_{i,j,L}^T \times \int_{-\infty}^{\text{MAX}[SDT_{L,n,0}+HT_{L,n}]} f_{AT_{i,0}}(x)dx \]

\[ \Delta_{L,n,0}^S \leftarrow \Delta \times \sum_{\text{All Runs } \Theta} E'(P_{i,j,L}^{TS}) + \Delta_{L,n,0}^S \]

Step-2) Update \( E'(P_{i,j,L}^{TS}) \) to \( E''(P_{i,j,L}^{TS}) \).

For all run \( j \) on route \( i \in \Theta \{ \)

As setting that \( \Delta_{L,n,0}^{TS} = \Delta \times E'(P_{i,j,L}^{TS}) \) and \( \Delta_{L,n,0}^{S} = \Delta_{L,n,0}^{S} - \Delta \times E'(P_{i,j,L}^{TS}) \),
Compute \( E''(P_{i,j,L}^{TS}) \) using Equation 3-29.

\[ \Delta_{L,n,0}^{S} \leftarrow \Delta \times \sum_{\text{All Runs } \Theta} E''(P_{i,j,L}^{TS}) + \Delta_{L,n,0}^{S} \]

Step-3) Compute \( E(T_{i,j,L}^{TS}) \), \( E(T_{i,j,L}^{TM}) \), \( E(P_{i,j,L}^{TS}) \) and \( E(P_{i,j,L}^{TM}) \) for each feeder run, and \( E_{i,j}^r(DT_{L,n,0}) \)'s of the receiving run against each feeder run.

For all run \( j \) on route \( i \in \Theta \{ \)

As setting that \( \Delta_{L,n,0}^{TS} = \Delta \times E''(P_{i,j,L}^{TS}) \) and \( \Delta_{L,n,0}^{S} = \Delta_{L,n,0}^{S} - \Delta \times E''(P_{i,j,L}^{TS}) \),
Compute \( E(T_{i,j,L}^{TS}) \) using Equation 3-25,
Compute \( E(T_{i,j,L}^{TM}) \) using Equation 3-26,
Compute \( E(P_{i,j,L}^{TS}) \) using Equation 3-29,
Compute \( E(P_{i,j,L}^{TM}) \) using Equation 3-30, and
Compute \( E_{i,j}^r(DT_{L,n,0}) \) using Equation 3-27.

\}

Step-4) Determine \( E(DT_{L,n,0}) \) taking the maximum of \( E_{i,j}^r(DT_{L,n,0}) \).

\[ E(DT_{L,n,0}) \leftarrow \max_{\text{All Runs } \Theta} \{ E_{i,j}^r(DT_{L,n,0}) \} \]

Step-5) Update \( E(T_{i,j,L}^{TS}) \) according to \( E(DT_{L,n,0}) \).

For all Run \( j \) on route \( i \in \Theta \{ \)

\[ E(T_{i,j,L}^{TS}) \leftarrow E(T_{i,j,L}^{TS}) + \{ E(DT_{L,n,0}) - E_{i,j}^r(DT_{L,n,0}) \} \]

\}

FIGURE 3-5 Approximation procedure for n:1-transfer
4 Downstream Model

As illustrated in FIGURE 3-2, a run on a local transit route leaves the transfer stop, serves the downstream stops, and returns to the stop again. The downstream model of this subsection specifies the arrival and departure process of the transit unit at the downstream stops in order to estimate the influence of modifying its schedule on the waiting times of the downstream stop passengers.

4.1 Model Framework

The arrival and departure process of a transit unit at stops along a route can be formulated in a deterministic way in the following two equations:

\[
AT_{L,n,k+1} = DT_{L,n,k} + TT_{L,n,k}
\]  
(EQ. 3-34)

\[
DT_{L,n,k+1} = \begin{cases} 
SDT_{L,n,k+1} & \text{if } AT_{L,n,k+1} \leq SDT_{L,n,k+1} - \Lambda_{L,n,k+1} \\
AT_{L,n,k+1} + \Lambda_{L,n,k+1} & \text{if } AT_{L,n,k+1} > SDT_{L,n,k+1} - \Lambda_{L,n,k+1}
\end{cases}
\]  
(EQ. 3-35)

Where:

- \(AT_{L,n,k}\) : Arrival time of run \(n\) on local transit route \(L\) at stop \(k\)
- \(DT_{L,n,k}\) : Departure time of run \(n\) on local transit route \(L\) at stop \(k\)
- \(TT_{L,n,k}\) : Travel time from Stop \(k\) to Stop \(k+1\) of run \(n\) on local transit route \(L\)
- \(SDT_{L,n,k}\) : Scheduled departure time of run \(n\) on local transit route \(L\) at stop \(k\)
- \(\Lambda_{L,n,k}\) : Dwell time of run \(n\) on local transit route \(L\) at stop \(k\)
The arrival time of run $n$ at the $(k+1)$th downstream stop is simply the sum of the departure time at the previous stop and the travel time from Stop $k$ to Stop $k+1$ ($TT_{L,n,k}$) as shown in Equation 3-34. The departure time at stop $k+1$ ($DT_{L,n,k+1}$) will be either the given scheduled departure time ($SDT_{L,n,k+1}$) or the sum of the arrival time and the dwell time at the stop ($\Lambda_{L,n,k+1}$) according to the schedule adherence of the transit run, as shown in Equation 3-35. For a route having a relatively long headway with schedule adherence control, no transit unit can leave a time point earlier than the scheduled departure time. $SDT_{L,n,k} - \Lambda_{L,n,k}$ is the latest arrival time for the run so as to depart according to its schedule. Accordingly, Equation 3-35 implies: if the run arrives earlier than $SDT_{L,n,k} - \Lambda_{L,n,k}$, it will depart as scheduled; if not, it will depart as soon as it completes boarding and alighting of passengers. This equation will also be valid at transit stops where schedule control is not applied by assigning negative infinity to the scheduled departure time ($SDT_{L,n,k} = -\infty$).

Once the initial departure time (i.e. departure time at a transfer point) is known and travel times, scheduled departure times, and dwell times at the downstream stops along a transit route are given, it is possible to estimate the arrival and departure times of the transit run on the route through applying these equations recursively.

In the real world, however, the arrival and departure times of a transit run are not deterministic variables but random variables. The above deterministic framework should be a probabilistic process. The downstream model of this study assumes that the travel times as well as the arrival and departure times are described by the Lognormal distribution. Wirasinghe and Liu developed a model for the transit arrival and departure process similar to the downstream model of this study, and employed Weibull distributions for their model (Wirasinghe and Liu, 1995A). Before moving to the details of the model, some properties of Lognormal distribution must first be examined.
4.2 Lognormal Distribution

The Lognormal distribution (also called Johnson $S_{L}$ distribution) is the probability density distribution of a random variable whose logarithm has a normal distribution. That is, if a random variable $A$ is normally distributed, $\exp(A)$ will be a Lognormal distribution. The probability density function (PDF) of the Lognormal distribution is:

$$g_{t}(t: \mu, \sigma, \theta) = \frac{1}{t\sigma \sqrt{2\pi}} \exp\left(-\frac{(\ln(t - \theta) - \mu)^2}{2\sigma^2}\right)$$

(EQ. 3-36)

Where:

- $g_{t}$: PDF of a Lognormal distribution ($t \sim LN$)
- $\mu$: Scale parameter of $g_{t}$
- $\sigma$: Shape parameter of $g_{t}$
- $\theta$: Location parameter of $g_{t}$

The Lognormal distribution with the three parameters, as in the above equation, is called three-parameter Lognormal distribution, while a Lognormal distribution with $\theta = 0$ is called standard Lognormal distribution or two-parameter Lognormal distribution. The first three raw moments of a Lognormal distribution are (Heyde, 1963):

$$E(t) = \theta + \exp(\mu + \frac{\sigma^2}{2})$$

(EQ. 3-37)

$$E(t^2) = \theta^2 + \theta \exp(\mu + \frac{\sigma^2}{2}) + \exp(2\mu + \frac{4}{2}\sigma^2)$$

(EQ. 3-38)
\[ E(t^3) = \theta^3 t + 3 \theta^2 t \exp(\mu t + \frac{1}{2} \sigma^2 t) + 3 \theta t \exp(2\mu + \frac{4}{2} \sigma^2 t) + \exp(3\mu + \frac{9}{2} \sigma^2 t) \]  

(EQ. 3-39)

Where:

\[ E(t) \] : Expected value of random variable \( t \)

Accordingly, the mean of the distribution is \( E(t) \) defined in Equation 3-37, and the variance and skewness are defined as:

\[ \text{Var}(t) = \{ \exp(\sigma^2) - 1 \} \exp(2\mu + \sigma^2 t) \]  

(EQ. 3-40)

\[ \text{Skw}(t) = \{ \exp(\sigma^2) + 2 \} \sqrt{\exp(\sigma^2) - 1} \]  

(EQ. 3-41)

Where:

\[ \text{Var}(t) \] : Variance of random variable \( t \)

\[ \text{Skw}(t) \] : Skewness value of random variable \( t \)

If the mean, variance and skewness of a Lognormal distribution are given (instead of the three parameters), it is possible to compute the three parameters conversely by solving Equations 3-37, 3-40, and 3-41 simultaneously. However, the solution of the simultaneous equations does not exist in a closed form. It must be found through a numerical search such as the Newton-Raphson method. Hill et al. suggest a numerical solution algorithm to find the solution (Hill et al., 1976).

The shape of the distribution is always skewed positively (i.e. skewness is always positive). FIGURE 3-6 (a) shows the shape of the distribution with different degrees of skewness. As
shown in FIGURE 3-6 (b), a Lognormal distribution is very close to a Normal distribution, when it has a very small skewness.

FIGURE 3-6 Shapes of Lognormal distributions with different degrees of skewness

The sum of multiple Lognormal distributions does not become a Lognormal distribution: the distribution of the sum is an unknown distribution skewed like a Lognormal distribution (Beaulieu et al., 1995). The computation of the sum is very critical in the downstream model, because the arrival time is the sum of the departure time and the travel time (Equation 3-34) whose distributions are assumed as Lognormal distribution. Several approximation methods have been proposed for the sum, the most famous and simplest of which is the Fenton-Wilkinson (FW) approximation method (Fenton, 1960).

If two Lognormal random variables, X and Y, are given, the FW method assumes the sum of the two, Z, as a random variable having a Lognormal distribution. It approximates Z through matching positive moments. Assuming that X and Y are independent of each other, the first three moments of Z will be:
\[ \begin{align*}
E(Z) &= E(X+Y) = E(X) + E(Y) \quad \text{(EQ. 3-42)} \\
E(Z^2) &= E(X^2) + 2E(X)E(Y) + E(Y^2) \quad \text{(EQ. 3-43)} \\
E(Z^3) &= E(X^3) + 3E(X^2)E(Y) + 3E(X)E(Y^2) + E(Y^3) \quad \text{(EQ. 3-44)}
\end{align*} \]

These three equations are called the first three moments matching, also known as the FW approximation method. From the equations, the variance and skewness of \( Z \) are:

\[ \begin{align*}
Var(Z) &= E(Z^2) - \{E(Z)\}^2 \quad \text{(EQ. 3-45)} \\
Skw(Z) &= \frac{E(Z^3)}{\{E(Z^2)\}^{3/2}} \quad \text{(EQ. 3-46)}
\end{align*} \]

Now, since \( E(Z), Var(Z), \) and \( Skw(Z) \) are available, the three parameters of \( g_t \) can be found. There are other methods which can be used to determine the sum, such as negative moment matching, cumulants matching, curve fitting, and so on (Szyszkowicz, 2007). Nevertheless, the FW method provides very efficient results, while at the same time being relatively simple to implement.

### 4.3 Model for Arrival and Departure Processes of Transit Run

Suppose that the PDF of the arrival time of run \( n \) on local transit route \( L \) at stop \( k, \ g_{AT_{L,n,k}} \), is a Lognormal distribution. If the scheduled departure time, \( SDT_{L,n,k} \) and the dwell time, \( \Lambda_{L,n,k} \), are given, the PDF of the departure time, \( g_{DT_{L,n,k}} \), can be written in the next piecewise formula:
\[ g_{DT_{L,n,k}} (DT_{L,n,k}) = \begin{cases} G_{AT_{L,n,k}} (DT_{L,n,k} - \Lambda_{L,n,k}) & \text{for } DT_{L,n,k} = SDT_{L,n,k} \\ g_{AT_{L,n,k}} (DT_{L,n,k} - \Lambda_{L,n,k}) & \text{for } DT_{L,n,k} > SDT_{L,n,k} \end{cases} \]  

(EQ. 3-47)

Where:

\[ g_{DT_{L,n,k}} : \text{Probability density function of } DT_{L,n,k} \]

\[ g_{AT_{L,n,k}} : \text{Probability density function of } AT_{L,n,k} \]

\[ G_{AT_{L,n,k}} : \text{Cumulative density function of } AT_{L,n,k} \]

Since run \( n \) cannot depart the stop before \( SDT_{L,n,k} \), the domain of \( g_{DT_{L,n,k}} \) is from \( SDT_{L,n,k} \) to \( \infty \).

The probability that the run departs at \( SDT_{L,n,k} \), \( g_{DT_{L,n,k}} (DT_{L,n,k} = SDT_{L,n,k}) \), is

\[ G_{AT_{L,n,k}} (DT_{L,n,k} - \Lambda_{L,n,k}) \), because it should be earlier than \( SDT_{L,n,k} - \Lambda_{L,n,k} \) in order to depart at \( SDT_{L,n,k} \). If \( AT_{L,n,k} \) is later than \( SDT_{L,n,k} - \Lambda_{L,n,k} \) (i.e., \( DT_{L,n,k} > SDT_{L,n,k} \)), \( g_{DT_{L,n,k}} \) will be \( g_{AT_{L,n,k}} \) shifted by \( \Lambda_{L,n,k} \). This formula implies that the PDF of \( DT_{L,n,k} \) has various shapes according to \( g_{AT_{L,n,k}} \), \( SDT_{L,n,k} \) and \( \Lambda_{L,n,k} \). For example, if the distribution of \( AT_{L,n,k} \) is located sufficiently earlier than \( SDT_{L,n,k} - \Lambda_{L,n,k} \), \( g_{DT_{L,n,k}} \) will be close to the deterministic value of \( SDT_{L,n,k} \) (because \( E (DT_{L,n,k}) \approx SDT_{L,n,k} \) and \( Var (DT_{L,n,k}) \approx 0 \)). In contrast, if it is very late, \( g_{DT_{L,n,k}} \) will be \( g_{AT_{L,n,k}} \) shifted by \( \Lambda_{L,n,k} \). FIGURE 3-7 illustrates the various possible shapes of \( g_{DT_{L,n,k}} \) : FIGURE 3-7 (a) and (d) exemplify the two extreme examples mentioned. In FIGURE 3-7 (b) and (c), \( AT_{L,n,k} \) is distributed around \( SDT_{L,n,k} - \Lambda_{L,n,k} \). Even though \( g_{DT_{L,n,k}} \) can vary with the schedule adherence of run \( n \) at stop \( k \), the figure shows that the approximation of \( g_{DT_{L,n,k}} \) to a positively skewed distribution is quite suitable: the left side of \( g_{DT_{L,n,k}} \) is truncated by \( SDT_{L,n,k} - \Lambda_{L,n,k} \) while the
right side mirrors $g_{AT_{L,n,k}}$ and has a long tail. Accordingly, this study approximates $g_{DT_{L,n,k}}$ as a Lognormal distribution.

Since $g_{AT_{L,n,k}}$, $SDT_{L,n,k}$ and $\Lambda_{L,n,k}$ are given, the first three moments of $DT_{L,n,k}$ will be computed as the following:

**FIGURE 3-7 Possible shapes of the probability density functions of departure times**
\[ E( DT_{L,n,k} ) = \int_{-\infty}^{SDT_{L,n,k} - \Lambda_{L,n,k}} (SDT_{L,n,k} \times g_{AT_{L,n,k}}(t)) dt + \int_{SDT_{L,n,k} - \Lambda_{L,n,k}}^{\infty} ((t + \Lambda_{L,n,k}) \times g_{AT_{L,n,k}}(t)) dt \]  
(EQ. 3-48)

\[ E( DT_{L,n,k}^2 ) = \int_{-\infty}^{SDT_{L,n,k} - \Lambda_{L,n,k}} (SDT_{L,n,k}^2 \times g_{AT_{L,n,k}}(t)) dt + \int_{SDT_{L,n,k} - \Lambda_{L,n,k}}^{\infty} ((t + \Lambda_{L,n,k})^2 \times g_{AT_{L,n,k}}(t)) dt \]  
(EQ. 3-49)

\[ E( DT_{L,n,k}^3 ) = \int_{-\infty}^{SDT_{L,n,k} - \Lambda_{L,n,k}} (SDT_{L,n,k}^3 \times g_{AT_{L,n,k}}(t)) dt + \int_{SDT_{L,n,k} - \Lambda_{L,n,k}}^{\infty} ((t + \Lambda_{L,n,k})^3 \times g_{AT_{L,n,k}}(t)) dt \]  
(EQ. 3-50)

The first moment will be the expected value of \( DT_{L,n,k} \), and the variance and skewness are computed based on the three moments:

\[ Var( DT_{L,n,k} ) = E( DT_{L,n,k}^2 ) - \{ E( DT_{L,n,k} ) \}^2 \]  
(EQ. 3-51)

\[ Skw( DT_{L,n,k} ) = E( DT_{L,n,k}^3 ) / [ E( DT_{L,n,k}^2 ) ]^{3/2} \]  
(EQ. 3-52)

From \( E( DT_{L,n,k} ) \), \( Var( DT_{L,n,k} ) \) and \( Skw( DT_{L,n,k} ) \), \( g_{DT_{L,n,k}} \) can be found. Because \( g_{DT_{L,n,k}} \) is now available, the arrival time of run \( n \) at the next stop, \( AT_{L,n,k+1} \), is the sum of the two Lognormal random variables, \( DT_{L,n,k} \) and \( TT_{L,n,k} \). The PDF of \( AT_{L,n,k+1} \) is approximated by a Lognormal
distribution through the FW method. Based on \( g_{AT_{L,n,k}} \), the above procedure is recursively applied to the next stop.

To apply the downstream model to the transfer optimization model of the previous section, \( g_{AT_{L,n,1}} \) (the PDF of the arrival time of run \( n \) at the first downstream stop), one of the inputs of the downstream model, should be provided to this model. Although the PDF of \( DT_{L,n,0} \) is an unknown distribution, \( E(DT_{L,n,0}) \) and \( Var(DT_{L,n,0}) \) can be derived (see Equations 3-28 and 3-29). Because of the holding control at the transfer point, however, \( DT_{L,n,0} \) does not have the shapes exemplified in FIGURE 3-7. As a simplification, \( g_{AT_{L,n,1}} \) is approximated by a Lognormal distribution, which has the mean of \( E(DT_{L,n,0})+E(TT_{L,n,0}) \), the variance of \( Var(DT_{L,n,0})+Var(TT_{L,n,0}) \), and the skewness of \( Skw(TT_{L,n,0}) \). (\( TT_{L,n,0} \) is the travel time from the transfer point to stop 1 being a Lognormal random variable.) That is, the skewness of \( DT_{L,n,0} \) is ignored in approximating \( g_{AT_{L,n,1}} \).

Even though it assumes that the travel time is a Lognormal random variable, this model is applicable to normally distributed travel times. It is possible to approximate a Normal distribution by a Lognormal distribution, because a Lognormal distribution with a small skewness is very close to a Normal distribution (FIGURE 3-6). If \( TT_{L,n,k} \) is given in a Normal distribution, accordingly, the next can be established:

\[
TT_{L,n,k} \sim N( E(TT_{L,n,k} ), Var(TT_{L,n,k} ) ) \]

\[
\approx TT_{L,n,k} \sim LN( E(TT_{L,n,k} ), Var(TT_{L,n,k} ), Skw(TT_{L,n,k})=0.001 ) \quad (EQ. 3-53)
\]

That is, the normally distributed travel times are considered to be Lognormal distributions with a small skewness, such as 0.001 in the model. However, the arrival and departure times are also described by Lognormal distributions, even though the travel times are normally distributed. The
departure times are skewed because of the scheduled departure times. FIGURE 3-8 summarizes the overall input and output of the model.

\[ \begin{align*}
E(DT_{L,n,0}) & \quad Var(DT_{L,n,0}) \\
TT_{L,n,0} & \quad SDT_{L,n,1} \quad \Lambda_{L,n,1} \\
 & \quad SDT_{L,n,2} \quad \Lambda_{L,n,2} \\
 & \quad SDT_{L,n,3} \quad \Lambda_{L,n,3} \\
\end{align*} \]

\( TT_{L,n,k}, \ (k = 0,1,2,\ldots, K-1) \sim \text{Lognormal (or Normal)} \)

\( DT_{L,n,k}, \ (k = 1,2,\ldots,K) \sim \text{Lognormal} \)

\( AT_{L,n,k}, \ (k = 1,2,\ldots,K) \sim \text{Lognormal} \) \{Unknown\ Approx. to Lognormal\}

FIGURE 3-8 Overall framework of downstream model

4.4 Estimation of Waiting Time of Stop Passengers

The waiting time of the downstream passengers, who wait for run \( n \) on a local transit route at the downstream stops, depends on the arrivals and departures of run \( n \), as well as the arrival patterns of passengers at those stops. For low-frequency services, it is known that two types of passengers exist, namely aware and unaware passengers. The aware passengers check the route timetable in order to reduce their waiting time at stops, whereas unaware passengers arrive randomly (O’Flaherty and ManGan, 1970; Bowman and Turnquist, 1981; Luethi \textit{et al.}, 2007; Csikos and Currie, 2008). This study explicitly considers the two types.

The number of unaware passengers is simply the product of a given constant arrival rate and headway between two consecutive transit units:

\[
P_{L,n,k}^U = \Lambda_{L,n,k} \times \{Max(E(AT_{L,n,k}),SDT_{L,n,k}) - Max(E(AT_{L,n-1,k}),SDT_{L,n-1,k})\} \quad (\text{EQ. 3-54})
\]
Where:

\[ P_{L,n,k}^U \] : Number of unaware stop passengers for run \( n \) on route \( L \) at stop \( k \)

\[ \lambda_{L,n,k}^U \] : Arrival rate of unaware stop passengers for run \( n \) on route \( L \) at stop \( k \)

Since a transit unit will not depart for the next stop earlier than the scheduled departure time, the number of passengers will vary according to the deviation of operations from the schedule. Therefore, Equation 3-54 indicates that if the expected arrival time of a transit unit is earlier than the scheduled departure time, it will serve passengers who come by the scheduled departure time; if the expected arrival time of a transit unit is later than its scheduled departure time, it will serve passengers who come by its expected arrival time. The passengers who arrive at the stop during the dwell time are ignored.

The expected waiting time of unaware passengers is determined by the equation developed by Osuna and Newell (Osuna and Newell, 1972):

\[
E(T_{L,n,k}^U) = \left[ 1/2 \times E(H) \times (1 + \text{Var}(H) / E(H)^2) \right]
\]

(EQ. 3-55)

Where:

\( E(T_{L,n,k}^U) \) : Expected waiting time of an unaware stop passenger (run \( n \) on route \( L \) at stop \( k \))

\( E(H) \) : Expected headway, \( E(DT_{L,n,k}) - E(DT_{L,n-1,k}) \)

\( \text{Var}(H) \) : Variance of headway, \( \text{Var}(DT_{L,n,k}) + \text{Var}(DT_{L,n-1,k}) \)

The arrival of the aware passengers is a much more complicated process and beyond the scope of this study. This study simplifies it by introducing the following assumptions: 1) the number of passengers are proportional to the scheduled headway; 2) they arrive at a stop according to a
given arrival rate; 3) they arrive earlier than the scheduled departure time of their intended transit unit by a given average time; and 4) they do not miss the intended transit unit. Accordingly, the number of aware passengers and their waiting time will be:

\[ P_{L,n,k}^A = \lambda_{L,n,k}^A \times (SDT_{L,n,k} - SDT_{L,n-1,k}) \]  
\[ E(T_{L,n,k}^A) = [E(DT_{L,n,k}) - (SDT_{L,n,k} - \tau_{L,n,k})] \]

Where:

- \( P_{L,n,k}^A \): Number of aware stop passengers for run \( n \) on route \( L \) at stop \( k \)
- \( E(T_{L,n,k}^A) \): Expected waiting time of an aware stop passenger (run \( n \) on route \( L \) at stop \( k \))
- \( \lambda_{L,n,k}^A \): Arrival rate of aware stop passengers for run \( n \) on route \( L \) at stop \( k \)
- \( \tau_{L,n,k} \): Average arrival time in advance of \( SDT_{L,n,k} \) of aware passengers for run \( n \) on route \( L \) at stop \( k \)

Equation 3-56 is derived from the first and second assumptions. Due to the last assumption, the expected waiting time, Equation 3-57, will be dependent only on the expected departure time of the transit unit, regardless of the distribution of their arrival. Consequently, the expected waiting time of the downstream passengers at stop \( k \) can be expressed as:

\[ P_{L,n,k}^{St} = P_{L,n,k}^U + P_{L,n,k}^A \]  
\[ E(T_{L,n,k}^{St}) = (P_{L,n,k}^U / P_{L,n,k}^{St}) \times E(T_{L,n,k}^U) + (P_{L,n,k}^A / P_{L,n,k}^{St}) \times E(T_{L,n,k}^A) \]

Where:
\[ E(T_{L,n,k}^{St}) \] : Expected waiting time of a stop passenger who starts trip at stop \( k \) through run \( n \) on local transit route \( L \)

Assuming the alighting rate of the in-vehicle passengers at each stop is available, the dwell time at each stop can be computed, which is then input to the model of arrival and departure processes. Given that the number of in-vehicle passengers when run \( n \) departs from the transfer point is known, it is possible to trace the number of passengers in run \( n \) at each downstream stop using the following equation:

\[
P_{L,n,k+1}^{InV} = (1 - \omega_{L,n,k}) \times P_{L,n,k}^{InV} + (P_{L,n,k}^{St})
\]  

(EQ. 3-60)

Where:

\( P_{L,n,k}^{InV} \) : Number of in-vehicle passengers on run \( n \) on route \( L \) at stop \( k \)

\( \omega_{L,n,k} \) : Alighting rate of in-vehicle passengers on run \( n \) on route \( L \) at stop \( k \)

Therefore, the dwell time is the maximum value of the boarding time and the alighting time (Equation 3-61), which are computed through Equations 3-62 and 3-63, respectively. The boarding time and the alighting time per passenger are referred to in the Transit Capacity and Quality of Service Manual (Kittelson & Associates et al., 2003). They recommend \( \Delta = 3.5 \text{sec.} \) and \( \Delta = 2.1 \text{sec.} \), assuming that passengers pay their fare with a single ticket and alight through the rear door of a transit unit.

\[
\Lambda_{L,n,k} = \max(\Delta_{L,n,k}^{St}, \Delta_{L,n,k}^{St})
\]  

(EQ. 3-61)

\[
\Delta_{L,n,k}^{St} = \Delta \times (P_{L,n,k}^{St})
\]  

(EQ. 3-62)
\[
\Delta_{L,n,k}^{St} = \Delta \times (\omega_{L,n,k} \times P_{L,n,k+1}^{lnV})
\]  
(EQ. 3-63)

Where:

\[\Delta_{L,n,k}^{St}\] : Boarding time of stop passengers to run \(n\) on route \(L\) at stop \(k\)

\[\Delta_{L,n,k}^{St}\] : Alighting time of run \(n\) on route \(L\) at stop \(k\)

\[\Lambda_{L,n,k}\] : Dwell time of run \(n\) on local transit route \(L\) at stop \(k\)

\[\Delta\] : Boarding time per one passenger (3.5 sec. \(\approx\) 0.0583 min.)

\[\hat{\Delta}\] : Alighting time per one passenger (2.1 sec. \(\approx\) 0.035 min.)

### 4.5 Evaluation of the Downstream Model

The performance of the downstream model is tested through a simulation. It is assumed that transit run \(n\) on route \(L\) is scheduled to depart at the 100th min. at a transfer stop \((SDT_{L,n,0} = 100)\) and is ready to depart at that time. There are 40 stops on the downstream of run \(n\); in order to observe the performance of the model, a relatively large number of the downstream stops is considered. The differences of the scheduled departure times between two stops are five min. (i.e., \(SDT_{L,n,1} = 105, SDT_{L,n,2} = 110, \ldots, SDT_{L,n,40} = 300\)) and the dwell time at each stop is given as one min. The travel times between two downstream stops are given as 3.8 min. Their standard deviation and skewness are 15 percent of the travel time and one, respectively (i.e., \(E(TT_{L,n,k}) = 3.8, Stdv(TT_{L,n,k}) = 3.8 \times 0.15, Skw(TT_{L,n,k}) = 1\)). That is, the route has 0.2 min. of a slack time at each stop. (Generally, this kind of slack time is inserted by a transit agency in order to stabilize the operation of a route against the variance of transit operation. i.e., \(SDT_{L,n,k} - SDT_{L,n,k-1} > E(TT_{L,n,k}) + \Lambda_{L,n,k}\)).

FIGURE 3-9 presents the model vs. the simulation when run \(n\) departs as scheduled without any holding. The number of simulation replications is 10,000. In FIGURE 3-10, there exists a feeder run (run \(i\) from route \(j\)) supposed to arrive at the transfer point according to the Lognormal
distribution of $E(\text{AT}_{i,j,0})=102$, $Stdv(\text{AT}_{i,j,0})=1.0$ and $Skw(\text{AT}_{i,j,0})=0.1$, and the receiving run is supposed to be held for two min. if the feeder is later than $SDT_{L,n,0}$. (a) and (b) of each figure plot the mean and the standard deviation of the departure time of run $n$ at each stop from the model against those from the simulation. (c) shows departure delays, $E(DT_{L,n,k}) - SDT_{L,n,k}$, and (d) plot the PDF and the CDF (cumulative density function) of the arrival time of run $n$ at the 40th stop.

Run $n$ departs as per the given schedule.

(The number of simulation replications = 10,000)

FIGURE 3-9 Downstream model vs. simulation (LN/3.8 without holding)
The figures show that the means and the standard deviation of the departure times, input to estimate the waiting times of the downstream passengers, correspond well with the simulated results for both cases. The PDF and the CDF of the arrival times are quite closely distributed with those of the simulation.

Also, the model reasonably explains the downstream operations of run \( n \). Even though the run departs according to the given schedule, the departure delays and the standard deviations of the departure times increase moderately due to the randomness of the travel time as it proceeds to the downstream stops. However, those do not increase continuously, but converge to a certain level.
because of the slack time, as shown in FIGURE 3-9 (b) and (c). When run $n$ is held, in contrast, the standard deviation significantly increases for the initial stops (FIGURE 3-10 (b)); because of the departure delay at the transfer point, the randomness of the travel times have direct effects on the departures times. As the run moves to the downstream, the delay is absorbed by the slack time. The standard deviation and the departure delay decrease and converge (FIGURE 3-10 (b) and (c)).

Zero slack time would undeniably result in the continuous divergence of the standard deviations and the departure delays. FIGURE 3-11 presents the results of zero slack time with the holding scenario (i.e., $E (TT_{L,n,k})=4.0, Stdv(TT_{L,n,k})=4.0 \times 0.15, Skw(TT_{L,n,k})=1$). As expected, the standard deviations and the departure delays rise continuously, as run $R$ proceeds to the downstream (FIGURE 3-11 (b) and (c)). Also, the PDF of $AT_{L,n,40}$ of FIGURE 3-11 (d) is more dispersed than that of FIGURE 3-10 (d). In terms of the accuracy of the model, some gaps are apparent on the standard deviations as shown in FIGURE 3-11 (b). Some errors are incurred and accumulated. The errors result from the two approximations in the model: the Lognormal approximation of the departure time and the Lognormal approximation of the sum of two Lognormal random variables. The results show that some underestimations of the standard deviations can occur for this diverging case. However, the model traces the pattern of the simulation very well, and the margin of error is not significant even with the 40 iterations of the arrival and departure processes. (The model underestimates the standard deviation by 0.09 at the 40th stop.)

The tests show that the model is efficient and provide a good approximation of the operation of the receiving run, through the downstream stops, in order to estimate the waiting times of the downstream passengers.
As commented earlier, this model is applicable when the travel times are given as Normal distributions. A Normal distribution can be approximated by a Lognormal distribution having a small skewness (i.e., 0.001). In the next three figures, the same analysis is performed, assuming that the travel times are Normal distributions. FIGURE 3-12, 3-13 and 3-14 correspond to FIGURE 3-9, 3-10 and 3-11, respectively. Note that the travel time given in the normally distributed is considered as a Lognormal distribution having the skewness of 0.001 in the model, while the travel times of the simulation are sampled from the Normal distribution. The observed patterns of the previous comparisons occur in these tests as well, and the model works well even
at the Normal distribution of travel times. That is, the developed model can explain the downstream operations of run \( n \) for both conditions: the travel times with Normal distributions and the travel times with Lognormal distributions. This flexibility would be beneficial in a real application of the model.

\[ T_{L,n,k} \sim N; E(T_{L,n,k}) = 3.8 \text{ and } Stdv(T_{L,n,k}) = 3.8 \times 0.15. \]

Run \( n \) departs as per the given schedule.

(The number of simulation replications = 10,000)

FIGURE 3-12 Downstream model vs. simulation (\( N / 3.8 / \) without holding)
(a) Mean of $DT_{L,n,k}$  
(b) Standard deviation of $DT_{L,n,k}$  

(c) Departure delay \( E(DT_{L,n,k}) - SDT_{L,n,k} \)  
(d) PDF and CDF of $AT_{L,n,40}$  

\[ TT_{L,n,k} \sim N; E(TT_{L,n,k}) = 3.8 \text{ and } Stdv(TT_{L,n,k}) = 3.8 \times 0.15. \]

Run \( n \) will be held for 2 min., if a feeder run (run \( i \) from route \( j \)) is late.  

(The number of simulation replications = 10,000)  

FIGURE 3-13 Downstream model vs. simulation ($N / 3.8 /$ with holding)
5 Genetic Algorithms

5.1 Overview of Genetic Algorithms

Genetic algorithms (GAs) are a heuristic search method motivated by the theory of evolutionary biology. Developed by Holland in 1975, GAs have since been widely used in optimization and search problems as an efficient way of finding exact or approximate solutions (Holland, 1975; Goldberg, 1989; Whitley, 1994). The algorithms are different from traditional optimization techniques according to the following (Goldberg, 1989): (a) GAs work with abstract
representations of the parameter set rather than the parameters themselves; (b) GAs simultaneously search an optimal solution at various points of the search space; (c) GAs use objective functions without requiring the mathematical features of the objective functions and constraints; and (d) GAs do not use deterministic rules, but probabilistic rules.

GAs are based on an iterative procedure: they move toward an optimal solution through iterations. FIGURE 3-15 shows a typical GA process. The algorithm starts with a population, which is a set of candidate solutions to a given problem. The candidate solutions are referred to as chromosomes, which are randomly generated from a search space (the feasible solution space to a problem) and are, in turn, composed of genes, their most basic element.

![FIGURE 3-15 Typical procedure of genetic algorithms](image)

Each chromosome has a fitness value, which is determined by objective functions and represents the suitability of each chromosome. In each generation, a new population is evolved from the previous one through genetic operations inspired by natural selection. Reproduction is a process...
to determine which chromosomes are selected for generating the new population. The chromosomes with more suitable fitness values have a higher probability of being selected for reproduction. Crossover combines genes from the selected chromosomes (parents), and creates new chromosomes (offsprings). This process mixes chromosomes from the parents, creating new chromosomes containing characteristics of the parents and turning them over to the offspring. Simultaneously, mutation randomly changes the genes of a chromosome. It takes place after crossover in order to keep the genetic diversity of the population. Mutation involves preventing GAs from being trapped into local minima, allowing chromosomes to occupy a different area of the solution space. This procedure is repeated until a termination condition is satisfied.

5.2 Encoding

The first issue in applying GAs is to find suitable encoding of a candidate solution to a given problem. Binary encoding is commonly used, because it is simple and has been used since the beginning. A chromosome is a binary string made up of binary values (0’s and 1’s). This encoding scheme is adopted for this study, because the decision variables – integer values – can be easily coded as binary strings. There are other possible methods and the selection of parameter encoding scheme depends on the given problem.

5.3 Reproduction

As explained earlier, chromosomes are selected from the current generation to be the parents of the next generation. The problem is how to select the chromosomes so that those with better fitness have more of a chance to survive than those with poor fitness. The following explains the reproduction methods commonly used, namely, the roulette wheel selection, the rank selection, and elitism.

In the roulette wheel selection, each chromosome is associated with a probability of being selected according to its fitness value. The probability is defined as the ratio of the fitness value of a chromosome to the sum of the fitness values of the chromosomes within the whole population. Therefore, a chromosome with higher fitness will be the most likely to be selected.
However, this method has potential problems: premature convergence and stagnation. If a chromosome in an early generation has a much larger fitness value than other chromosomes in the population, the genes of the super chromosome will spread over the next population, very quickly. Accordingly, crossover cannot create new chromosomes anymore (at that point, new chromosomes can be generated by mutation only). The lack of diversity in the population results in premature convergence; i.e., the population converges to local optima too early. Alternatively, stagnation can occur at the end of runs where all chromosomes have a relatively high and similar fitness value. Then, the probability that each chromosome is selected becomes almost uniform, and selective pressure does not exist.

In rank selection, all chromosomes are ranked according to their fitness values and the rank of each chromosome is regarded as its fitness value instead of the fitness value itself. This method is free from the problems of premature convergence and stagnation, more likely to occur in the roulette wheel selection (Whitley, 1989). However, this method can cause slower convergence, and requires reordering overhead.

When generating the next generation through crossover and mutation, there exists a high chance of losing a chromosome with the best fitness value. Elitism ensures that the best chromosome survives and is passed onto the next generation (Coley, 1999). As it preserves the best solution found, elitism guarantees the convergence of GAs and can improve the performance of GAs.

### 5.4 Crossover

Crossover is a genetic operator to create new offsprings by exchanging some genes of the parents. It allows the offspring to inherit characteristics of the parents. To this end, the one-point and two-point crossovers are commonly used.

In the one-point crossover, a single crossover point is randomly selected on chromosomes of the parents. The chromosomes are cut at a point, and the sub-chromosomes beyond that point are swapped between the parents.
The two-point crossover randomly cuts the chromosomes of the parents at two points. The subchromosomes between the two points are exchanged between the parents in order to generate new offsprings. The probability that the crossover occurs is called the crossover rate.

5.5 Mutation

The role of mutation is to maintain the diversity of chromosomes in a population. As in biological mutation, it randomly alters one or more genes on a chromosome. Using binary encoding, it flips genes randomly selected (i.e., 0 to 1, or 1 to 0). The probability that the mutation takes place is called the mutation rate.

5.6 Consideration in Applying GAs

GAs are practical in problems which are non-linear and discrete in nature. However, GAs do not guarantee that the found solution is the global optima, although they are able to search a solution space on a global level. It is also important to run the GAs with appropriate GA parameters, such as the population size, crossover rate and mutation rate. An unsuitable set of parameters get GAs trapped into local minima or converged more slowly.

6 Case Study

6.1 Testbed

The developed model is evaluated using schedule data from a local transit service in the City of Brampton, a suburban municipality located northwest of Toronto, Canada. A commuter transit service (GO-Transit) links the CBD of Toronto with Brampton. During the PM-peak period, commuters travel back to Brampton using GO-Transit and transfer to the local transit buses to access their final destinations. In the AM-peak period, the local transit service collects the local residents from a local stop and carries them to a transfer stop of the commuter transit service.
FIGURE 3-16 illustrates the schedules of some selected routes at Brampton City Centre, one of the inter-modal transfer stations between the two services in the city, from 3:00PM to 7:00PM. The commuter transit service from Toronto is comprised of Route-31 (referred to as G31) and Route-34 (referred to as G34), while Route-25 (referred to as L25) and Route-52 (referred to as L52) are the local transit routes serving the station. L25 provides a total of eight runs every 30 minutes, using two buses. One bus undertakes the first trip chain, consisting of runs 1, 3, 5, and 7, while the other undertakes the second trip chain, runs 2, 4, 6 and 8, as indicated in the figure. The two runs outside of the boundary—runs 0 and 9—are assumed to be independently operated from the runs within the boundary. L52 has a 15-minute headway and four buses serving four trip chains. FIGURE 3-17 shows the configuration of the downstream stops of the two local transit routes and the scheduled departure times at each stop of a single run. The transfer passengers require a walking time of one minute from the commuter transit platform to the local transit loading bay. Other input parameters, such as the number of transfer passengers and stop passengers and the parameters of the probability density functions, are assigned appropriate values without referring to historical statistics. However, if a filed survey is available, we can estimate new parameters from the data in that survey. TABLES 3-1, 3-2, 3-3 and 3-4 present those parameters of the model. Note that the expected arrival times of the commuter transit services correspond to their scheduled arrival times (i.e., $E(\text{AT}_{G31,j,0}) = \text{SAT}_{G31,j,0}$ and $E(\text{AT}_{G34,j,0}) = \text{SAT}_{G34,j,0}$).

(a) Commuter transit route-31 (G31): feeder route

FIGURE 3-16 Existing schedules at transfer stop (PM-peak)

(Continued on next page)
(b) Commuter transit route-34 (G34): feeder route

(c) Local transit route-25 (L25): receiving route

(d) Local transit route-52 (L52): receiving route

FIGURE 3-16 Existing schedules at transfer stop (PM-peak)
TABLE 3-1 Input parameters of commuter transit route-31 (PM-peak)

<table>
<thead>
<tr>
<th>Run No. (j)</th>
<th>Arrival Time ($AT_{G31,j,0}$)</th>
<th>Mean² ($SAT_{G31,j,0}$)</th>
<th>Standard deviation</th>
<th>Skewness</th>
<th>Number of Transfer Pass. to each local transit service ($P_{G31,j,L25}^{Tr}$, $P_{G31,j,L52}^{Tr}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (Bus)</td>
<td>3:35</td>
<td>2</td>
<td>0.1</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>1 (Rail)</td>
<td>3:57</td>
<td>1</td>
<td>0.1</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>2 (Bus)</td>
<td>4:05</td>
<td>2</td>
<td>0.1</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>3 (Bus)</td>
<td>4:40</td>
<td>2</td>
<td>0.1</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>4 (Rail)</td>
<td>4:57</td>
<td>1</td>
<td>0.1</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>5 (Bus)</td>
<td>5:15</td>
<td>2</td>
<td>0.1</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>6 (Rail)</td>
<td>5:27</td>
<td>1</td>
<td>0.1</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>7 (Bus)</td>
<td>5:45</td>
<td>2</td>
<td>0.1</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>8 (Rail)</td>
<td>5:57</td>
<td>1</td>
<td>0.1</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>9 (Bus)</td>
<td>6:15</td>
<td>2</td>
<td>0.1</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>10 (Rail)</td>
<td>6:27</td>
<td>1</td>
<td>0.1</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>11 (Bus)</td>
<td>6:40</td>
<td>2</td>
<td>0.1</td>
<td></td>
<td>8</td>
</tr>
</tbody>
</table>

NOTE#: $E(AT_{G31,j,0}) = SAT_{G31,j,0}$
### TABLE 3-2 Input parameters of commuter transit route-34 (PM-peak)

<table>
<thead>
<tr>
<th>Run No. ((j))</th>
<th>Arrival Time ((AT_{G34,j,0}))</th>
<th>Number of Transfer Pass. to each local transit service ((P_{G34,j,L25}^{T_b}, P_{G34,j,L52}^{T_b}))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean ((SAT_{G34,j,0}))</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>0 (Bus)</td>
<td>3:25</td>
<td>2</td>
</tr>
<tr>
<td>1 (Bus)</td>
<td>3:45</td>
<td>2</td>
</tr>
<tr>
<td>2 (Bus)</td>
<td>4:30</td>
<td>2</td>
</tr>
<tr>
<td>3 (Bus)</td>
<td>5:30</td>
<td>2</td>
</tr>
<tr>
<td>4 (Bus)</td>
<td>6:20</td>
<td>2</td>
</tr>
</tbody>
</table>

**NOTE#:** \(E(\ AT_{G34,j,0}\) = SAT_{G34,j,0}\)

### TABLE 3-3 Input parameters of local transit route-25 (PM-peak)

<table>
<thead>
<tr>
<th>Stop ((k))</th>
<th>Scheduled det. time ((SDT_{L25,n,k}))</th>
<th>Travel Time ((TT_{L25,n,k}))</th>
<th>Demand</th>
<th>Alighting Rate ((\omega_{L25,n,k}))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Standard deviation</td>
<td>Skewness</td>
<td>Rnd. Pass. ((\lambda_{L25,n,k}^U))</td>
</tr>
<tr>
<td>0</td>
<td>3:55</td>
<td>-</td>
<td>-</td>
<td>0.1</td>
</tr>
<tr>
<td>1</td>
<td>4:03</td>
<td>7.2</td>
<td>7.2(\times)0.15</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>4:09</td>
<td>5.4</td>
<td>5.4(\times)0.15</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>4:14</td>
<td>4.5</td>
<td>4.5(\times)0.15</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>4:20</td>
<td>5.4</td>
<td>5.4(\times)0.15</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>4:25</td>
<td>4.5</td>
<td>4.5(\times)0.15</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>4:31</td>
<td>5.4</td>
<td>5.4(\times)0.15</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>4:37</td>
<td>5.4</td>
<td>5.4(\times)0.15</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>4:55</td>
<td>7.2</td>
<td>7.2(\times)0.15</td>
<td>1</td>
</tr>
</tbody>
</table>

**NOTE #:** Average arrival time in advance of \(SDT_{L25,n,k}\) of aware passengers (\(\tau_{L25,n,k}\)) is 2.5 min.
TABLE 3-4 Input parameters of local transit route-52 (PM-peak)

<table>
<thead>
<tr>
<th>Stop (k)</th>
<th>Scheduled det. time (SDT&lt;sub&gt;L,52,n,k&lt;/sub&gt;)</th>
<th>Travel Time (TT&lt;sub&gt;L,52,n,k&lt;/sub&gt;)</th>
<th>Demand</th>
<th>Alighting Rate (ω&lt;sub&gt;L,52,n,k&lt;/sub&gt;)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Scheduled det. time (SDT&lt;sub&gt;L,52,n,k&lt;/sub&gt;)</td>
<td>Mean</td>
<td>Standard deviation</td>
<td>Skew -ness</td>
</tr>
<tr>
<td>0</td>
<td>3:45</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>3:51</td>
<td>5.4</td>
<td>5.4×0.15</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>3:57</td>
<td>5.4</td>
<td>5.4×0.15</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>4:05</td>
<td>7.2</td>
<td>7.2×0.15</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>4:12</td>
<td>6.3</td>
<td>6.3×0.15</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>4:16</td>
<td>3.2</td>
<td>3.2×0.15</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>4:19</td>
<td>2.7</td>
<td>2.7×0.15</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>4:25</td>
<td>5.4</td>
<td>5.4×0.15</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>4:33</td>
<td>7.2</td>
<td>7.2×0.15</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>4:39</td>
<td>5.4</td>
<td>5.4×0.15</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>4:45</td>
<td>4.8</td>
<td>4.8×0.15</td>
<td>1</td>
</tr>
</tbody>
</table>

Note#: Average arrival time in advance of SDT<sub>L,52,n,k</sub> of aware passengers (τ<sub>L,52,n,k</sub>) is 2.5 min.

For the AM-peak period, the commuter transit services become the receiving routes, while the local transit services become the feeder routes. FIGURE 3-18 shows the scheduled departure times of Route-31 and Route-34 of the commuter transit service and Route-25 and Route-52 of the local transit service at the transfer point from 6:00AM to 10:00AM. The runs of G31 consist of two groups: some runs are destined to Union Station and some are destined to York Mills, even though they share the same route number. Therefore, Route-31 is divided into G31A for Union Station and G31B for York Mills. As in the PM-peak period, L25 has 30 min. of headway and has two buses serving two trip chains; L52 is served every 15 min. by four buses based on four trip chains.

Since the schedules of the local transit services do not specify the scheduled arrival times, the scheduled arrival times are assumed to be the sum of the scheduled departure time at the stop prior to the transfer point and the travel time from the stop to the transfer point. Note that the arrival times of the local transit service at the transfer point are provided by the downstream model. Therefore, there can exist some gap between the expected arrival times and the scheduled arrival times (i.e., $E( AT_{L,25,n,0} ) \neq SAT_{L,25,n,0}$ and $E( AT_{L,52,n,0} ) \neq SAT_{L,52,n,0}$). That is, the connection...
pairs between feeder runs and receiving runs are determined based on the scheduled arrival times of the feeders, but the estimation of the transfer cost is based on the estimated arrival times by the downstream model. TABLE 3-5, 3-6, 3-7, 3-8 and 3-9 present the input parameters of the AM-peak case in detail.

FIGURE 3-18 Existing schedules at transfer stop (AM-peak)
(Continued on next page)
(e) Local transit route-52 (L52): feeder route

FIGURE 3-18 Existing schedules at transfer stop (AM-peak)

TABLE 3-5 Input parameters of commuter transit route-31A (AM-peak)

<table>
<thead>
<tr>
<th>Run No. ( (j) )</th>
<th>Arrival time ( AT_{31A,j,0} )</th>
<th>Scheduled dep. time ( SDT_{31A,j,0} )</th>
<th>Number of transfer pass.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Standard deviation</td>
<td>Skewness</td>
</tr>
<tr>
<td>1 (Train)</td>
<td>6:41</td>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>2 (Train)</td>
<td>7:06</td>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>3 (Train)</td>
<td>7:21</td>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>4 (Train)</td>
<td>7:56</td>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>5 (Bus)</td>
<td>8:23</td>
<td>2</td>
<td>0.1</td>
</tr>
<tr>
<td>6 (Bus)</td>
<td>9:38</td>
<td>2</td>
<td>0.1</td>
</tr>
</tbody>
</table>

TABLE 3-6 Input parameters of commuter transit route-31B (AM-peak)

<table>
<thead>
<tr>
<th>Run No. ( (j) )</th>
<th>Arrival time ( AT_{31B,j,0} )</th>
<th>Scheduled dep. time ( SDT_{31B,j,0} )</th>
<th>Number of transfer pass.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Standard deviation</td>
<td>Skewness</td>
</tr>
<tr>
<td>1 (Bus)</td>
<td>6:30</td>
<td>2</td>
<td>0.1</td>
</tr>
<tr>
<td>2 (Bus)</td>
<td>6:40</td>
<td>2</td>
<td>0.1</td>
</tr>
<tr>
<td>3 (Bus)</td>
<td>6:55</td>
<td>2</td>
<td>0.1</td>
</tr>
<tr>
<td>4 (Bus)</td>
<td>7:25</td>
<td>2</td>
<td>0.1</td>
</tr>
<tr>
<td>5 (Bus)</td>
<td>7:55</td>
<td>2</td>
<td>0.1</td>
</tr>
<tr>
<td>6 (Bus)</td>
<td>8:35</td>
<td>2</td>
<td>0.1</td>
</tr>
<tr>
<td>7 (Bus)</td>
<td>9:25</td>
<td>2</td>
<td>0.1</td>
</tr>
</tbody>
</table>
TABLE 3-7 Input parameters of commuter transit route-34 (AM-peak)

<table>
<thead>
<tr>
<th>Run No. (j)</th>
<th>Arrival time ($AT_{G34,j,0}$)</th>
<th>Scheduled dep. time ($SDT_{G34,j,0}$)</th>
<th>Number of transfer pass.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Standard deviation</td>
<td>Skewness</td>
</tr>
<tr>
<td>0 (Bus)</td>
<td>6:05</td>
<td>2</td>
<td>0.1</td>
</tr>
<tr>
<td>1 (Bus)</td>
<td>6:55</td>
<td>2</td>
<td>0.1</td>
</tr>
<tr>
<td>2 (Bus)</td>
<td>7:35</td>
<td>2</td>
<td>0.1</td>
</tr>
<tr>
<td>3 (Bus)</td>
<td>7:40</td>
<td>2</td>
<td>0.1</td>
</tr>
<tr>
<td>4 (Bus)</td>
<td>8:25</td>
<td>2</td>
<td>0.1</td>
</tr>
<tr>
<td>5 (Bus)</td>
<td>8:55</td>
<td>2</td>
<td>0.1</td>
</tr>
<tr>
<td>6 (Bus)</td>
<td>9:25</td>
<td>2</td>
<td>0.1</td>
</tr>
</tbody>
</table>

TABLE 3-8 Input parameters of local transit route-25 (AM-peak)

<table>
<thead>
<tr>
<th>Stop (k)</th>
<th>Scheduled det. time ($SDT_{L25,n,k}$)</th>
<th>Travel Time ($TT_{L25,n,k}$)</th>
<th>Demand</th>
<th>Alighting Rate ($\omega_{L25,n,k}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Standard deviation</td>
<td>Skewness</td>
<td>Rnd. Pass. ($\lambda^U_{L25,n,k}$)</td>
</tr>
<tr>
<td>0</td>
<td>5:55</td>
<td>-</td>
<td>-</td>
<td>0.1</td>
</tr>
<tr>
<td>1</td>
<td>6:03</td>
<td>72</td>
<td>72×0.15</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>6:09</td>
<td>54</td>
<td>54×0.15</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>6:14</td>
<td>45</td>
<td>45×0.15</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>6:20</td>
<td>54</td>
<td>54×0.15</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>6:25</td>
<td>45</td>
<td>45×0.15</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>6:31</td>
<td>54</td>
<td>54×0.15</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>6:37</td>
<td>54</td>
<td>54×0.15</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>6:55</td>
<td>72</td>
<td>72×0.15</td>
<td>1</td>
</tr>
</tbody>
</table>

Note#: Average arrival time in advance of $SDT_{L52,n,k}$ of aware passengers ($\tau_{L52,n,k}$) is 2.5 min.
TABLE 3-9 Input parameters of local transit route-52 (AM-peak)

<table>
<thead>
<tr>
<th>Stop (k)</th>
<th>Scheduled det. time (SDT_{L,25,n,k})</th>
<th>Travel Time (TT_{L,25,n,k})</th>
<th>Demand</th>
<th>Alighting Rate (\omega_{L,25,n,k})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Standard deviation</td>
<td>Skewness</td>
<td>Rnd. Pass. (\lambda^L_{L,25,n,k})</td>
</tr>
<tr>
<td>0</td>
<td>6:00</td>
<td>-</td>
<td>-</td>
<td>0.2</td>
</tr>
<tr>
<td>1</td>
<td>6:06</td>
<td>54</td>
<td>5.4x0.15</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>6:12</td>
<td>54</td>
<td>5.4x0.15</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>6:20</td>
<td>72</td>
<td>7.2x0.15</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>6:27</td>
<td>63</td>
<td>6.3x0.15</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>6:31</td>
<td>32</td>
<td>3.2x0.15</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>6:34</td>
<td>27</td>
<td>2.7x0.15</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>6:40</td>
<td>54</td>
<td>5.4x0.15</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>6:48</td>
<td>72</td>
<td>7.2x0.15</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>6:54</td>
<td>54</td>
<td>5.4x0.15</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>7:00</td>
<td>48</td>
<td>4.8x0.15</td>
<td>1</td>
</tr>
</tbody>
</table>

Note#: Average arrival time in advance of SDT_{L,52,n,k} of aware passengers (\tau_{L,52,n,k}) is 2.5 min.

As explained earlier, the transfer passengers of the AM-peak case are defined as \( P_{L,i,j}^{Tr^*} \), the number of transfer passengers who aim to transfer from local transit route \( L \) to run \( j \) of commuter transit route \( i \) (i.e., real demand for run \( j \) on commuter transit route \( i \) from local route \( L \)). If the passengers have to take local transit run \( n \) to reach a specific commuter run (run \( j \) of route \( i \)), \( P_{L,i,j}^{Tr^*} \) can be broken down into a stop level as shown in Figure 3-19. \( P_{L,i,j,k}^{Tr^*} \) is the number of transfer passengers who aim transfer from stop \( k \) of route \( L \) to run \( j \) of route \( i \).
FIGURE 3-19 Breakdown of transfer passengers for AM-peak case

For the AM peak cases, therefore, $P_{L,i,j}^{Tr*}$ is defined at each stop level, based on $P_{L,i,j,k}^{Tr*}$ (Equation 3-64). $P_{L,i,j,k}^{Tr*}$ is included in the group of aware passengers at each stop in the downstream model. As shown in Equation 3-65, each $P_{L,i,j,k}^{Tr*}$ is defined by a distribution ratio, $\nu_{L,i,j,k}$. TABLE 3-10 presents the used values of $\nu_{L,i,j,k}$.

\[ P_{L,i,j}^{Tr*} = \sum_{k=1}^{K_i} P_{L,i,j,k}^{Tr*} \quad \text{(EQ. 3-64)} \]

\[ P_{L,i,j,k}^{Tr*} = \nu_{L,i,j,k} \times P_{L,i,j}^{Tr*} \quad \text{(EQ. 3-65)} \]

Where:

$P_{L,i,j,k}^{Tr*}$ : Number of transfer passengers from stop $k$ of local transit route $L$ aiming to transfer to run $j$ on commuter transit route $i$ (i.e., real demand for run $j$ on route $i$ from each stop of route $L$)

$\nu_{L,i,j,k}$ : Distribution rate of transfer passengers, $P_{L,i,j}^{Tr*}$, at stop $k$. ($\sum_{k=1}^{K_i} \nu_{L,i,j,k} = 1.0$)
6.2 Results: PM-Peak Case

The developed model provides schedulers of transit services with a tool to devise and test various schemes to reduce the overall transfer cost. In this study, the model is used to test the following four schemes:

- Scheme-1: shifting the entire local bus schedule only, thus maintaining the scheduled headway
- Scheme-2: allowing headway changes, as well as schedule shifting
- Scheme-3: allowing schedule shifting, headway changes, plus the holding policy
- Scheme-4: similar to Scheme-3, but determining separately the offsets and holding times through two stages: Stage-a) using the optimized offsets of Scheme-2, and Stage-b) adding the holding times to the timetable resulting from Stage-a. Scheme-3 simultaneously optimizes the schedule shifting, headway changes, and the holding policy. However, Scheme-4 applies the holding policy in order to ensure the transfers of a timetable optimized by Scheme-2.

This study optimizes the schedules of local transit Route-25 (L25) and Route-52 (L52) separately using these schemes, since transfers between the 2 routes are ignored. TABLE 3-11 summarizes
the solution space of and the solution approach to each scheme. In Scheme-1, the ranges of the offset values are given based on the headways of the existing schedules. The minimum and maximum of the shifting ranges are \(-\text{Headway}/2\) and \(\text{Headway}/2\), respectively; in Scheme-1 of Route-25, for instance, \(\alpha_{L,25}\) would range from \(-15\) \(=30/2\) to 15. Since the size of the solution space is small enough to search all the possible offset values, the optimal solutions of this scheme can be found by inspecting every offset value.

In Scheme-2 of Route-25, \(\beta_{L,25,n}\) is from \(-8\) to 7 (4-bit binary string) to prevent the headway being too large. By setting \(\alpha_{L,25}\) from \(-7\) to 8 (4-bit binary string), simultaneously, the total range of the shift is from \(-15\) to 15. \(HT_{L,25,n}\) of Scheme-3 is arbitrarily set to range from 0 to 3 mins (2-bit binary string). The solution spaces of Route-52 are determined in the same way. The solutions of Schemes 2 and 3 are searched \(\text{via}\) a genetic algorithm. The solution of Scheme-1, however, is found by examining all the possible solutions, since the solutions space is sufficiently small. Scheme-4 is to assign holding times in order to ensure the transfers of a timetable optimized by Scheme-2 \(\text{via}\) a heuristic rule, which will be explained in detail later.

### TABLE 3-11 Solution spaces according to optimization schemes (PM-peak)

<table>
<thead>
<tr>
<th>Routes</th>
<th>Optimization Schemes</th>
<th>Scheme-1</th>
<th>Scheme-2</th>
<th>Scheme-3</th>
<th>Scheme-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>L25</td>
<td>Solution Space</td>
<td>(\alpha_{L,25}^{\min} / \alpha_{L,25}^{\max})</td>
<td>(-15/15)</td>
<td>(-7/8)</td>
<td>(-7/8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(\beta_{L,25,n}^{\min} / \beta_{L,25,n}^{\max})</td>
<td>N/A</td>
<td>(-8/7)</td>
<td>(-8/7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(HT_{L,25}^{\max})</td>
<td>N/A</td>
<td>N/A</td>
<td>(3)</td>
</tr>
<tr>
<td></td>
<td>Possible number of Solution Space</td>
<td>31</td>
<td>(2^12) ((=4096))</td>
<td>(2^28)</td>
<td>(2^28)</td>
</tr>
<tr>
<td>L52</td>
<td>Solution Space</td>
<td>(\alpha_{L,52}^{\min} / \alpha_{L,52}^{\max})</td>
<td>(-7/8)</td>
<td>(-3/4)</td>
<td>(-3/4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(\beta_{L,52,n}^{\min} / \beta_{L,52,n}^{\max})</td>
<td>N/A</td>
<td>(-4/3)</td>
<td>(-4/3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(HT_{L,52}^{\max})</td>
<td>N/A</td>
<td>N/A</td>
<td>(3)</td>
</tr>
<tr>
<td></td>
<td>Possible number of Solution Space</td>
<td>16</td>
<td>(2^15) ((=32768))</td>
<td>(2^47)</td>
<td>(2^47)</td>
</tr>
<tr>
<td></td>
<td>Solution Method</td>
<td>Exhaustive sequential search</td>
<td>GA</td>
<td>GA</td>
<td>GA</td>
</tr>
</tbody>
</table>
TABLE 3-12 and 3-13 present the optimized results of the two local transit routes using the developed model. Each table includes the optimal solutions and the associated cost components for the four optimization schemes. FIGURE 3-20 and 3-21 illustrate how the schedule of L25 is revised according to Scheme-1, Scheme-2 and Scheme-3. FIGURE 3-20 presents the revised schedule with connection pairs and FIGURE 3-21 shows the successful connection cost, the missed connection cost and the probability of the successful connection from each commuter transit to its primary transit run in detail. FIGURE 3-22 and 3-23 show the results of L52 in the same way.

<table>
<thead>
<tr>
<th>Optimization Schemes</th>
<th>Base Case</th>
<th>Scheme-1</th>
<th>Scheme-2</th>
<th>Scheme-3</th>
<th>Scheme-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solution</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_{L25}$</td>
<td>N/A</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>$\beta_{L25,n}$</td>
<td>N/A</td>
<td>N/A</td>
<td>2,0</td>
<td>0,-2</td>
<td>2,0</td>
</tr>
<tr>
<td>$HT_{L25,n}$</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>2,3,3,3,3,3,3,3,3,3,1,0</td>
<td>0,0,3,0,3,0,0,0</td>
</tr>
<tr>
<td>Transfer from GO-31</td>
<td>Succ. Conn. Cost</td>
<td>2904.2</td>
<td>1190.3</td>
<td>1300.9</td>
<td>1149.3</td>
</tr>
<tr>
<td></td>
<td>Miss. Conn. Cost</td>
<td>0.0</td>
<td>61.4</td>
<td>39.4</td>
<td>4.0</td>
</tr>
<tr>
<td>Transfer from GO-34</td>
<td>Succ. Conn. Cost</td>
<td>614.5</td>
<td>279.7</td>
<td>339.2</td>
<td>276.0</td>
</tr>
<tr>
<td></td>
<td>Miss. Conn. Cost</td>
<td>196.4</td>
<td>385.8</td>
<td>156.9</td>
<td>79.6</td>
</tr>
<tr>
<td>Total Transfer Cost</td>
<td>3715.1</td>
<td>1917.1</td>
<td>1836.4</td>
<td>1508.8</td>
<td>1713.3</td>
</tr>
<tr>
<td>Total Stop Waiting Cost</td>
<td>3229.1</td>
<td>3247.6</td>
<td>3263.3</td>
<td>3319.1</td>
<td>3277.4</td>
</tr>
<tr>
<td>Total Cost</td>
<td>6944.2</td>
<td>5164.7</td>
<td>5099.6</td>
<td>4828.0</td>
<td>4990.72</td>
</tr>
</tbody>
</table>

(Unit of Cost: pass.·min.)
### TABLE 3-13 Optimized results of local transit route-52 (PM peak)

<table>
<thead>
<tr>
<th>Optimization Schemes</th>
<th>Base Case</th>
<th>Scheme-1</th>
<th>Scheme-2</th>
<th>Scheme-3</th>
<th>Scheme-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solution</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_{L25}$</td>
<td>N/A</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\beta_{L25,n}$</td>
<td>N/A</td>
<td>N/A</td>
<td>2,1,1,0</td>
<td>1,0,0,-1</td>
<td>2,1,1,0</td>
</tr>
<tr>
<td>$HT_{L25,n}$</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>0,1,3,2,0, 3,1,2,3,3, 3,2,3,1,1, 0, 3,0,3,0,0, 3,0,3,0,0, 0,</td>
<td></td>
</tr>
<tr>
<td>Transfer from GO-31</td>
<td>Succ. Conn. Cost</td>
<td>474.2</td>
<td>474.2</td>
<td>561.7</td>
<td>483.4</td>
</tr>
<tr>
<td>Miss. Conn. Cost</td>
<td>234.5</td>
<td>234.5</td>
<td>120.7</td>
<td>30.8</td>
<td>31.4</td>
</tr>
<tr>
<td>Transfer from GO-34</td>
<td>Succ. Conn. Cost</td>
<td>169.2</td>
<td>169.2</td>
<td>200.5</td>
<td>176.19</td>
</tr>
<tr>
<td>Miss. Conn. Cost</td>
<td>243.5</td>
<td>243.5</td>
<td>176.6</td>
<td>55.10</td>
<td>25.8</td>
</tr>
<tr>
<td>Total Transfer Cost</td>
<td>1121.5</td>
<td>1121.5</td>
<td>1059.6</td>
<td>745.4</td>
<td>847.6</td>
</tr>
<tr>
<td>Total Stop Waiting Cost</td>
<td>4572.5</td>
<td>4572.5</td>
<td>4591.5</td>
<td>4695.0</td>
<td>4649.5</td>
</tr>
<tr>
<td>Total Cost</td>
<td>5694.0</td>
<td>5694.0</td>
<td>5651.1</td>
<td>5440.4</td>
<td>5497.2</td>
</tr>
</tbody>
</table>

(Unit of Cost: pass·min.)

![Diagram](a) Original schedule

FIGURE 3-20 Optimized schedules (L25:PM-peak)

(Continued on next page)
(b) Scheme-1

(c) Scheme-2 and Scheme-4

(d) Scheme-3

Arrival of GO31, Arrival of GO34, Departure of L25, Connection Pair.

*: Run ID, **: Route Name, ***: Schedule Shift($\alpha_{L25} + \beta_{L25,n}$)

FIGURE 3-20 Optimized schedules (L25:PM-peak)
FIGURE 3-21 Transfer costs according to optimization schemes (L25:PM-peak)
(a) Original schedule and Scheme-1

(b) Scheme-2 and Scheme-4

(c) Scheme-3

Arrival of GO31, Arrival of GO34, Departure of L52, Connection Pair

*: Run ID, **: Route Name, ***: Schedule Shift($\alpha_{l52} + \beta_{l52,n}$)

FIGURE 3-22 Optimized schedules (L52:PM-peak)
The results reveal that the tested schemes effectively reduce the transfer times, depending on the given current schedules; the total costs of L25 are significantly reduced even with the simple shifting of the existing schedule, while the improvement on L52 is less pronounced. The results of L52 imply that, in terms of transfer, it is currently scheduled properly. The results make sense in that the schedule of L25 is revised in Scheme-1 so that its first run starts at 3:30PM with 30
min. of headway, in the same way that Route-52 operates with 15 min. of headway from 3:15PM.

Scheme-2 does not change the schedule radically, but tweaks the schedule one or two minutes more, compared to Scheme-1 (see TABLE 3-12 and 3-13). In general, Scheme-2 results in better transfers than Scheme-1 due to its additional flexibility: it improves the probabilities of successful connections, and decreases the transfer costs, as shown in FIGURE 3-21 (c) and FIGURE 3-23 (b). Because it changes the given headways among consecutive transit units, however, the waiting times of the downstream passengers can be seen to increase, as presented in TABLE 3-12 and 3-13. In fact, it is the unaware passengers among the downstream passengers who are affected by this change of headway.

The results of Scheme-3 indicate that the holding policy is beneficial in further reducing the transfer costs. It makes use of the holding time to reduce the missed connections, and also to reduce the transfer slack time (the difference between the arrival time of a feeder and the scheduled departure time of a receiver). Under Scheme-3, while transfer passengers experience less missed connections, their transfer time is substantially decreased because of the reduced slack time. However, this scheme causes further delays to stop passengers, as shown in TABLE 3-12 and 3-13. In total, the overall cost experienced by all passengers is reduced under this scheme.

To better understand the behaviour of the developed model, it is necessary to explore how the transfer costs from feeder runs to a receiving run are changed according to the given optimization schemes. FIGURE 3-24 provides detailed analysis focusing on one specific transfer case of Route-25: the transfer to run 2 of Route-25 (L25-2) from the three feeder runs of GO-31 and GO-34 (G31-1: run 1 of GO-31, G34-2: run 2 of GO-34, and G31-2: run 2 of GO-31). As seen in FIGURE 3-24 (a), the optimization model shifts the original scheduled departure time of L25-2 by 5 min. under Scheme-1 and 2, while shifting it by 3 min. with 3 min. of holding time under Scheme-3.
(a) Scheduled departure time of L25-2 according to optimization schemes

(b) Transfer costs of G31-1, G34-2 and G31-2 according to schedule shifts of L25-2
(Scheme-1 and Scheme-2)

(c) Transfer costs and the departure delay against some possible schedules of L25-2
(Scheme-3)

FIGURE 3-24 Analysis of a transfer from feeder routes to a receiving route
(Continued on next page)
FIGURE 3-24 Analysis of a transfer from feeder routes to a receiving route

FIGURE 3-24 (b) explains the trade-off procedure occurring during the transfer optimization with Scheme-1 and Scheme-2, which revise the given schedule through the schedule shift. It shows the transfer passengers’ costs of the three feeders change as the schedule of L25-2 is shifted to the right side. It can be seen that the transfer costs of G31-1 and G34-2 (‘Tr. Cost of G31-1’ and ‘Tr. Cost of G34-2’ in the figure) increase slightly as the scheduled departure time of L25-2 moves to the right side. Since G31-1 and G34-2 arrive much earlier than the scheduled departure time of L25-2, their missed connection costs are almost zero regardless of the shift of L25-2. Their transfer costs only consist of the successful connection costs. However, the transfer cost of G31-2 is directly influenced by the shift. When L25-2 is shifted by 2 min., the transfer slack time between G31-2 and L25-2 is zero (i.e., the scheduled arrival time of G31-2 is equal to the scheduled departure time of L25-2); accordingly, there exists a high risk that the transfer from G31-2 to L25-2 will fail. The figure plots the missed connection cost (‘MC Cost of G31-2’) and the successful connection cost (‘SC Cost of G31-2’) of G31-2 together with its total transfer cost (‘Tr. Cost of G31-2’). The missed connection cost is quite high with zero transfer slack time, but it gradually decreases as the slack time increases (i.e., the scheduled departure of L25-2 is shifted to the right side); on the other hand, as the slack time increases, so too does the successful connection. That is, the decision of the optimal departure time of L25-2 with respect to the total transfer cost of G31-1, G34-2 and G31-2 is the trade-off between their missed connection costs and their successful connection costs, and the provision of too-long slack times...
causes the increase of the total transfer cost. The figure indicates that the optimal departure time of L25-2 is 4:00PM; that is, the schedule of Route-25—revised according to Scheme-1 and Scheme-2—allows the transfer passengers from G31-1, G34-2 and G31-2 to transfer to L25-2 with the minimum transfer cost. However, it should be noted that the revised schedule is not beneficial for all of the transfer passengers, since the objective function is to minimize overall cost. With the optimized results of Route-25 by implementing Scheme-1 (FIGURE 3-21 (b)), for example, the transfer from run 3 of GO-34 (G34-3) to Route-25 has about 30 percent of success for that transfer (i.e., the missed connection cost of that transfer is very high). By implementing Scheme-2, the additional flexibility allows for a much greater chance to increase the probability of a successful transfer significantly; so, a 30 percent of success could potentially be more than double.

If local bus holding is allowed, such as in Scheme-3, it is possible to optimize the schedule of L25-2 making use of the holding time. From the perspective of the transfer passengers, a schedule shift of two min. with three min. of holding time seems to be equivalent to five min. of schedule shift without the holding time, in that the transfer is successful if G31-2 arrives at the transfer point earlier than 4:00PM. However, the transfer cost of the former is less than that of the latter: if L25-2 is scheduled to depart at 3:57PM and can be held until 4:00PM, it will depart as soon as it processes the passengers of G31-2 instead of staying at the transfer point until 4:00PM; however, if L25-2 is scheduled to depart at 4:00PM without the holding time, it must depart at 4:00PM. That is, the dispatching of L25-2 with the holding time is absolutely favourable for the transfer passengers.

However, it should be noted that the holding would break the schedule adherence of L25-2 at the transfer point, because the holding allows L25-2 to depart beyond the scheduled departure time if G32-2 is delayed. FIGURE 3-24 (c) plots the transfer costs together with the departure delay of L25-2 for some selected schedules of L25-2. It shows that the schedule of Scheme-3 (SS=+3, Hold=+3) is better than that of Schemes 1 and 2 (SS=+5, Hold=0) in terms of the transfer cost. However, Scheme-3 causes about one min. departure delay. As seen in TABLE 3-12 and 3-13, Scheme-3 lessens the overall transfer cost, while resulting in further delays to stop passengers; however, the overall cost of all passengers is reduced in total. That is, Scheme-3 can bring a further reduction of the total cost by making use of the holding time with a small margin of slack time, instead of increasing the transfer slack time. However, the margin of benefit by allowing
the holding of the receiving run depends on transfer passengers’ demand and stop passengers’ demand. If the demand of transfer passengers is much smaller than that of stop passengers, the benefit of the holding is less than the cost of the holding (i.e., the increase of the downstream passengers’ cost incurred by the departure delay).

FIGURE 3-24 (d) presents the relationship between the slack time and the holding time. The effect of the holding on the transfer cost is closely related to the transfer slack time. The provision of the holding time is only beneficial when the probability of successful connection is relatively low. However, the probability of successful connection is high if enough slack time is given, rendering holding useless. For the case of SS=+6 of FIGURE 3-24 (d), where the scheduled departure time of L25-2 is 4:01PM and, accordingly, enough slack time (i.e., 4 min.) is provided for the transfer from G31-2 to L25-2, most of the transfer would be successful without the holding time. The provision of the holding policy has no effect on the transfer cost because L25-2 does not stay at the transfer point until the end of the given holding time but departs as soon as it meets G31-2 after its scheduled departure time. Therefore, the cost does not vary regardless of any holding time as shown in the figure. In the case of SS=+5, there exists some benefit with the holding time of one min., but the cost does not decrease any more from the holding time of two min. In contrast, the total transfer cost of SS=+3, where the probability of the successful connection is low, decreases substantially as the holding time increases.

Although Scheme-3 reduces the total cost markedly, it can result in timetables that possibly break the schedule adherence of downstream stops due to the holding time, as discussed. This problem motivates Scheme-4 where the holding time is assigned to ensure the transfers of a timetable optimized by Scheme-2. In Scheme-2 of FIGURE 3-21 and 3-23, there exist some transfers that have a relatively low probability of a successful transfer (for example, the transfers from GO34-3 and GO34-4 in FIGURE 3-21 (c); the transfers from GO31-6, GO31-8, GO31-10, GO34-2, GO34-3 and GO34-4 in FIGURE 3-23 (b)). It is possible to assign holding times to those transfers via revising the results of Scheme-2. One possible way is to incrementally assign a holding time to a transfer where the margin of the total cost decreases mostly by adding the holding time by 1, and repeat this holding time assignment until the margin is trivial. As shown in TABLE 3-12 and 3-13, and FIGURE 3-25, Scheme-4 not only improves the successful probabilities but also reduces the overall costs when compared with Scheme-2. Even though the overall costs under this scheme are inferior to those of Scheme-3, Scheme-4 establishes a
timetable where transit units serve the downstream stops in a more stable manner, and consequently reduces the waiting costs of the stop passengers.

The results show that the developed model can be applied to revise the existing schedules with various optimization schemes, and reasonably improves the overall transfer-related costs.
6.3 Results: AM-Peak Case

The optimization of AM-peak is basically the same as that of PM-peak. In contrast to the PM-peak case, though, the commuter transit service becomes the receiving route during the AM-peak, while the local transit service becomes the feeder route. However, the holding policy is not applied to the commuter transit routes because of their long-haul status. Therefore, the two optimization schemes tested are without the holding control: Scheme-1 shifts the entire schedule of the local transit route, and Scheme-2 is the schedule shift allowing for headway changes.

Accounting for the fact that the additional cost term, the hidden waiting cost, may not be equivalent to the transfer passengers’ cost and the non-transfer passengers’ cost in terms of the perception of passengers, three different weights, 0, 0.5, and 1.0, are applied to the hidden waiting cost: the weight of 0 completely ignores the hidden waiting cost; the weight of 0.5 represents a 50 percent reduction to the passengers perception of the hidden waiting cost; and the weight of 1.0 equalizes the hidden waiting cost with the actual waiting cost of both the transfer and non-transfer passengers.

The solution spaces of the two schemes are defined in the same manner as that of PM peak case, as seen in TABLE 3-14. The solution methods are the same as well.

<table>
<thead>
<tr>
<th>Routes</th>
<th>Optimization Schemes</th>
<th>Scheme-1</th>
<th>Scheme-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>L25 Solution Space</td>
<td>$\alpha_{\text{min}}^{L25} / \alpha_{\text{max}}^{L25}$</td>
<td>$\min / 25$</td>
<td>$\max / 25$</td>
</tr>
<tr>
<td>Number of Possible Solutions</td>
<td>31</td>
<td>4096 (12 bit)</td>
<td></td>
</tr>
<tr>
<td>L52 Solution Space</td>
<td>$\alpha_{\text{min}}^{L52} / \alpha_{\text{max}}^{L52}$</td>
<td>$\min / 52$</td>
<td>$\max / 52$</td>
</tr>
<tr>
<td>Number of Possible Solutions</td>
<td>16</td>
<td>32768 (15 bit)</td>
<td></td>
</tr>
<tr>
<td>Solution Method</td>
<td>Exhaustive sequential search</td>
<td>GA</td>
<td></td>
</tr>
</tbody>
</table>
TABLE 3-15 and 3-16 present the optimized results of the two local transit routes with a hidden waiting cost weight of 0.5. The tables show the offset values and the cost components for the two schemes. FIGURE 3-26, 3-27, 3-28, and 3-29 illustrate the results in detail (with a hidden waiting cost weight of 0.5). FIGURE 3-26 and 3-28 draw the revised schedule with connection pairs. FIGURE 3-27 and 3-29 plot the successful connection cost, the missed connection cost and the probability of the successful connection from each commuter transit to its primary transit run in detail.

**TABLE 3-15 Optimized results of local transit route-25 (AM peak)**

<table>
<thead>
<tr>
<th>Optimization Schemes</th>
<th>Base</th>
<th>Scheme-1</th>
<th>Scheme-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solution</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha_{25} )</td>
<td>-</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>( \beta_{25,n} )</td>
<td>-</td>
<td>-1,1</td>
<td>-1,1</td>
</tr>
<tr>
<td>Transfer to GO-31A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Succ. Conn. Cost</td>
<td>1330.9</td>
<td>715.3</td>
<td>742.5</td>
</tr>
<tr>
<td>Miss. Conn. Cost</td>
<td>1.9</td>
<td>200.3</td>
<td>98.2</td>
</tr>
<tr>
<td>Hidden. Conn. Cost</td>
<td>0.0</td>
<td>228.0</td>
<td>228.0</td>
</tr>
<tr>
<td>Transfer to GO-31B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Succ. Conn. Cost</td>
<td>672.4</td>
<td>311.6</td>
<td>287.1</td>
</tr>
<tr>
<td>Miss. Conn. Cost</td>
<td>172.7</td>
<td>85.1</td>
<td>84.7</td>
</tr>
<tr>
<td>Hidden. Conn. Cost</td>
<td>60.0</td>
<td>40.0</td>
<td>40.0</td>
</tr>
<tr>
<td>Transfer to GO-34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Succ. Conn. Cost</td>
<td>825.2</td>
<td>401.5</td>
<td>393.4</td>
</tr>
<tr>
<td>Miss. Conn. Cost</td>
<td>276.7</td>
<td>75.5</td>
<td>78.2</td>
</tr>
<tr>
<td>Hidden. Conn. Cost</td>
<td>0.0</td>
<td>20.0</td>
<td>20.0</td>
</tr>
<tr>
<td>Total Transfer Cost</td>
<td>3339.7</td>
<td>2077.2</td>
<td>1972.1</td>
</tr>
<tr>
<td>Total Stop Waiting Cost</td>
<td>3561.8</td>
<td>3638.3</td>
<td>3649.6</td>
</tr>
<tr>
<td>Total Cost</td>
<td>6901.6</td>
<td>5715.5</td>
<td>5621.6</td>
</tr>
</tbody>
</table>

(Unit of Cost: pass. · min. / Weight of hidden waiting cost = 0.5)
### TABLE 3-16 Optimized results of local transit route-52 (AM peak)

<table>
<thead>
<tr>
<th>Optimization Schemes</th>
<th>Base</th>
<th>Scheme-1</th>
<th>Scheme-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solution</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_{L52}$</td>
<td></td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>$\beta_{L52,m}$</td>
<td></td>
<td></td>
<td>0.0,0.1</td>
</tr>
<tr>
<td>Transfer to GO-31A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Succ. Conn. Cost</td>
<td>788.8</td>
<td>317.8</td>
<td>301.4</td>
</tr>
<tr>
<td>Miss. Conn. Cost</td>
<td>2.9</td>
<td>151.3</td>
<td>153.4</td>
</tr>
<tr>
<td>Hidden. Conn. Cost</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Transfer to GO-31B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Succ. Conn. Cost</td>
<td>121.7</td>
<td>298.0</td>
<td>281.4</td>
</tr>
<tr>
<td>Miss. Conn. Cost</td>
<td>671.3</td>
<td>135.7</td>
<td>136.4</td>
</tr>
<tr>
<td>Hidden. Conn. Cost</td>
<td>0.0</td>
<td>40.0</td>
<td>40.0</td>
</tr>
<tr>
<td>Transfer to GO-34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Succ. Conn. Cost</td>
<td>163.1</td>
<td>352.8</td>
<td>336.2</td>
</tr>
<tr>
<td>Miss. Conn. Cost</td>
<td>836.9</td>
<td>13.4</td>
<td>13.8</td>
</tr>
<tr>
<td>Hidden. Conn. Cost</td>
<td>0.0</td>
<td>20.0</td>
<td>20.0</td>
</tr>
<tr>
<td>Total Transfer Cost</td>
<td>2584.7</td>
<td>1329.1</td>
<td>1282.6</td>
</tr>
<tr>
<td>Total Stop Waiting Cost</td>
<td>5067.0</td>
<td>5128.5</td>
<td>5149.3</td>
</tr>
<tr>
<td>Total Cost</td>
<td>7651.6</td>
<td>6457.6</td>
<td>6431.9</td>
</tr>
</tbody>
</table>

(Unit of Cost: pass.·min. / Weight of hidden waiting cost = 0.5)
(a) Original schedule

(b) Scheme-1

(c) Scheme-2

Departure of GO31A, Departure of GO31B, Departure of GO34,
Arrival of L25, Connection Pair, Connection pair incurring hidden waiting cost.

*: Run ID, **: Route Name, ***: Schedule Shift($\alpha_{25} + \beta_{25,n}$)

FIGURE 3-26 Optimized schedules (L25:AM-peak)
FIGURE 3-27 Transfer cost according to optimization schemes (L25:AM-peak)
(a) Original schedule

(b) Scheme-1

(c) Scheme-2

Departure of GO31A, Departure of GO31B, Departure of GO34,
Arrival of L52, Connection Pair, Connection pair incurring hidden waiting cost

*: Run ID, **: Route Name, ***: Schedule Shift($\alpha_{L52} + \beta_{L52,n}$)

FIGURE 3-28 Optimized schedules (L52:AM-peak)
The results show that the two schemes significantly reduce the total costs for the two routes. The two schemes efficiently decrease the successful and missed connection costs of transfer passengers to the commuter transit routes, as seen in FIGURE 3-27 and 3-29. FIGURE 3-29 (a) shows the original schedule of L52 does not have enough transfer slack time. The two
optimization schemes improve the overall quality of the transfers efficiently. The revisions of the schedules are based on the trade-off process exemplified in FIGURE 3-24.

Compared to Scheme-1, Scheme-2 does not reform the schedules entirely, but alters them by adding or subtracting one or two minutes, while breaking the original headway of the routes. Therefore, Scheme-2 reduces the transfer passengers’ cost more than Scheme-1 does. However, the transfer optimization approach cannot provide a satisfactory connection to all the transfers. The transfer to GO31B-1 in FIGURE 3-29 (c) has a large amount of missed connection even when applying Scheme-2. These observations are consistent with the results of the PM peak case.

As seen in TABLE 3-15 and 3-16, the revised schedules likely increase the hidden waiting cost. That is, the revised schedules prevent some transfer passengers from getting the commuter transit run they want, in effect forcing them to take another run earlier than their desired target. Even though the hidden waiting cost increases, the total cost incurred by all passengers is reduced in total. As noted previously, the weights of 0.0 and 1.0 are tested, but do not result in different offset values. The hidden waiting cost is relatively trivial to the benefit of shifting the given schedules and, consequently, the optimized results are not influenced by the weight of hidden waiting costs.

Even though the effect of the hidden waiting cost across the given routes is negligible, its value to optimizing transfers cannot be underestimated. Actually, the hidden waiting costs can be related to not only the inconvenience of passengers transferring to a run they do not want, but also to the occupancies of the runs of a receiving route. The increase of the hidden waiting cost can imply a risk that transfer passengers take an overcrowded transit run. The revised schedule of Route-25 (FIGURE 3-26) exemplifies this risk. Under the revised schedules, the transfer passengers transferring to GO31A-3 can take GO31A-2 instead. Consequently, the revision of the schedule can cause the over-saturation of run GO31A-2. In spite of increasing the weight of the hidden waiting cost to 1.0, the benefit of shifting the schedule is larger than the consequence of the hidden waiting costs in the given case. Taking the hidden waiting cost into account, however, the model is helpful in detecting the possibility of the over-saturation of a receiving run. This risk should be carefully reviewed as to whether it causes significant over-saturation on the receiving run or not.
6.4 Recommendations for GA Parameters

Determining appropriate GA features and parameters, such as selection method, population size, crossover rate and mutation rate, is essential when applying GAs to an optimization problem; an unsuitable set of parameters may trap GAs into local minima or converge more slowly. General recommendations suggested by De Jong (Goldberg, 1989) include: a high crossover rate, low mutation rate and a moderate population size. Following the suggestions, Goldberg puts forth the following parameters as a possible alternative: crossover rates (CR) of 0.6, mutation rates (MR) of 0.033, and Population sizes (PS) of 30. The rank selection is generally the preferred selection method, because it is more robust than the roulette wheel selection against the risk of premature convergence and stagnation (Whitley, 1989). However, there does not exist a unique rule for the selection of the parameters. The parameters interact nonlinearly with one another, and the best combination of the parameters depends on the specifics of the problem at hand. The GA of this study is tested with different combinations of the GA parameters in order to analyze the efficiency of the GA according to the parameters, and to make recommendations about the selection of the parameters. TABLE 3-17 summarizes the tested parameters.

<table>
<thead>
<tr>
<th>GA Parameters</th>
<th>PM-peak schedule optimization using Scheme-2 (for L25 and L52)</th>
<th>PM-peak schedule optimization using Scheme-3 (for L25 and L52)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection Method</td>
<td>Rank Selection + Elitism</td>
<td>Rank Selection + Elitism</td>
</tr>
<tr>
<td>Population Size (PS)</td>
<td>10, 30</td>
<td>10, 50</td>
</tr>
<tr>
<td>Crossover Rate (CR)</td>
<td>0.6, 0.8</td>
<td>0.6, 0.8</td>
</tr>
<tr>
<td>Mutation Rate (MR)</td>
<td>0.01, 0.03, 0.05</td>
<td>0.01, 0.03, 0.05</td>
</tr>
<tr>
<td>Total number of parameter combinations</td>
<td>12 (FIGURE 3-30 and 3-31 for L25) (FIGURE 3-32 and 3-33 for L52)</td>
<td>12 (FIGURE 3-34 and 3-35 for L25) (FIGURE 3-36 and 3-37 for L52)</td>
</tr>
</tbody>
</table>

The performance analysis of the GA against the parameters is conducted for the PM-peak schedule optimizations of L25 and L52 using Scheme-2 (i.e., schedule optimization by shifting an existing schedule) and using Scheme-3 (i.e., schedule optimization by shifting an existing
schedule and adding holding times). The selection method is a fixed combination of rank selection and elitism. As noted earlier, the rank selection is more preferable than the roulette wheel selection, and elitism is adopted to ensure the convergence of the GA. Accounting for the recommendations by De Jong (Goldberg, 1989), CR of 0.6 and 0.8 are considered. Tested MR are 0.01, 0.03 and 0.05. PS of 10 and 30 are tested for Scheme-2. For Scheme-3, a population size of 50 is included instead of that of 30, accounting for the size of the solution space of Scheme-3. Therefore, a total of 12 combinations of parameters exist for the following four schedule optimizations: L25 using Scheme-2, L52 using Scheme-2, L25 using Scheme-3 and L52 using Scheme-3. The results are presented from FIGURE 3-30 to FIGURE 3-37.

FIGURE 3-30 and FIGURE 3-31 demonstrate the results of optimizing the PM-peak schedules of L25 using Scheme-2. Each chart of FIGURE 3-30 plots the solutions found by the GA against the iterations (generations) when PS=10, CR=0.6 and 0.8 and MR=0.01, 0.03 and 0.05. The GA is terminated at the 90\textsuperscript{th} iteration; therefore, the total number of the tested candidate solutions is 900. The GA is executed three times per each combination of parameters. The three numbers in each chart are the costs minimized by the GA at the last iteration. FIGURE 3-31 shows the results of the combinations of PS=30, CR=0.6 and 0.8 and MR=0.01, 0.03 and 0.05. The maximum iteration size is given as 30. In terms of the number of examined candidate solutions, therefore, the 30\textsuperscript{th} iteration of FIGURE 3-30 corresponds to the 10\textsuperscript{th} iteration of FIGURE 3-31; the 90\textsuperscript{th} iteration of FIGURE 3-30 corresponds to the 30\textsuperscript{th} iteration of FIGURE 3-31.

FIGURE 3-30, where the population size is 10, demonstrates that the higher crossover and mutation rates bring better results. The results with MR=0.05 are better than the others, because the cost curves of FIGURE 3-30 (e) and (f) converge earlier than the others. Furthermore, between the two figures (where MR=0.05), the GA with CR=0.08 performs better than that where CR=0.6. It should be noted that the GA with CR=0.6 and MR=0.01 (FIGURE 3-30 (a)) is much worse than the others: one of the three runs is beyond the boundary and cannot be seen in the figure. It results from the premature convergence caused by the small population size and the small crossover and mutation rates. This condition makes the population lose the diversity at the early stages of iterations, and the GA consequently fails to escape from the trap of a local minimum even with 90 iterations. In FIGURE 3-31 (a), where the population size increases to 30, the premature problem of the parameter set (i.e., CR=0.3 and MR=0.01) is relaxed. However, FIGURE 3-31 (a) is worse than the other GAs of FIGURE 3-31: the GAs with the higher CR and
MR show better performance as well. Higher population size does not show a significant improvement if crossover and mutation rates are high (compare FIGURE 3-30 (f) to FIGURE 3-31 (f)).

The same analysis is performed against the optimizations of the PM-peak schedules of L52 using Scheme-2. The results of L52 are presented in FIGURE 3-32 and 3-33, and bear a great similarity to the results of L25: higher crossover and mutation rates are better; the GA with CR=0.6 and MR=0.01 suffers from premature convergence when PS=10; and the increase of population size is not beneficial when the crossover and mutation rates are sufficiently high. It is not easy to choose a specific combination due to the small sample size. However, the identified patterns based on the above observations are used to determine the best combination. Accordingly, the parameter set of PS=10, CR=0.8 and MR=0.05 proves to be the best combination for optimizing a schedule using Scheme-2, as shown in the four figures.

While the figures show the convergence process in detail, TABLE 3-18 summarizes the average of the three convergence values at the last iteration according to the various GA parameters. The rank of each parameter combination is in the bracket (a set of parameters ranked in the first three places are accented in bold character). The selected parameter set is commonly the best for both L25 and L52, as highlighted in the table.
(a) Crossover=0.6 / Mutation=0.01
(b) Crossover=0.8 / Mutation=0.01
(c) Crossover=0.6 / Mutation=0.03
(d) Crossover=0.8 / Mutation=0.03
(e) Crossover=0.6 / Mutation=0.05
(f) Crossover=0.8 / Mutation=0.05

FIGURE 3-30 Scheme-2 of local transit route-25 (Population Size=10)
FIGURE 3-31 Scheme-2 of local transit route-25 (Population Size=30)
FIGURE 3-32 Scheme-2 of local transit route-52 (Population Size=10)
FIGURE 3-33 Scheme-2 of local transit route-52 (Population Size=30)
TABLE 3-18 Averages of three minimized costs at the last iteration (Scheme-2)

<table>
<thead>
<tr>
<th>Scheme-2</th>
<th>PS</th>
<th>MR</th>
<th>CR = 0.6</th>
<th>CR = 0.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>L25</td>
<td>10</td>
<td>0.01 FIGURE 3-30 (a) 5373.9 (Rank:12)</td>
<td>FIGURE 3-30 (b) 5099.6 (Rank:1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.03 FIGURE 3-30 (c) <strong>5099.6 (Rank:1)</strong></td>
<td>FIGURE 3-30 (d) 5106.9 (Rank:9)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.05 FIGURE 3-30 (e) <strong>5099.6 (Rank:1)</strong></td>
<td>FIGURE 3-30 (f) <strong>5099.6 (Rank:1)</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.01 FIGURE 3-31 (a) 5106.9 (Rank:9)</td>
<td>FIGURE 3-31 (b) 5106.9 (Rank:9)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.03 FIGURE 3-31 (c) <strong>5099.6 (Rank:1)</strong></td>
<td>FIGURE 3-31 (d) <strong>5099.6 (Rank:1)</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.05 FIGURE 3-31 (e) <strong>5099.6 (Rank:1)</strong></td>
<td>FIGURE 3-31 (f) <strong>5099.6 (Rank:1)</strong></td>
<td></td>
</tr>
<tr>
<td>L52</td>
<td>10</td>
<td>0.01 FIGURE 3-32 (a) 5674.2 (Rank:12)</td>
<td>FIGURE 3-32 (b) 5667.9 (Rank:11)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.03 FIGURE 3-32 (c) 5653.6 (Rank:7)</td>
<td>FIGURE 3-32 (d) <strong>5651.1 (Rank:1)</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.05 FIGURE 3-32 (e) <strong>5651.1 (Rank:1)</strong></td>
<td>FIGURE 3-32 (f) <strong>5651.1 (Rank:1)</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.01 FIGURE 3-33 (a) 5653.0 (Rank:4)</td>
<td>FIGURE 3-33 (b) 5654.7 (Rank:9)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.03 FIGURE 3-33 (c) 5655.4 (Rank:10)</td>
<td>FIGURE 3-33 (d) 5653.0 (Rank:4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.05 FIGURE 3-33 (e) 5653.9 (Rank:8)</td>
<td>FIGURE 3-33 (f) 5653.0 (Rank:4)</td>
<td></td>
</tr>
</tbody>
</table>

(Iteration size=90 for PS=10, and iteration size=30 for PS=30)

: Recommended parameters

From FIGURE 3-34 to FIGURE 3-37, the same performance analysis of the GA is conducted when revising the schedules of L25 and L52 using Scheme-3. The total number of examined candidate solutions is limited to 2,500. Therefore, the maximum iterations are 250 for a population size of 10 and 50 for a population size of 50. FIGURE 3-34 and 3-35 show the results of L25, and FIGURE 3-36 and 3-37 demonstrate the results of L52. Premature convergence also occurs with a GA having a small population size with low crossover and mutation rates, as shown in FIGURE 3-34 (a) and (b), and FIGURE 3-36 (a). Generally, a mutation rate of 0.03 resulted in a better solution compared to a mutation rate of 0.01 or 0.05. The increase of population size slows the convergence speed. Based on the observation of the figures, the GA with PS=10, CR=0.8 and MR=0.03 seems to be the best. TABLE 3-19, presenting the average of the three convergence values at the last iteration, also supports the selection of this parameter.
set. Its convergence speed is faster than that of the other GAs with other combinations of parameters, while it converges to an excellent solution at the end of iterations.

FIGURE 3-34 Scheme-3 of local transit route-25 (Population Size=10)
(a) Crossover=0.6 / Mutation=0.01
(b) Crossover=0.8 / Mutation=0.01
(c) Crossover=0.6 / Mutation=0.03
(d) Crossover=0.8 / Mutation=0.03
(e) Crossover=0.6 / Mutation=0.05
(f) Crossover=0.8 / Mutation=0.05

FIGURE 3-35 Scheme-3 of local transit route-25 (Population Size=50)
Population Size=10 / Crossover=0.6 / Mutation=0.01
Population Size=10 / Crossover=0.8 / Mutation=0.01
Population Size=10 / Crossover=0.6 / Mutation=0.03
Population Size=10 / Crossover=0.8 / Mutation=0.03
Population Size=10 / Crossover=0.6 / Mutation=0.05
Population Size=10 / Crossover=0.8 / Mutation=0.05

Run1: 5436.7  Run2: 5504.5  Run3: 5448.3
Run1: 5436.7  Run2: 5460.6  Run3: 5451.6
Run1: 5463.0  Run2: 5451.5  Run3: 5452.2
Run1: 5442.8  Run2: 5453.6  Run3: 5444.9
Run1: 5471.2  Run2: 5457.3  Run3: 5456.2
Run1: 5470.5  Run2: 5462.0  Run3: 5460.7

(a) Crossover=0.6 / Mutation=0.01  (b) Crossover=0.8 / Mutation=0.01
(c) Crossover=0.6 / Mutation=0.03  (d) Crossover=0.8 / Mutation=0.03
(e) Crossover=0.6 / Mutation=0.05  (f) Crossover=0.8 / Mutation=0.05

FIGURE 3-36 Scheme-3 of local transit route-52 (Population Size=10)
FIGURE 3-37 Scheme-3 of local transit route-52 (Population Size=50)
TABLE 3-19 Averages of three minimized costs at the last iteration (Scheme-3)

<table>
<thead>
<tr>
<th>PS</th>
<th>MR</th>
<th>CR = 0.6</th>
<th>CR = 0.8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.01</td>
<td>FIGURE 3-34 (a)</td>
<td>4872.5 (Rank: 12)</td>
</tr>
<tr>
<td></td>
<td>0.03</td>
<td>FIGURE 3-34 (c)</td>
<td>4834.0 (Rank: 7)</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>FIGURE 3-34 (e)</td>
<td>4829.9 (Rank: 4)</td>
</tr>
<tr>
<td>R25</td>
<td>0.01</td>
<td>FIGURE 3-35 (a)</td>
<td>4826.8 (Rank: 1)</td>
</tr>
<tr>
<td></td>
<td>0.03</td>
<td>FIGURE 3-35 (c)</td>
<td>4830.8 (Rank: 6)</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>FIGURE 3-35 (e)</td>
<td>4835.4 (Rank: 9)</td>
</tr>
<tr>
<td>Scheme-3</td>
<td>10</td>
<td>FIGURE 3-36 (a)</td>
<td>5463.5 (Rank: 9)</td>
</tr>
<tr>
<td></td>
<td>0.03</td>
<td>FIGURE 3-36 (c)</td>
<td>5455.6 (Rank: 7)</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>FIGURE 3-36 (e)</td>
<td>5461.6 (Rank: 8)</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>FIGURE 3-37 (a)</td>
<td>5452.9 (Rank: 6)</td>
</tr>
<tr>
<td></td>
<td>0.03</td>
<td>FIGURE 3-37 (c)</td>
<td>5449.9 (Rank: 4)</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>FIGURE 3-37 (e)</td>
<td>5478.6 (Rank: 12)</td>
</tr>
</tbody>
</table>

(Iteration size=250 for PS=10, and iteration size=50 for PS=50)

: Recommended parameters

The performance analysis of the GA shows the importance of carefully selecting the GA parameters. TABLE 3-20 presents the recommended GA parameters based on the analysis. Generally, the larger the population size, the greater the chance that a population will contain a candidate solution close to the optimal solution. However, the GA with a small population size can be more efficient than that with a large population size, taking into account the risk of premature convergence. Also, the appropriate combination of crossover and mutation rates depends on the given problem; a crossover rate of 0.8 is commonly recommended, but different values of mutations are recommended for Scheme-2 and Scheme-3.
TABLE 3-20 Recommended GA parameters

<table>
<thead>
<tr>
<th>GA Parameters</th>
<th>Schedule optimization using Scheme-2</th>
<th>Schedule optimization using Scheme-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection Method</td>
<td>Rank Selection + Elitism</td>
<td>Rank Selection + Elitism</td>
</tr>
<tr>
<td>Population Size (PS)</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Crossover Rate (CR)</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>Mutation Rate (MR)</td>
<td>0.05</td>
<td>0.03</td>
</tr>
</tbody>
</table>

6.5 Review of the Solution Found by GA

As noted earlier, GAs search for a solution close to the optimal solution of a given problem, rather than the optimal solution itself. Accordingly, it is essential to review the solutions found by the GA of this study. For Scheme-2, which has a relatively small solution space, an exhaustive sequential search can be performed. The solutions of the sequential search completely correspond to the solution found by the GA. The solution space of Scheme-3 is too large to handle through the sequential search; however, as seen in the analysis of a transfer shown in FIGURE 3-24, the found solutions are reasonably explained.

Nevertheless, some refinement is required to the solutions of Scheme-3. The review of several solutions by the GA shows strong consistency in the offset values of Scheme-3, but the holding times, though marginally different, do not dramatically change the total cost. The next figure shows how the inconsistency occurs with the holding times.
In FIGURE 3-38 (b), the holding times of L52-1 and L52-16 set by the GA (Raw solution by GA) are extreme examples of a dummy holding time. Those two receiving runs do not have any feeder. Consequently, their holding times have no influence on the cost of the objective function and do not result in any selection pressure on the GA. Also, the holding times determined by the GA for L52-2 and L52-14 are somewhat high.

Recall, however, that with regard to the property of the holding times shown in FIGURE 3-24 (d), the validity of the holding time depends on the transfer slack time; the output of the objective function will not be sensitive to the holding time, if the transfer slack time is sufficient. Although the holding times of L52-2 and L25-14 are reduced to 1 min., the total cost function is barely
affected (See SS=+5 in FIGURE 3-24 (d)). In contrast, even though the GA properly determines the critical holding times that have a strong impact on the cost function, such as L52-6, L52-10 and L52-11 in FIGURE 3-38 (b), the holding time of L52-12 is given as 2 min.; if it is increased to 3 min., the total cost decreases by approximately 4 pass.min., but this difference is negligible.

This review implies that the GA has enough ability to derive the overall framework of the revised schedule (i.e., the decision of the offset values and the critical holding times), but is not sophisticated enough to tune the holding times, since GAs are basically a global search without using gradient information. Therefore, a refinement of holding times may be required by reviewing the results. The refined solutions of FIGURE 3-38 are set by applying the following rules iteratively:

- if the increase of the total cost by reducing the holding time by 1 min. is less than 1 pass.min., decrease the holding time by 1 min.; and
- if the decrease of the total cost by raising the holding time by 1 min. is larger than 1 pass.min., increase the holding time by 1 min.

7 Summary

This chapter focuses on coordinating the inter-modal transfer from commuter transit services to local transit services at a suburban transfer point. It attempts to minimize the transfer associated costs through modifying the existing schedule of the local transit service. Accounting for the direction of transfers, AM-peak transfer is formulated differently from PM-peak transfer. For the provision of better transfers, the developed optimization model allows for changing the scheduled headway as well as shifting the whole schedule. Unlike other studies that focused on producing the optimal schedules only, this study simultaneously considers an additional dispatching policy that further reduces transfer related costs. The models explicitly incorporate the variability of transit vehicle arrivals, assumed to follow a Lognormal distribution.

The case study, which is based on actual schedule data, indicates that the developed optimization model is promising and favourable to improving timetables in terms of transfer related costs. The GA-based solution approach demonstrates sufficient ability in solving the given optimization
problem. This case study also shows that the model provides schedulers of transit services with various alternatives to revise their existing schedules. Although the design of the optimization schemes must be dependent on the decision and the policy of transit agencies, this flexibility of the model is very beneficial, facilitating the development of a timetable that maximizes convenience for transit users and adheres to their policy simultaneously.
Chapter 4
Real-Time Coordination Strategy

1 Introduction

Transit agencies need an operational control method to protect the coordinated transfers against unexpected delays of transit vehicles at a transfer stop, even though they may have well-coordinated timetables. As shown in FIGURE 4-1, a transit run can be delayed beyond its historical pattern due to several reasons, such as road construction or maintenance, traffic accidents and so on. This unexpected delay may break the planned coordination between two runs. One possible approach is real-time Connection Protection (CP) through holding strategies, accounting for these possible daily variances of transit operations.

![FIGURE 4-1 Broken coordination by unexpected delay of feeder run](image)

Generally, the CP problem involves the transfer from one transit route (feeder route) to another transit route (receiving route) where the scheduled coordinated transfer has been broken. The delay of a feeder transit run, beyond the scheduled departure time of a receiving transit run, would cause failure of the coordinated transfer. A CP control strategy would hold the receiving run for the transfer passengers. The optimal CP holding control is basically a trade-off between benefits and costs of applying the holding control in terms of passenger waiting times and transit operational costs. The holding control would cause some passengers to experience reduced waiting times, while lengthening the waiting and travel times of other passengers and bringing
additional expenditures to the transit agency (e.g., labour, fuel costs, etc.). The time-space diagram in FIGURE 4-2 shows how holding control can affect the waiting times.

FIGURE 4-2 Influence of CP control on passengers’ waiting times

Suppose that a transit run $R$ on a transit route is scheduled to pick up passengers who are transferring from a transit run $d$ on another transit route at a transfer stop. Due to the delay of run $d$, $R$ is held. This holding control will affect the schedule adherence of run $R$ at the downstream stops unless it recovers that delay (the dotted line of the figure shows the affected bus trajectory due to holding). The holding control reduces the waiting times of the transfer passengers. However, it delays the in-vehicle passengers who are already aboard run $R$, and results in delays for those passengers waiting for the run $R$ at the downstream stops. Additionally, the holding
control can result in a by-product for the passengers who are supposed to get on run $R+1$ (the next run of run $R$) at the transfer point. Since some of them can take run $R$, the average waiting time of those who were originally waiting for run $R+1$ is reduced. This is also propagated downstream depending on whether run $R$ can compensate for the delay due to holding. The key point to implement a model for the CP control is to evaluate the described costs and benefits so as to decide how long run $R$ should wait. The CP control would be beneficial when the number of transfer passengers is large and the headway of the receiving route is long.

This figure also indicates how prediction modules are critical for effective CP control in real-time. In the figure, the right side of the vertical line indicates future events (to be predicted). At minimum, we need to predict when run $d$ will arrive, how many passengers will transfer to run $R$, how many passengers are currently in run $R$, and when the next run, run $R+1$, will be available if the transfer passengers miss run $R$. Additionally, in order to evaluate how a certain holding time increases the waiting times of the downstream passengers, the demands of the passengers and the travel times of run $R$ along the downstream stops should be provided. FIGURE 4-3 summarizes the input for the CP control model that should be provided or predicted.

The CP control model of this study assumes an external prediction module providing the input, making use of AVL (Automatic Vehicle Location) and APC (Automatic Passenger Counting) systems. Additionally, the following assumptions are made:

- If the transfer passengers miss their target, run $R$, they will wait for the next available run, run $R+1$.
- The CP control does not cause any additional costs for a transit agency.
- The influence of the CP control on the passengers for run $R+1$ is ignored.
- The transit vehicle capacity is unconstrained.

Some transfer passengers may decide not to wait, but will try to find an alternative mode to continue their trip if they miss the connecting run. As noted in the previous chapter, an additional model would be required in order to estimate how they choose the alternative modes and what
those costs are, but that is beyond the scope of this study. Nevertheless, a separate study on the passengers’ behaviour, i.e. choosing the alternative mode and related cost, could be incorporated in the CP model.

This study excludes the transit operational cost and the waiting times of passengers served by run R+1. The difference in attributes between the operational cost and the waiting times requires conversion factors to unify their units. However, the decision about the factors themselves is a big topic, and the cost can be added to the CP model quite easily with the defined topic. As illustrated in Figure 4-2, the holding of run R possibly changes the waiting times of customers of run R+1. Some passengers who are supposed to get on run R+1 can take run R because of the CP control, thereby experiencing a reduced waiting time (i.e., passengers numbered as 4 and 5 in Figure 4-2). However, the influence of the CP control on these passengers is ignored, because this is a secondary effect, which is not critical to the given problem.

This study does not consider the capacity constraint of a transit unit. There is certainly the possibility that, even though run R is held by the CP control, some transfer passengers from the delayed feeder cannot get on run R due to its limited capacity. In a future study, therefore, the additional waiting time caused by the capacity issue should be incorporated into the CP model.

It is difficult to handle the complexity of the model for a real-time application, as the number of delayed feeder runs increases. Although this study tries to formulate a CP model in a more sophisticated way than previous models, we limit the number of delayed feeder runs here to two. In a real application, the simultaneous delays of multiple feeders (three or more) would be uncommon.

The next section will formulate and analyze the CP model to apply the holding control to a receiving run in order to protect the scheduled connection against the delay of one feeder run. Section 3 deals with the case of two delayed feeder runs. Section 4 reviews the difference between the CP model of this study and the previous studies in detail. Finally, Section 5 presents concluding remarks on the analysis.
Delayed Feeder Transit Runs ($d=1,2$)
At Transfer Point
* Predicted Arrival Time
* Predicted Number of Transfer Passengers

Receiving Transit Run $R$
At Transfer Point
* Number of In-Vehicle Passengers

For Downstream Segment
* Predicted Downstream Travel Times of Run $R$ (from Stop $k$ to Stop $k+1$)
* Predicted Downstream Passenger Demands of Run $R$ (at each stop)
  - Arrival Rate
    (Unaware Passengers who arrive randomly)
  - Number of Passengers
    (Aware Passengers who arrive in reference to the transit schedule)

Run $R+1$, (Next Run of Run $R$)
At Transfer Point
* Predicted Departure Time (or Scheduled Departure Time)

Run $R-1$, (Previous Run of Run $R$)
For Downstream Segment
* Actual Departure Times (or Predicted Departure Times)
  (Required in order to estimate the number of downstream passengers
   using the arrival rate of the unaware passengers.)

FIGURE 4-3 Required input for real-time CP control
2 Connection Protection (CP) Model: Delay of One Feeder Run

2.1 Formulation of Model

Suppose that a transit run, run \( R \) (receiving run), is ready to be dispatched, but another run, run \( d=1 \) (feeder run), is late and has not arrived at a transfer point. As assumed, at this moment, the predicted arrival time of the delayed feeder, \( E(A\hat{T}_1) \), is available from an external prediction model based on real-time observations using AVL and APC systems. If run \( R \) is held, there will be three groups of passengers affected by the control as described earlier: 1) the transfer passengers who transfer from run 1 to run \( R \); 2) the in-vehicle passengers who are already in run \( R \); and 3) the downstream passengers who are waiting for run \( R \) at the downstream stops. It is assumed that the holding of run \( R \) does not cause any additional costs for the transit agency. The total cost of all the passengers related to the CP control can be expressed in the following equation:

\[
E(T_{R^{Total}}) = E(T_{1,R^{Tr}}) + E(T_{R^{InV}}) + E(T_{R^{Dwn}}) \] (EQ-4.1)

Where:

\( E(T_{R^{Total}}) \) : Expected total cost of all passengers related to the CP control

\( E(T_{1,R^{Tr}}) \) : Expected transfer cost of transfer passengers from run \( d \) to run \( R \)

\( E(T_{R^{InV}}) \) : Expected waiting cost of in-vehicle passengers in run \( R \)

\( E(T_{R^{Dwn}}) \): Expected waiting cost of downstream passengers waiting for run \( R \) at the downstream stops of the transfer point

If run 1 arrives as predicted (i.e., it arrives exactly at \( E(A\hat{T}_1) \)), the holding time, \( HT_R \), can be decided simply based on the next rule:
IF \( E(T_R^{Total} | HT_R = 0) < E(T_R^{Total} | HT_R = E(A\hat{T}) - SDT_R) \)

\[ HT_R = 0 \]

ELSE

\[ HT_R = E(A\hat{T}) - SDT_R \]

Where:

- \( HT_R \): Holding time of run \( R \)
- \( SDT_R \): Scheduled departure time of connecting run \( R \) at a transfer point
- \( E(A\hat{T}) \): Predicted arrival time of delayed feeder run

However, not only are the transit arrival and departure processes stochastic in the real world, but the predicted arrival times by any prediction model are also random variables having some probability distribution. Of course, the delayed run does not always arrive exactly at the predicted time. As shown in FIGURE 4-4, there exists some risk of failing the connection (shaded area), even while holding run \( R \) beyond \( E(A\hat{T}) \).

![FIGURE 4-4 Randomness of predicted arrival time of feeder run](image-url)
Generally, prediction models assume error terms following Normal distributions with zero mean and some amount of variance (i.e., white noise). The assumption is commonly adopted in building prediction models based on regression methods and time-series approaches (e.g., autoregressive model, moving average model and autoregressive integrated moving average model). Accordingly, the predicted arrival times can be written in the next equation.

\[ A\hat{T}_i = E(\hat{A}T_i) + \varepsilon_{\hat{A}T_i} \]  

(EQ. 4-2)

Where:

\( \varepsilon_X \) : Prediction error in predicted value \( X \), i.e., white noise ~ Normal \( (0, \sigma_X^2) \)

\( \sigma_X \) : Standard deviation of \( \varepsilon_X \)

Therefore, the probability density function of the random variable \( A\hat{T}_i \) follows a Normal distribution, shown as:

\[ A\hat{T}_i \sim \text{Normal} \left( E(\hat{A}T_i), \sigma_{\hat{A}T_i}^2 \right) \]  

(EQ. 4-3)

Considering the error of the predicted value, the cost of the transfer passengers is defined by the following equations:

\[ E(T_{1R}^{Tr}) = E(T_{1R}^{TrS}) + E(T_{1R}^{TrM}) \]  

(EQ 4-4)
\[ E(T_{1,R}^{TS}) = P_{1,R}^{Tr} \times \int_{SDT_R}^{SDT_R + HT_R}(\Delta_{1,R}^{Tr})f_{\hat{A}_{T_1}}(x)dx \]  
(EQ 4-5)

\[ E(T_{1,R}^{TM}) = P_{1,R}^{Tr} \times \int_{SDT_R}^{\infty}(SDT_{R+1} - x)f_{\hat{A}_{T_1}}(x)dx \]  
(EQ 4-6)

\[ \Delta_{d,R}^{Tr} = \Delta \times P_{1,R}^{Tr} \]  
(EQ 4-7)

Where:

- \( E(T_{1,R}^{TS}) \) : Expected successful connection cost of transfer passengers from run 1 to run \( R \)
- \( E(T_{1,R}^{TM}) \) : Expected missed connection cost of transfer passengers from run 1 to run \( R \)
- \( P_{1,R}^{Tr} \) : Number of transfer passengers from run 1 to run \( R \)
- \( f_{\hat{A}_{T_1}} \) : Probability density function of \( \hat{A}_{T_1} \)
- \( SDT_{R+1} \) : Scheduled departure time of the next available receiving run, run \( R+1 \)
- \( \Delta_{1,R}^{Tr} \) : Boarding time of the transfer passengers from run 1 to run \( R \)
- \( \Delta \) : Boarding time per one passenger (3.5 sec. \( \approx 0.0583 \) min.)

\( E(T_{1,R}^{Tr}) \) consists of successful connection cost, \( E(T_{1,R}^{TS}) \), and missed connection cost, \( E(T_{1,R}^{TM}) \), as shown in Equation 4-4. The former cost represents the expected waiting time passengers experience if the delayed run 1 reaches run \( R \) by holding run \( R \). The latter cost is their expected waiting time if they miss run \( R \) and take the next available run. If the connection is successful (represented by the blank area of FIGURE 4-4), the transfer cost will be the total boarding time of the transfer passengers (\( \Delta_{1,R}^{Tr} \)) as formulated in Equation 4-5. Note that run \( R \) does not remain at the stop any longer than necessary, but departs immediately for the next stop as soon as the passengers finish boarding. \( \Delta_{1,R}^{Tr} \), experienced by each transferring passenger, is obtained from
Equation 4-7. If run 1 is later than the holding time of run $R$, the waiting time of the transferring passengers will be the difference between the arrival time of run 1 and the scheduled departure time of the next run (run $R+1$). Accordingly, the expected cost will be defined as Equation 4-6.

The in-vehicle passengers already in run $R$ would be delayed if it was held, and the delay will depend on its expected departure time. Their cost will be:

$$E \left( T_{inv}^R \right) = P_{inv}^R \times (E (DT^R) - SDT^R) \quad \text{(EQ. 4-8)}$$

$$E (DT^R) = \int_{SDT^R}^{\infty} (x + \Delta_{1,R}^{Tr}) f_{AT_1} (x) dx + \int_{SDT^R + HT^R}^{\infty} (SDT^R + HT^R) f_{AT_1} (x) dx \quad \text{(EQ. 4-9)}$$

Where:

- $P_{inv}^R$ : Number of in-vehicle passengers in run $R$ at $SDT^R$
- $E (DT^R)$ : Expected departure time of run $R$

Equation 4-9 defines the expected departure time of run $R$. The first term of the right side of the equation implies that it departs upon having processed the transfer passengers. The second term indicates that if run 1 has not arrived before the holding time of run $R$, then run $R$ is released. The term in the brackets in Equation 4-8 represents the expected departure delay of run $R$ due to the holding control.

The holding of run $R$ causes the downstream passengers’ waiting time to increase, since it makes run $R$ later than its original schedule. Accordingly, $E (T_{down}^R)$ is

$$E \left( T_{down}^R \right) = \text{fun}_{dus} \left( \lfloor DT^R \rfloor HT_R \right) - \text{fun}_{dus} \left( \lfloor DT^R \rfloor HT_R = 0 \right) \quad \text{(EQ. 4-10)}$$
The downstream model returns the expected downstream passengers' waiting time according to the departure time of run \( R \). The first term of the right side of the equation is the expected value when the run is held (\( R_{HT} > 0 \)), and the second term is the expected value when the run departs according to the given schedule (\( R_{HT} = 0 \)). Hence, \( E(D_{RT}) \) reflects the additional waiting time caused by the holding control and will be zero when the run departs at the scheduled departure time. The CP model also uses the downstream model explained in Section 4 of Chapter 3.

It should be noted that \( D_{RT} \) is also a random variable whose distribution is unknown; it is possible to estimate its expected value (Equation 4-9) and variance (Equation 4-11 and 4-12).

The downstream model makes use of the two parameters in approximating the probability density function of the arrival time at the next stop.

\[
E(D_{RT}) = E(D_{RT}^2) - E(D_{RT})
\]

\[
\text{Var}(D_{RT}) = E(D_{RT}^2) - E(D_{RT})^2
\]

The cost function of the CP model is the sum of the cost components as defined in Equation 4-1, and the solution of the problem is the holding time minimizing the cost function:

\[
HT^* = \{ t : \min \{ E(T_{Tr}^R | HT = t) \} \text{ for } 0 \leq t \}
\]

Where:

- \( HT^* \): Optimal holding time

\[
E(T_{Tr}^R | HT = t) = \int (x + \bar{x}) f_{\bar{x}}(x) dx + \int (SDT + HT^*) f_{\bar{x}}(x) dx
\]

(EQ. 4-11)

\[
\text{Var}(D_{RT}) = \int (x + \bar{x})^2 f_{\bar{x}}(x) dx - \bar{x}^2
\]

(EQ. 4-12)

Section 4 of Chapter 3.
2.2 Analysis of Model

2.2.1 Testbed

As a testbed to analyze the CP model, two transit services in the City of Brampton are considered as in the previous chapter: a local bus route of the city and a commuter transit service (GO-Transit) linking the city with downtown Toronto. The commuter transit serves commuter travel between Brampton and the CBD of Toronto and the local transit line provides the residents of Brampton with access to/from the commuter line. The CP control is applied to the afternoon peak period transfer from the commuter line to the local bus line at Brampton City Centre, a transfer point where the two transit services meet.

As shown in FIGURE 4-5, it is assumed that the local bus line (i.e., receiving run, run $R$) is scheduled to depart from the transfer stop at 3:00PM, and the commuter transit run (i.e., feeder run, run 1) is scheduled to arrive earlier but got delayed unexpectedly. The local transit run, belonging to Route-25 of the Brampton local transit, travels eight downstream stops and the service headway is 30 min. The downstream passengers’ cost incurred by the CP control is restricted to the downstream passengers of the eight stops, and other influences of the CP control on other local transit routes are ignored.
The predicted arrival time of the late feeder is provided by a prediction model that takes input from AVL and APC systems. The distribution of the travel time of the delayed feeder can be defined with respect to the distance to the transfer stop as shown in FIGURE 4-6. When making
holding decisions, hence, it is possible to estimate the mean and the standard deviation of the travel time using the current location of the delayed feeder available from the AVL system. The mean and standard deviation of the predicted travel time will increase as the distance from the transfer stop increases, as shown in the figure. If the holding decision is made at the scheduled departure time of the receiving run, the error of the predicted arrival time, $\sigma_{A_{f_t}}$, can be expressed by the following equation:

$$
\sigma_{A_{f_t}} = c_v \times \{ E(A_{f_t}) - SDT_R \} \quad \text{(EQ. 4-14)}
$$

Where:

$c_v$ : Coefficient of variation (the ratio of standard deviation to mean)

$E(A_{f_t}) - SDT_R$ represents the mean of the travel time of the delayed feeder to the transfer stop at the decision time. $c_v$ is the coefficient of variation measuring the relative variance with respect to the mean. Accordingly, a higher coefficient of variation is associated with a higher variability in the predicted arrival time. The study of Rajbhandari showed that $c_v$ ranged from 0.1 to 0.3 based on data collected in New Jersey, U.S. Its range in the non-CBD is between 0.1 and 0.17, while it rises to 0.3 for the CBD area (Rajbhandari, 2005).

For simplification, the following is assumed: walking time required to transfer between two routes is zero; and the previous and following runs of the receiving run operate as scheduled. If the transfer requires significant walking time, however, the predicted arrival time of the feeder run can be shifted by the walking time. The arrival and departure times of the previous and following runs of the receiving run can be provided by a prediction model as well.

TABLE 4-1 presents the travel times and the scheduled departure times of run $R$ along the downstream stops, and the arrival rates of downstream passengers at each stop. It is assumed that the travel times are predicted with $c_v$ of 0.15 and the prediction errors are normally distributed
like the predicted arrival time of run 1. As noted earlier, the downstream model of Section 4 of Chapter 3 is compatible with the Normally distributed travel times.

### TABLE 4-1 Input parameters for downstream stops

<table>
<thead>
<tr>
<th>STOP</th>
<th>Scheduled departure time</th>
<th>Downstream Travel Time (Normally distributed)</th>
<th>Downstream Passengers</th>
<th>Alighting Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean (min) Std. dev. (min)</td>
<td>Unaware Pass. (arr. rate: pass/min)</td>
<td>Aware Pass. (arr. rate: pass/min)</td>
</tr>
<tr>
<td>Tr. Point</td>
<td>3:00PM</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>3:08PM</td>
<td>7.2</td>
<td>7.2×0.15</td>
<td>0.1</td>
</tr>
<tr>
<td>2</td>
<td>3:14PM</td>
<td>5.4</td>
<td>5.4×0.15</td>
<td>0.1</td>
</tr>
<tr>
<td>3</td>
<td>3:19PM</td>
<td>4.5</td>
<td>4.5×0.15</td>
<td>0.1</td>
</tr>
<tr>
<td>4</td>
<td>3:25PM</td>
<td>5.4</td>
<td>5.4×0.15</td>
<td>0.1</td>
</tr>
<tr>
<td>5</td>
<td>3:30PM</td>
<td>4.5</td>
<td>4.5×0.15</td>
<td>0.1</td>
</tr>
<tr>
<td>6</td>
<td>3:36PM</td>
<td>5.4</td>
<td>5.4×0.15</td>
<td>0.1</td>
</tr>
<tr>
<td>7</td>
<td>3:42PM</td>
<td>5.4</td>
<td>5.4×0.15</td>
<td>0.1</td>
</tr>
<tr>
<td>8 (Tr.)</td>
<td>3:50PM</td>
<td>7.2</td>
<td>7.2×0.15</td>
<td>0.1</td>
</tr>
</tbody>
</table>

2.2.2 Sensitivity Analysis

This subsection includes a sensitivity analysis to explore the relationship between the input parameters and the CP holding strategy.

FIGURE 4-7 shows the cost functions according to predicted arrival times of the feeder with \( c_s \) of 0.1. The total cost increases with short holding times. After reaching some peak, the total cost functions drop as the probability of a successful connection rises with increasing holding time. Then, the costs converge, as a result of early dispatching; this convergence is an important feature of the cost function of this study. In order to provide insights on the relationship between the total cost function and the cost components over different holding times, FIGURE 4-8 (a) shows the total cost function \( E(T_{R\text{Total}}) \) with the cost components of the total cost function, namely the transfer passenger cost \( E(T_{1,R}) = E(T_{1,R}^{TrS}) + E(T_{1,R}^{TrM}) \), the in-vehicle passenger cost
\( E (T^V_R) \), and the downstream passenger cost \( E (T^D_R) \), when \( E (A_T) = 3:02:30 \). FIGURE 4-8 (b) shows the departure delay of run \( R \) due to holding control, \( E (D_T_R) - SDT_R \).

\[
P^T_{\infty} = 10, \quad P^V_R = 35, \quad SDT_R = 3:00, \quad SDT_{R+1} = 3:30
\]

FIGURE 4-7 Variation of cost function against different \( E (A_T) \)'s \( (c_v = 0.1) \)
FIGURE 4-8 Cost function and its cost components
The total cost function is partitioned into three segments for better explanation as seen in FIGURE 4-8 (a). In Seg-1, the total cost function, \( E(T^\text{Total}_R) \), increases as the holding time increases because the probability of a successful connection is still low. Both \( E(T^\text{InV}_R) \) and \( E(T^\text{Down}_R) \) increase. \( E(T^\text{Tr}_{1,R}) \) is much higher than both of them and is decreasing very slightly as the holding time increases. Within this segment, the penalty due to holding, \( E(T^\text{InV}_R) + E(T^\text{Down}_R) \), is much higher than its benefit, \( E(T^\text{Tr}_{1,R}) \).

In Seg-2, it can be seen that, as the holding time is increased, \( E(T^\text{Total}_R) \) decreases, which is an effect caused largely by the rapid reduction of \( E(T^\text{Tr}_{1,R}) \) or more specifically the missed connection cost incurred by the transfer passengers.

In Seg-3, all functions start to converge to constant values due to sufficiently long holding times. \( E(T^\text{Tr}_{1,R}) \) becomes approximately \( P^\text{Tr}_{1,R} \times \Delta^\text{Tr}_{1,R} \), since the missed connection cost will be close to zero while the successful connection cost will be nearly \( P^\text{Tr}_{1,R} \times \Delta^\text{Tr}_{1,R} \), because the probability of a successful connection is almost 1.0. It should be noted that \( E(T^\text{InV}_R) \), and \( E(T^\text{Down}_R) \) do not increase proportional to the holding time, since those two times depend mainly on the departure delay of run \( R \). As seen in FIGURE 4-8 (b), the departure delay converges to a constant value indicating that the holding times are sufficiently long to ensure early arrival of the feeder, hence early dispatching of run \( R \). In other words, setting a holding time of say 5 minutes has the same associated cost as a holding time of 8 minutes because in both scenarios the feeder will most likely arrive sufficiently earlier than those holding times and the local bus will be dispatched immediately after the transfer passengers board the local bus (due to the early dispatching policy).

FIGURE 4-8 (c) provides further examples of the cost function using different input values. In FIGURE 4-8 (c), while all inputs remain similar to the previous example, the expected arrival time of the feeder is set to be later than the corresponding value of FIGURE 4-8 (a). As shown, the functions behave in a similar pattern to that of FIGURE 4-8 (a), with the difference being the point of convergence (i.e. first point of guaranteed successful connection) is further delayed. In this example, the cost of guaranteed connection protection is higher than the cost of no control (i.e. with holding time of 0), while for the case of FIGURE 4-8 (a) the opposite is observed.
Accordingly, FIGURE 4-7 shows that if the predicted delay of the feeder run is small, it is more beneficial to hold the receiving run, as the total cost of holding at the first point of convergence is smaller than the total cost of no holding (i.e., holding time of 0). As the expected delay increases, however, the benefit begins to decrease until a point beyond which it becomes more costly to hold the receiving run. This is clear in the case of the predicted arrival time of 3:03:30, where the optimal strategy is to let the receiving run depart without holding.

FIGURE 4-9 presents the relationship between the expected benefit of holding and the error (standard deviation) of the predicted arrival time of the delayed feeder run. It presents the variation of the total cost function when $c_v = 0.2$. That is, the prediction errors of FIGURE 4-9 are twice as large as those of FIGURE 4-7. The higher the error of predicting the arrival time, the more dispersed the cost functions are. That is, the connecting run should be held longer in order to obtain the benefit from the holding control when the arrival time is predicted with a large error.

\[
P_{1,R}^{TR} = 10, \quad P_{R}^{Inv} = 35, \quad SDT_R = 3:00, \quad SDT_{R+1} = 3:30
\]

FIGURE 4-9 Variation of cost function against different $E(\hat{A}_1)$'s ($c_v = 0.2$)
The total cost function, the holding decision, and the expected benefit are obviously influenced by other input parameters. Those include the number of transfer passengers \( P_{i,R}^{Tr} \), the number of passengers already on board the receiving run \( P_{R}^{inv} \), the demand of passengers waiting for the receiving runs at the downstream stops, and the time at which the next run is available, in the event transfer passengers miss the receiving run \( SDT_{R+1} \). A comprehensive sensitivity analysis against those parameters is presented from FIGURE 4-10 to FIGURE 4-13. It shows that the holding control is beneficial if:

- the number of the transfer passengers is high,
- the number of the in-vehicle passengers is low,
- the demand of downstream passengers is low, and
- the time at which the next run is available is long.

![Graph showing variation of cost functions against different \( P_{i,R}^{Tr} \)’s](image)

\[ E(\hat{AT}_i)=3:02:30, \quad c_v = 0.2, \quad P_{R}^{inv} = 35, \quad SDT_R = 3:00, \quad SDT_{R+1} = 3:30 \]

FIGURE 4-10 Variation of cost functions against different \( P_{i,R}^{Tr} \)’s
$E( \hat{A}_1 ) = 3:02:30, \ c_v = 0.2, \ P_{1,R}^{Tr} = 5, \ SDT_R = 3:00, \ SDT_{R+1} = 3:30$

**FIGURE 4-11** Variation of cost functions against different $P_R^{InV}$’s

$E( \hat{A}_1 ) = 3:02:30, \ c_v = 0.2, \ P_{1,R}^{Tr} = 10, \ P_R^{InV} = 35, \ SDT_R = 3:00, \ SDT_{R+1} = 3:30$

(‘x %’ represents the percentage of the downstream passengers’ demand level in TABLE 4-1)

**FIGURE 4-12** Variation of cost functions against different demand levels of downstream stops
2.2.3 Decision of Optimal Holding Time

Because of the complexities associated with the underlying stochastic processes, it is difficult to find an analytical solution to the models formulated in Section 2.1. The simplest solution approach would be an exhaustive search that tests all the feasible solutions through discretizing the solution space. However, an efficient solution algorithm is necessary for the real-time application of the CP control strategy.

The shapes of the cost functions shown in the sensitivity analysis provide an important clue to find the optimal holding time of the CP problem that minimizes the cost function. The total cost does not have a specific minimum point, but has one peak and then converges as the probability of a successful connection rises sufficiently (i.e., the probability that the delayed feeder run arrives at the transfer point before the holding time expires is close to 1). When the cost of the holding control is much higher than its benefit, alternatively, the cost function rises continuously and converges exhibiting no peaking. Therefore, the decision of the holding control depends on whether the cost of the converged point is lower than the cost of no control or not: an optimal holding time, $HT^*_R$, will exist at the two boundary points, namely, the holding time of 0 or the holding time of the converged point. It is possible to define the converged point as

$$E(A\hat{T}_1) = 3:02:30, \ c_v = 0.2, \ P_{1,R}^{TV} = 10, \ P_{R}^{TV} = 35, \ SDT^*_R = 3:00$$

FIGURE 4-13 Variation of cost functions against different $SDT^*_{R+1}$'s
$E(\hat{A}_{T_i}) + 3 \times \sigma_{\hat{A}_{T_i}}$, since the predicted arrival time is normally distributed and the probability that the delayed run arrives earlier than the 3-sigma point is over 0.997. As such, the following procedure is proposed in order to decide the optimal holding time.

\[
\begin{align*}
\text{IF } E(T_{R|HT=0}) > E(T_{R|HT=E(\hat{A}_{T_i})+3\times \sigma_{\hat{A}_{T_i}}}) \\
\text{HT}_{R}^* = \{ E(\hat{A}_{T_i}) + 3 \times \sigma_{\hat{A}_{T_i}} \} - SDT_{R}
\end{align*}
\]

\[
\begin{align*}
\text{ELSE } \\
\text{HT}_{R}^* = 0
\end{align*}
\]

FIGURE 4-14 Rule to find the optimal holding time (delay of one feeder run)

The above rule decides if run $R$ should wait for the delayed run by comparing the total cost of no holding with that of holding until the converged point: if the cost of no holding is larger than that of holding, run $R$ is held until the 3-sigma point; and if not, it departs without holding. If the range of the holding time is restricted for some reason, due, for instance, to a service standard of the transit agency, the 3-sigma point of the above procedure can be replaced by the maximum holding time allowed.

3 Connection Protection (CP) Model: Delay of Two Feeder Runs

3.1 Formulation of Model

When two feeder runs (run $d=1$ and run $d=2$) from two feeder routes are late, the total cost function of the CP model is written as:

\[
E(T_{R|Total}) = E(T_{1,R}^{Tr}) + E(T_{2,R}^{Tr}) + E(T_{R}^{Inv}) + E(T_{R}^{Dyn})
\]  
(EQ. 4-15)
The connection costs of the transfer passengers from run 1 and run 2 ($T_{1,R}^{Tr}$ and $T_{2,R}^{Tr}$, respectively) and the departure time of the receiving run $R$ ($DT_R$) are defined in the following equations:

$$T_{1,R}^{Tr} = \begin{cases} 
\Delta_{1,R}^{Tr} + \Delta_{2,R}^{Tr} & \text{for Case 1} \\
(A\hat{T}_2 - A\hat{T}_1) + \Delta_{1,R}^{Tr} + \Delta_{2,R}^{Tr} & \text{for Case 1'} \\
(A\hat{T}_2 - A\hat{T}_1) + \Delta_{2,R}^{Tr} & \text{for Case 2} \\
\Delta_{1,R}^{Tr} & \text{for Case 2'} \\
SDT_R + HT_R - A\hat{T}_1 & \text{for Case 3} \\
SDT_{R+1} - A\hat{T}_1 & \text{for Case 3'} \\
\Delta_{1,R}^{Tr} & \text{for Case 4} \\
SDT_{R+1} - A\hat{T}_1 & \text{for Case 4'} \\
SDT_{R+1} - A\hat{T}_1 & \text{for Case 5} 
\end{cases}$$

(EQ. 4-16)

$$T_{2,R}^{Tr} = \begin{cases} 
(A\hat{T}_1 - A\hat{T}_2) + \Delta_{1,R}^{Tr} + \Delta_{2,R}^{Tr} & \text{for Case 1} \\
\Delta_{1,R}^{Tr} + \Delta_{2,R}^{Tr} & \text{for Case 1'} \\
\Delta_{2,R}^{Tr} & \text{for Case 2} \\
(A\hat{T}_1 - A\hat{T}_2) + \Delta_{1,R}^{Tr} & \text{for Case 2'} \\
SDT_{R+1} - A\hat{T}_2 & \text{for Case 3} \\
SDT_R + HT_R - A\hat{T}_2 & \text{for Case 3'} \\
SDT_{R+1} - A\hat{T}_2 & \text{for Case 4} \\
\Delta_{2,R}^{Tr} & \text{for Case 4'} \\
SDT_{R+1} - A\hat{T}_2 & \text{for Case 5} 
\end{cases}$$

(EQ. 4-17)
As shown in FIGURE 4-15, nine domains are defined according to the combination of the related variables. The X-axis and Y-axis represent the predicted arrival times of run 1 and run 2, respectively. The applied assumptions are the same as those of the single run delay.
For Case 1 through Case 4, run 1 successfully connects with run R because it arrives at the transfer point earlier than the given holding time while being ahead of run 2. Run 2 is also connected to run R successfully for the first two cases, but fails to connect with the receiving run for the last two cases. In Case 1, run 2 comes to the stop while run R is loading the passengers of run 1. Therefore, the departure time of run R is $A\hat{T}_1 + \Delta_{1,R}^{Tr} + \Delta_{2,R}^{Tr}$, i.e., once the passengers of run 2 have completely boarded run R. The transfer times of passengers from run 1 and run 2 are, respectively, $\Delta_{1,R}^{Tr} + \Delta_{2,R}^{Tr} (= DT_R - A\hat{T}_1)$ and $(A\hat{T}_1 + \Delta_{1,R}^{Tr}) - A\hat{T}_2) + \Delta_{2,R}^{Tr} (= DT_R - A\hat{T}_2)$. Since run 2 arrives at the stop after the passengers from run 1 finish boarding in Case 2, run R departs at $A\hat{T}_2 + \Delta_{2,R}^{Tr}$. $T_{1,R}^{Tr}$ and $T_{2,R}^{Tr}$ are $A\hat{T}_2 - A\hat{T}_1 + \Delta_{2,R}^{Tr}$ and $\Delta_{2,R}^{Tr}$, respectively. Case 3 and Case 4 indicate that run 2 is later than the holding time.

In Case 3, run R completes the boarding of the passengers of run 1 and waits for run 2 until the given holding time. Accordingly, its departure time will be $SDT_R + HT_R$ and the waiting time for run 1 passengers is expressed as $SDT_R + HT_R - A\hat{T}_1$. Case 4 represents the departure of run R as soon as it processes the passengers because it cannot wait any longer ($A\hat{T}_1 + \Delta_{1,R}^{Tr} > SDT_R + HT_R$).

$DT_R$ and $T_{1,R}^{Tr}$ of this case should be $A\hat{T}_1 + \Delta_{1,R}^{Tr}$ and $\Delta_{1,R}^{Tr}$, respectively. $T_{2,R}^{Tr}$ of both Case 3 and Case 4 is $SDT_R + 1 - A\hat{T}_2$ because run 2 misses run R in both cases.

Case 1’ to Case 4’ symmetrically correspond to Case 1 to Case 4 with a condition where the arrival time of run 2 is earlier than the given holding time and that of run 1. Run 2 is successfully connected to run R for the four cases. Case 5 is the worst and costliest case where both of the feeder runs miss run R. Therefore, run R does not leave for the next stop until $SDT_R + HT_R$ and all the transfer passengers must wait for the next available run.

Based on the above definitions, the expected values of the successful connection cost and the missed connection cost of the transfer passengers from run 1 are written as Equations 4-19 and 4-20, respectively. Their total connection cost is the sum of the two (Equation 4-21). In the same way, the costs of the passengers from run 2 are defined as Equations 4-22, 4-23, and 4-24.

The in-vehicle passenger cost, $E(T_R^{Inv})$, and downstream passenger cost, $E(T_R^{Down})$, can be estimated by taking $E(DT_R)$ of Equation 4-25 and $Var(DT_R)$ of Equation 4-26. $E(T_R^{Inv})$ is
obtained similar to the case where one feeder run is delayed (Equation 4-8 of Section 2.1).

$E( DT_R )$ and $Var( DT_R )$ are also passed to the downstream model in order to estimate $E( T_{R}^{Down} )$ as seen in Equation 4-10 of Section 2.1.

\[
E( T_{1,R}^{TrS} ) = P_{1,R}^{Tr} \times \int_{R^{Case^S_1}} \int [T_{1,R}^{Tr} \times f_{A_f_1}(x) \times f_{A_f_2}(y)] dxdy
\]  
(EQ. 4-19)

\[
E( T_{1,R}^{TrM} ) = P_{1,R}^{Tr} \times \int_{R^{Case^M_1}} \int [T_{1,R}^{Tr} \times f_{A_f_1}(x) \times f_{A_f_2}(y)] dxdy
\]  
(EQ. 4-20)

\[
E( T_{1,R} ) = E( T_{1,R}^{TrS} ) + E( T_{1,R}^{TrM} )
\]  
(EQ. 4-21)

\[
E( T_{2,R}^{TrS} ) = P_{2,R}^{Tr} \times \int_{R^{Case^S_2}} \int [T_{2,R}^{Tr} \times f_{A_f_1}(x) \times f_{A_f_2}(y)] dxdy
\]  
(EQ. 4-22)

\[
E( T_{2,R}^{TrM} ) = P_{2,R}^{Tr} \times \int_{R^{Case^M_2}} \int [T_{2,R}^{Tr} \times f_{A_f_1}(x) \times f_{A_f_2}(y)] dxdy
\]  
(EQ. 4-23)

\[
E( T_{2,R} ) = E( T_{2,R}^{TrS} ) + E( T_{2,R}^{TrM} )
\]  
(EQ. 4-24)

\[
E( DT_R ) = \int [DT_R \times f_{A_f_1}(x) \times f_{A_f_2}(y)] dxdy
\]  
(EQ. 4-25)

\[
Var( DT_R ) = E( DT_R^2 ) - \{ E( DT_R ) \}^2
\]  
(EQ. 4-26)

Where:

$CASE^{S_1}$: Set defined as \{Case1, Case1’, Case2, Case2’, Case3, Case4\}

$CASE^{M_1}$: Set defined as \{Case3’, Case4’, Case5\}

$CASE^{S_2}$: Set defined as \{Case1, Case1’, Case2, Case2’, Case3’, Case4’\}

$CASE^{M_2}$: Set defined as \{Case3, Case4, Case5\}
Therefore, when two feeder runs are delayed, the optimal holding time is obtained by minimizing the total cost function of Equation 4-15. It must be noted that the above formulations assume that run $R$ is held to wait for both run 1 and run 2. However, there exist other possible scenarios in applying the CP control when two feeder runs are delayed: run $R$ does not have to wait for both, but does have to wait for either run 1 or run 2. If the holding control is provided only to run 1, for example, then run $R$ will instantly depart as soon as it loads the passengers from run 1. When run 2 arrives earlier than run 1, of course, run 2 is successfully connected to run $R$; that is, run $R$ is not subject to run 2.

Denote $T_{1,R}^{Tr}$, $T_{2,R}^{Tr}$ and $DT_R$ as $\hat{T}_{1,R}^{Tr}$, $\hat{T}_{2,R}^{Tr}$ and $D\tilde{T}_R$ when the CP control is applied to run 1 only. Based on the formulation of $T_{1,R}^{Tr}$, $T_{2,R}^{Tr}$ and $DT_R$ of Equations 4-16, 4-17 and 4-18, $\hat{T}_{1,R}^{Tr}$, $\hat{T}_{2,R}^{Tr}$ and $D\tilde{T}_R$ are formulated as the following:

\[
\hat{T}_{1,R}^{Tr} = \begin{cases} 
\Delta_{1,R}^{Tr} + \Delta_{2,R}^{Tr} & \text{for Case1} \\
(A\hat{T}_2 - A\hat{T}_1) + \Delta_{1,R}^{Tr} + \Delta_{2,R}^{Tr} & \text{for Case1'} \\
\Delta_{1,R}^{Tr} & \text{for Case2} \\
\Delta_{1,R}^{Tr} & \text{for Case2'} \\
\Delta_{1,R}^{Tr} & \text{for Case3} \\
E(DT_{R+1}) - A\hat{T}_1 & \text{for Case3'} \\
\Delta_{1,R}^{Tr} & \text{for Case4} \\
E(DT_{R+1}) - A\hat{T}_1 & \text{for Case4'} \\
E(DT_{R+1}) - A\hat{T}_1 & \text{for Case5} 
\end{cases}
\] (EQ. 4-27)
\[
T_{2,R}^T = \begin{cases}
(A\hat{T}_1 - A\hat{T}_2) + \Delta_{1,R}^T + \Delta_{2,R}^T & \text{for Case1} \\
\Delta_{1,R}^T + \Delta_{2,R}^T & \text{for Case1'} \\
E(DT_{R+1}) - A\hat{T}_2 & \text{for Case2} \\
(A\hat{T}_1 - A\hat{T}_2) + \Delta_{1,R}^T & \text{for Case2'} \\
E(DT_{R+1}) - A\hat{T}_2 & \text{for Case3} \\
SDT_R + HT_R - A\hat{T}_2 & \text{for Case3'} \\
E(DT_{R+1}) - A\hat{T}_2 & \text{for Case4} \\
\Delta_{2,R}^T & \text{for Case4'} \\
E(DT_{R+1}) - A\hat{T}_2 & \text{for Case5}
\end{cases}
\text{(EQ. 4-28)}
\]

\[
D\hat{T}_R = \begin{cases}
A\hat{T}_1 + \Delta_{1,R}^T + \Delta_{2,R}^T & \text{for Case1} \\
A\hat{T}_2 + \Delta_{1,R}^T + \Delta_{2,R}^T & \text{for Case1'} \\
A\hat{T}_1 + \Delta_{1,R}^T & \text{for Case2} \\
A\hat{T}_1 + \Delta_{1,R}^T & \text{for Case2'} \\
A\hat{T}_1 + \Delta_{1,R}^T & \text{for Case3} \\
SDT_R + HT_R & \text{for Case3'} \\
A\hat{T}_1 + \Delta_{1,R}^T & \text{for Case4} \\
A\hat{T}_2 + \Delta_{2,R}^T & \text{for Case4'} \\
SDT_R + HT_R & \text{for Case5}
\end{cases}
\text{(EQ. 4-29)}
\]

Case2 and Case3 of Equations 4-16, 4-17 and 4-18 are revised due to the applied scenario where the holding of run R targets run 1 only. In Case2, even though the arrival time of run 2 is earlier than the given holding time, run 2 cannot reach run R. Run R does not stay until the holding time but leaves for the next stop instantly, because it has served run 1. Therefore, run R departs at \(A\hat{T}_1 + \Delta_{1,R}^T\cdot T_{1,R}^T\) and \(T_{2,R}^T\) of this case are respectively \(\Delta_{1,R}^T\) and \(SDT_{R+1} - A\hat{T}_2\). In Case3, run 2 is even later beyond the maximum holding time. \(D\hat{T}_R\) is formulated as \(A\hat{T}_1 + \Delta_{1,R}^T\cdot T_{1,R}^T\) and \(T_{2,R}^T\) become \(\Delta_{1,R}^T\) and \(SDT_{R+1} - A\hat{T}_2\), as in Case2. \(T_{1,R}^T\), \(T_{2,R}^T\) and \(D\hat{T}_R\), which are \(T_{1,R}^T\), \(T_{2,R}^T\) and \(D\hat{T}_R\) of a scenario where the CP control is applied to run 2, can be formulated in the same way.
Using $\tilde{T}_{1,R}^{Tr}$, $\tilde{T}_{2,R}^{Tr}$, $D\tilde{T}_{1,R}$, $\tilde{T}_{1,R}^{Tr}$, $\tilde{T}_{2,R}^{Tr}$ and $D\tilde{T}_{R}$, the following total cost functions can be obtained for each scenario:

$$E(\tilde{T}_{R}^{Total}) = E(\tilde{T}_{1,R}^{Tr}) + E(\tilde{T}_{2,R}^{Tr}) + E(\tilde{T}_{R}^{Inv}) + E(\tilde{T}_{R}^{Dwn})$$  \hspace{1cm} (EQ. 4-30)

$$E(\tilde{T}_{R}^{Total}) = E(\tilde{T}_{1,R}^{Tr}) + E(\tilde{T}_{2,R}^{Tr}) + E(\tilde{T}_{R}^{Inv}) + E(\tilde{T}_{R}^{Dwn})$$  \hspace{1cm} (EQ. 4-31)

Where:

$E(\tilde{T}_{R}^{Total})$: Expected total cost of all related passengers when the CP control is provided for run 1 only

$E(\tilde{T}_{R}^{Total})$: Expected total cost of all related passengers when the CP control is provided for run 2 only

Therefore, the optimal holding time should be found through the three cost functions, $E(\tilde{T}_{R}^{Total})$, $E(\tilde{T}_{1,R}^{Total})$ and $E(\tilde{T}_{2,R}^{Total})$ as in the following equation:

$$HT_{R}^{*} = Min \{ HT_{R}^{B*} = [t : Min\{ E(\tilde{T}_{R}^{Total} | HT_{R} = t) \} \] for \ 0 \leq t \},$$  \hspace{1cm} (EQ. 4-32)

$$HT_{R}^{1*} = [i : Min\{ E(\tilde{T}_{R}^{Total} | HT_{R} = i) \} \] for \ 0 \leq i \},$$

$$HT_{R}^{2*} = [i : Min\{ E(\tilde{T}_{R}^{Total} | HT_{R} = i) \} \] for \ 0 \leq i \}$$

Where:

$HT_{R}^{B*}$: Optimal holding time if the CP control is applied to both runs

$HT_{R}^{1*}$: Optimal holding time if the CP control is applied to run 1
\( HT_{R}^{2*} \) : Optimal holding time if the CP control is applied to run 2

Consequently, the next rule should be applied according to \( HT_{R}^{*} \):

If \( HT_{R}^{*} = 0 \), dispatch run R immediately

If \( HT_{R}^{*} = HT_{R}^{1*} \), hold run R until \( SDT_{R} + HT_{R}^{*} \) for run 1 and run 2

If \( HT_{R}^{*} = HT_{R}^{1*} \), hold run R until \( SDT_{R} + HT_{R}^{*} \) for run 1

If \( HT_{R}^{*} = HT_{R}^{2*} \), hold run R until \( SDT_{R} + HT_{R}^{*} \) for run 2

### 3.2 Analysis of Model

Unlike the case of Section 2 where one feeder run is late, the case of two late feeder runs requires the consideration of three cost functions in order to find the optimal holding time as described above. Also, the shape of the cost functions is more complex than that of the previous case. This section performs a sensitivity analysis for the case of two late feeder runs. The environment for the analysis is the same as that of the delay of one feeder run explained in Section 2.2.1, except that here there are two delayed feeder runs.

#### 3.2.1 Sensitivity Analysis

FIGURE 4-16 shows the cost functions according to the predicted arrival times of two delayed feeders, run 1 and run 2. The predicted arrival time of run 1, \( E(\hat{A}T_{1}) \), is fixed at 3:01:00, while the predicted arrival time of run 2, \( E(\hat{A}T_{2}) \), is varied from 3:01:00 to 3:04:00. The figure shows the three cost functions together: \( E(T_{R}^{Total}) \) is the cost function assuming that run R is held for
both; $E(\hat{T}_{R_{Total}})$ is the cost function assuming that run $R$ is held for run 1 only; and $E(\tilde{T}_{R_{Total}})$ is the cost function assuming that run $R$ is held for run 2 only.

In FIGURE 4-16 (a), the two delayed runs have the same predicted arrival time with the same number of transfer passengers. $E(T_{R_{Total}})$ has only one maximum point and converges in the same way as the total cost function of one delayed run, while the probabilities of successful connections of the two runs rise simultaneously with increasing holding times. The curves of $E(\hat{T}_{R_{Total}})$ and $E(\tilde{T}_{R_{Total}})$ are identical because $E(A\hat{T}_1)=E(A\hat{T}_2)$ and $P_{1,R}^{TR}=P_{2,R}^{TR}$. It can be seen that applying the CP control to both run 1 and run 2 is better than applying it to only one of them ($E(T_{R_{Total}}) < E(\hat{T}_{R_{Total}})=E(\tilde{T}_{R_{Total}})$).

As $E(A\hat{T}_2)$ becomes 3:02:00 in FIGURE 4-16 (b), $E(T_{R_{Total}})$ is more dispersed than that of FIGURE 4-16 (a), but it also has one maximum point. $E(\hat{T}_{R_{Total}})$ increases for $HT_R=0.5$ min., but starts to converge earlier than $E(T_{R_{Total}})$, because run $R$ is supposed to be held for run 1 under the scenario of $E(\hat{T}_{R_{Total}})$. $E(\tilde{T}_{R_{Total}})$ is almost identical to $E(T_{R_{Total}})$, as the probability that run 1 is ahead of run 2 is higher than that of the opposite case. That is, run 1 can successfully connect to run $R$, even though run $R$ is held only for run 2.

In FIGURE 4-16 (c) and (d) where $E(A\hat{T}_2)$ is given as 3:03:00 and 3:04:00 respectively, run 2 is most likely later than run 1 by a significant margin. Therefore, $E(T_{R_{Total}})$ has two maximum points and one local minimum point. $E(\hat{T}_{R_{Total}})$ is shaped in the same way as that of FIGURE 4-16 (b) and $E(\tilde{T}_{R_{Total}})$ coincides with $E(T_{R_{Total}})$ as well.

In order to observe the cost functions in detail, FIGURE 4-17 separately plots the $E(T_{R_{Total}})$, $E(\hat{T}_{R_{Total}})$ and $E(\tilde{T}_{R_{Total}})$ with their cost components, when $E(A\hat{T}_1)=3:01:00$ and $E(A\hat{T}_2)=3:03:00$. 
$E(\hat{T}_1) = 3:01:00, c_v = 0.2, P_{1,R}^{tr} = 7, P_{2,R}^{tr} = 7, P_R^{inv} = 20, SDT_R = 3:00, SDT_{R+1} = 3:30$

FIGURE 4-16 Variation of three cost functions against different $E(\hat{T}_2)$'s
(a) $E(T_{\text{Total}}^R)$ and cost components

(b) $E(\hat{T}_{R})$ and cost components

(c) $E(\hat{T}_{R}^{\text{Total}})$ and cost components

$E(A\hat{T}_1)=3:01$, $E(A\hat{T}_2)=3:03$, $c_v=0.2$, $P_{1,R}^{Tr}=7$, $P_{2,R}^{Tr}=7$, $P_{R}^{\text{hoV}}=20$, $SDT_{R}=3:00$, $SDT_{R+1}=3:30$

FIGURE 4-17 Three cost functions and their cost components
As seen in Seg-1 of FIGURE 4-17 (a), $E(T_{R,Total}^{TR})$ increases as the holding time increases, because $E(T_{R,Inv}^{TR})$ and $E(T_{R,Dwn}^{TR})$ rise. As the successful connection probability of run 1 grows in Seg-2, the cost of the transfer passengers of run 1 ($E(T_{L,R}^{TR})$) goes down sharply. Accordingly, $E(T_{R,Total}^{TR})$ decreases. The total cost increases again in Seg-3. Note that $E(T_{R,Inv}^{TR})$ and $E(T_{R,Dwn}^{TR})$ continuously rise because run R must be held for run 2 as well. In Seg-4, the decline in $E(T_{2,R}^{TR})$ makes $E(T_{R,Total}^{TR})$ decrease, as the successful connection probability of run 2 increases. $E(T_{R,Total}^{TR})$ and its cost components converge in Seg-5 because run R departs as soon as the transfer passengers of run 2 complete boarding onto run R.

$E(T_{R,Total}^{TR})$ of FIGURE 4-17 (b) varies significantly from $E(T_{R,Total}^{TR})$. Because run R waits only for run 1, $E(T_{R,Total}^{TR})$ converges around $HT_{R} \approx 1.5$. It can be seen that $E(T_{2,R}^{TR})$ remains almost constant and does not decrease even with sufficient holding times: the transfer from run 2 to run R is highly unlikely, because, in most cases, run 2 is later than run 1. FIGURE 4-17 (c) shows that $E(T_{R,Total}^{TR})$ is almost the same as $E(T_{R,Total}^{TR})$. In contrast to the previous cost function, run R is supposed to be held only for run 2. However, run 1 is successfully connected to run R, because run 1 most likely arrives earlier than run 2.

Accordingly, FIGURE 4-16 shows that if the predicted delay of run 2 is small, it is more beneficial to hold the receiving run for both until the first convergence point of $E(T_{R,Total}^{TR})$, since any cost of $E(T_{R,Total}^{TR})$ is lower than that of $E(T_{R,Total}^{TR})$. As the delay of run 2 increases, however, the benefit of holding for both begins to decrease. When the predicted arrival time of run 2 is 3:04:00, the optimal strategy is to let the receiving run wait for run 1 until the first convergence point of $E(T_{R,Total}^{TR})$. Of course, if run 1 and run 2 are significantly delayed, the costs of the three cost functions at their first convergence points will be larger than the cost of no holding.

The holding decision depends on the transfer demands from the two late runs as well. FIGURE 4-18 shows the variation of the cost functions according to the different combinations of the transfer passengers. If $P_{1,R}^{TR} = 11$ and $P_{2,R}^{TR} = 3$, it is clear that the optimal holding time exists on $E(T_{R,Total}^{TR})$: the receiving run must be held only for run 1 until the first convergence point of $E(T_{R,Total}^{TR})$. When $P_{2,R}^{TR}$ is relatively larger than $P_{1,R}^{TR}$, the right tail of $E(T_{R,Total}^{TR})$ will be located
under that of $E(T^{\text{Total}}_R)$. Accordingly, the optimal strategy is to hold the receiving run for run 2 as well as run 1.

$$E(\hat{T}_{1})=3:01, \; E(\hat{T}_{2})=3:03, \; c_v=0.2, \; SDT_R=3:00, \; SDT_{R+1}=3:30PM, \; P_{R}^{\text{InV}}=20$$

FIGURE 4-18 Variation of three cost functions against different $P_{1,R}^{Tr}$ and $P_{2,R}^{Tr}$

The holding decision is influenced by other parameters, such as the number of in-vehicle passengers ($P_{R}^{\text{InV}}$), the demand of downstream passengers, and the time at which the next run is available ($SDT_{R+1}$). FIGURE 4-19 shows the influence of $P_{R}^{\text{InV}}$ on the cost functions. As $P_{R}^{\text{InV}}$ increases, the decision of the CP model varies and the expected benefit of CP control decreases. When $P_{R}^{\text{InV}}=10$, it is beneficial to hold run R for both run 1 and run 2, because $E(T^{\text{Total}}_R)$, at its convergence point, is smaller than that of no holding and $E(T^{\text{Total}}_R)$ at its convergence point. As $P_{R}^{\text{InV}}$ increases, the overall benefit of CP control decreases. When $P_{R}^{\text{InV}}=40$, it is better to hold run R only for run 1, not for both. As seen in FIGURE 4-20, the number of downstream passengers has the same influence as $P_{R}^{\text{InV}}$ on the cost functions. With the default demand of TABLE 4-1, the cost of guaranteeing the connection of the two runs is smaller than the cost of
no control and the cost of holding run $R$ for run 1. When the demand increases four times, the opposite is observed. Accordingly, the optimal strategy is to let run $R$ depart without holding. FIGURE 4-21 presents the relationship between $SDT_{R+1}$ and the cost functions. As $SDT_{R+1}$ decreases, the benefit of the CP control declines. If the next run is available in 10 min., as seen in the figure, any holding of run $R$ is not beneficial. However, the benefit of the holding significantly increases, as the time gap to the next run becomes longer.

\[ E(\hat{A}_1) = 3:01, \quad E(\hat{A}_2) = 3:03, \quad c_v = 0.2, \quad P_{1,R}^{Tr} = 7, \quad P_{2,R}^{Tr} = 7, \quad SDT_R = 3:00, \quad SDT_{R+1} = 3:30 \]

FIGURE 4-19 Variation of three cost functions against different $P_R^{InV}$'s
FIGURE 4-20 Variation of three cost function against different demand levels of downstream stops

\[E(\hat{A}_{T1})=3:01/ E(\hat{A}_{T2})=3:03/ c_v = 0.2/ P_{1,R}^{Tr} = 7/ P_{2,R}^{Tr} = 7/ SDT_R = 3:00/ SDT_{R+1} = 3:30/ P_R^{Inv} = 20\]

FIGURE 4-21 Variation of three cost functions against different SDT_{R+1}'s

\[E(\hat{A}_{T1})=3:01/ E(\hat{A}_{T2})=3:03/ c_v = 0.2/ P_{1,R}^{Tr} = 7/ P_{2,R}^{Tr} = 7/ SDT_R = 3:00/ P_R^{Inv} = 20\]
3.2.2 Decision of Optimal Holding Time

For the real-time application of the CP control, an efficient solution algorithm is essential for the case of the two late feeders. The analysis of the previous section offers some insights which can be used to develop a heuristic rule to find the optimal holding time. If it is predicted that the two delayed runs come to a transfer point without significant gaps, such as the cases of FIGURE 4-16 (a) and (b), the cost functions will have one maximum point, such as the cost function of one late feeder run. Generally, holding run $R$ for both runs would be better than providing the holding to only one. Accordingly, the optimal holding time could be either at the left boundary (i.e., holding time of 0) or the first convergence point of the three functions ($E(T_{R}^{Total})$, $E(\dot{T}_{R}^{Total})$ and $E(\ddot{T}_{R}^{Total})$) as in the case of a single late feeder.

If $E(T_{R}^{Total})$ has two maximum points and one local minimum point (i.e., if the predicted arrivals of the two feeder runs are somewhat spaced out), such as FIGURE 4-16 (c) and (d), the local minimum point of $E(T_{R}^{Total})$ between the two maximum points must be found and examined whether it is the global minimum of $E(T_{R}^{Total})$ or not. Fortunately, there is an easier way to decide on the optimal holding time without finding the local minimum point.

Denote $HT_{R}^{B-L}$ as the local minimum point for $E(T_{R}^{Total})$. FIGURE 4-16 (c) and (d) show that $E(\dot{T}_{R}^{Total} | HT_{R} = HT_{R}^{B-L})$ is very similar to $E(T_{R}^{Total} | HT_{R} = HT_{R}^{B-L})$. Beyond $HT_{R}^{B-L}$, however, $E(\ddot{T}_{R}^{Total})$ moves down, then converges, because the departure of run $R$ is subject only to run 1, while $E(T_{R}^{Total})$ increases again due to run 2. FIGURE 4-22 provides this relationship among the three cost functions through conceptual sketches. It implies that the search of $HT_{R}^{B-L}$ can be omitted by utilizing $E(\ddot{T}_{R}^{Total})$. 
Based on the figure, the following rule can be established (suppose that $CP^1$, $CP^2$ and $CP^B$ are the first convergence points of $E(\hat{T}_R^{\text{Total}})$, $E(\hat{T}_R^{\text{Total}})$ and $E(T_R^{\text{Total}})$, respectively): If $E(\hat{T}_R^{\text{Total}} \mid HT_R = CP^1)$ is less than $E(T_R^{\text{Total}} \mid HT_R = 0)$ and $E(T_R^{\text{Total}} \mid HT_R = CP^B)$ (FIGURE 4-22 (a)), then, the optimal holding decision is to hold run $R$ until $CP^1$, but to have it depart as soon as it processes the transfer passengers of run 1; If $E(T_R^{\text{Total}} \mid HT_R = CP^B)$ is less than $E(\hat{T}_R^{\text{Total}} \mid HT_R = CP^1)$ and $E(T_R^{\text{Total}} \mid HT_R = 0)$ (FIGURE 4-22 (b)), then, the optimal holding decision is to hold run $R$ until $CP^B$, but to have it depart as soon as it processes the transfer passengers of run 1 and run 2; If not, the optimal holding decision is to have run $R$ depart as scheduled. That is, the optimal holding time can be found by checking the convergence points of the cost functions without finding $HT_R^{B-L}$.

Like the rule established in Section 2.2.3 (FIGURE 4-14), the convergence point is also defined as the 3-sigma point. It is possible to write a heuristic in some formal form as seen in FIGURE 4-23. If there exists a maximum holding time, the 3-sigma point of the procedure can be replaced by the holding time maximally allowed.
Step-0) Define \( CP_1, CP_2 \) and \( CP^B \)
\[
CP_1 = E(A \hat{T}_1) + 3 \times \sigma_{A \hat{T}_1}
\]
\[
CP_2 = E(A \hat{T}_2) + 3 \times \sigma_{A \hat{T}_2}
\]
\[
CP^B = \text{Max}(CP_1, CP_2)
\]

Step-1) Find \( HT_R^{B*}, HT_R^{1*} \) and \( HT_R^{2*} \)
\[
\text{IF } E(T_{R \text{Total}} \mid HT_R = 0) > E(T_{R \text{Total}} \mid HT_R = CP^B) \]
\[
HT_R^{B*} = CP^B - SDT_R
\]
\[
\text{ELSE}
\]
\[
HT_R^{B*} = 0
\]
\[
\text{IF } E(T_{R \text{Total}} \mid HT_R = 0) > E(T_{R \text{Total}} \mid HT_R = CP_1) \]
\[
HT_R^{1*} = CP_1 - SDT_R
\]
\[
\text{ELSE}
\]
\[
HT_R^{1*} = 0
\]
\[
\text{IF } E(T_{R \text{Total}} \mid HT_R = 0) > E(T_{R \text{Total}} \mid HT_R = CP_2) \]
\[
HT_R^{2*} = CP_2 - SDT_R
\]
\[
\text{ELSE}
\]
\[
HT_R^{2*} = 0
\]

Step-2) Find \( HT_R^* \) and make holding decision for run \( R \)
\[
HT_R^* = \text{Min}(HT_R^{B*}, HT_R^{1*}, HT_R^{2*})
\]
\[
\text{IF } HT_R^* = 0
\]
\[
\text{Depart Now,}
\]
\[
\text{ELSE IF } HT_R^* = HT_R^{B*}
\]
\[
\text{Hold run R for run 1 and run 2 for } HT_R^*
\]
\[
\text{ELSE IF } HT_R^* = HT_R^{1*}
\]
\[
\text{Hold run R only for run 1 for } HT_R^*
\]
\[
\text{ELSE IF } HT_R^* = HT_R^{2*}
\]
\[
\text{Hold run R only for run 2 for } HT_R^*
\]

FIGURE 4-23 Heuristic to find the optimal holding time (case of two late feeder runs)
4 Comparison with Previous Studies

This section compares the proposed CP model with previous models. As reviewed in Section 3 of Chapter 2, there have been several models for holding control that were developed in previous studies. TABLE 4-2 summarizes the differences among the key models, including the model of this study.

TABLE 4-2 Comparison of CP models

<table>
<thead>
<tr>
<th>CP models</th>
<th>Arrival of feeder</th>
<th>Cost function components</th>
<th>Time value of delay</th>
<th>Downstream model</th>
<th>Early dispatching</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dessouky et al. (1999)</td>
<td>D</td>
<td>Tr-P, InV-P &amp; Dn-P</td>
<td>No</td>
<td>Yes: P</td>
<td>N/A</td>
</tr>
<tr>
<td>Chung and Shalaby (2007)</td>
<td>D</td>
<td>Tr-P, InV-P &amp; Dn-P</td>
<td>No</td>
<td>Yes: D</td>
<td>N/A</td>
</tr>
<tr>
<td>Lee (1993)</td>
<td>P (N)</td>
<td>Tr-P, InV-P, Dn-P, &amp; OC</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Ting (1997)</td>
<td>P (N)</td>
<td>Tr-P, InV-P, Dn-P, &amp; OC</td>
<td>Yes</td>
<td>Yes: D</td>
<td>No</td>
</tr>
<tr>
<td>Chowdhury and Chien (2001)</td>
<td>P (N &amp; LN)</td>
<td>Tr-P, InV-P &amp; OC</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Hall et al. (2001)</td>
<td>P (N)</td>
<td>Tr-P, InV-P &amp; Dn-P</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>This Study</td>
<td>P (N)</td>
<td>Tr-P, InV-P &amp; Dn-P</td>
<td>No</td>
<td>Yes: P</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Where:
D: Deterministic, P: Probabilistic, N: Normal Distribution, LN: Lognormal Distribution
Tr-P: Transfer Passenger Cost, InV-P: In-Vehicle Passenger Cost
Dn-P: Downstream Passenger Cost, OC: Transit Operator Cost

This table shows that all the models share the same fundamental approach: the optimal holding time is found based on the costs of the related passengers. Even though our model ignores the cost of a transit operator incurred by the holding control, it can be added to the total cost function easily enough through a linear function of the departure delay. Also, the time value of delay (the weight for each group of passengers in $/pass.min) can be considered by assigning an appropriate weight to each cost component. For example, Lee applied different time values of
delay to the groups of passengers: 0.2 $/pass.min, to transfer passengers, and 0.25 $/pass.min. to
in-vehicle passengers and downstream passengers. In the previous studies, however, such values
were simply assumed rather than based on actual studies.

Some models have the downstream passengers’ cost, but do not model the downstream operation
explicitly (Lee, 1993; Hall et al., 2001). Those models assume that the number of the
downstream passengers is known, and estimate the associated cost by multiplying this number
with the receiver’s departure delay.

The model of this study is very similar to that of Hall et al. in that both allow for early
dispatching; i.e., the receiving run is allowed to depart immediately once all the feeder runs have
arrived. When the number of delayed feeders is one, the two models are almost the same except
that this study accounts for the boarding time of transfer passengers and has the downstream
model. When the number of delayed feeders is two, however, the decision of the model will be
different from that of Hall et al., because this study allows the receiving run to be held for either
one of the two delayed feeders or for both of them.

Based on the table, it is possible to categorize the models as deterministic and probabilistic
according to whether the arrivals of late feeders are probabilistic or not. In the case of the
probabilistic approach, a further categorization of the models can be made based on whether they
allow early dispatching or not. The next section explains in detail the deterministic approach
used by Dessouky et al. and Chung and Shalaby, while the following section presents the
probabilistic model without early dispatching used by Lee, Ting, and Chowdhury and Chien.
Next, we compare the two approaches with the model of this study through a simulation.

4.1 Deterministic CP Model (PM-1)

This approach is based on the assumption that the model knows the exact arrival times of the
delayed run (Dessouky et al., 1999; Chung and Shalaby 2007). Suppose that a single feeder run,
run 1, is delayed and its predicted arrival time, \( AT_1 \), is available. Since the predicted arrival time
of run 1 is assumed as a deterministic value, the transfer to a receiving run, run \( R \), will be
successful with the holding time of $A_{\hat{T}_1} - SDT_R$. Accordingly, if run $R$ is held, the successful connection cost and the missed connection cost of run 1 will be:

\[
T_{1,R}^{TrS} = D_R^{Tr} \times (DT_R - A_{\hat{T}_1}) = 0 \quad \text{(EQ. 4-33)}
\]

\[
T_{1,R}^{TrM} = 0 \quad \text{(EQ. 4-34)}
\]

The successful connection cost is defined as the waiting time of transfer passengers when the delayed run 1 reaches run $R$. As seen in the Equation 4-33, it is the difference between the departure time of run $R$ ($DT_R$) and the arrival time of run 1 ($A_{\hat{T}_1}$). By ignoring the boarding time of the transfer passengers, however, $DT_R$ becomes $A_{\hat{T}_1}$ (all previous studies ignored the boarding time) and $T_{1,R}^{TrS}$ will be zero. The missed connection cost of Equation 4-34 is also zero, because the transfer is successful. In-vehicle passengers’ cost ($T_R^{Inv}$) and downstream passengers’ cost ($T_R^{Dsn}$) depend on $DT_R$. Since we assume $DT_R = A_{\hat{T}_1}$, the two costs will be the following:

\[
T_R^{Inv} = P_R^{Inv} \times (DT_R - SDT_R) = P_R^{Inv} \times (A_{\hat{T}_1} - SDT_R) \quad \text{(EQ. 4-35)}
\]

\[
T_R^{Dsn} = fun_{dsn} ( [ DT_R \mid HT_R = DT_R - SDT_R ] ) - fun_{dsn} ( [ DT_R \mid HT_R = 0 ] ) \quad \text{(EQ. 4-36)}
\]

Therefore, the total cost when run $R$ is held is:

\[
[T_R^{Total} \mid HT_R = A_{\hat{T}_1} - SDT_R] = \{ P_R^{Inv} \times (A_{\hat{T}_1} - SDT_R) \}
\]
\[ +\{ \text{fun}_{\text{dns}} \left( \left[ DT_R \mid HT_R = DT_R - SDT_R \right] \right) - \text{fun}_{\text{dns}} \left( \left[ DT_R \mid HT_R = 0 \right] \right) \} \text{ (EQ. 4-37)} \]

When run \( R \) departs as scheduled without any holding, the transfer passengers will miss run \( R \), and consequently have to wait for the next available run, run \( R+1 \). Therefore, \( T_{\text{TrS}}^{1,R} \) and \( T_{\text{TrM}}^{1,R} \) are:

\[ T_{\text{TrS}}^{1,R} = 0 \text{ (EQ. 4-38)} \]
\[ T_{\text{TrM}}^{1,R} = P_{1,R}^{\text{Tr}} \times (SDT_{R+1} - \hat{A}_1) \text{ (EQ. 4-39)} \]

Since run \( R \) departs without any delay, both \( T_{R}^{\text{InV}} \) and \( T_{R}^{\text{Dwn}} \) are zero. Therefore, the total cost when run \( R \) is held is:

\[ [T_{R}^{\text{Total}} \mid HT_R = 0] = P_{1,R}^{\text{Tr}} \times (SDT_{R+1} - \hat{A}_1) \text{ (EQ. 4-40)} \]

Finally, the holding decision is made by comparing the two costs of Equations 4-37 and 4-40.

\[ HT_R^* = \begin{cases} 0 & \text{if } [T_{R}^{\text{Total}} \mid HT_R = 0] < [T_{R}^{\text{Total}} \mid HT_R = \hat{A}_1 - SDT_R] \\ \hat{A}_1 - SDT_R & \text{Otherwise} \end{cases} \text{ (EQ. 4-41)} \]
By adding into the formulation the constraint of maximum holding time, if $\hat{A}T_1$ is larger than $SDT_R + HT_R^{Max}$ (i.e., run 1 does not arrive within the maximum holding time), $HT_R^*$ will always be zero. This model can be easily extended to the case that two or more feeder runs are delayed, as presented in FIGURE 4-24. More details of this deterministic CP model can be found in Chung and Shalaby (2007), the initial work of this study, which also includes the deterministic version of the downstream model. Unfortunately, the effectiveness of this approach is undermined by the fact that it ignores the variances of the predicted arrival times of the delayed runs.

\[
D (=\text{the number of delayed run}) \geq 1, \text{ The optimal holding time, } HT_R^*, \text{ is:}
\]

\[
HT_R^* = \left\{ t : \text{Min}\{(T_{R}^{Total}|HT_R=t) \text{ for } t=0, \hat{A}T_1 - SDT_R, \hat{A}T_2 - SDT_R, \ldots, \hat{A}T_D - SDT_R \} \right\}
\]

Where:

\[
T_{R}^{Total} = \sum_{d=1}^{D} (T_{T_{d,R}}^{TrS} + T_{T_{d,R}}^{TrM}) + T_{R}^{InV} + T_{R}^{Dwn},
\]

\[
T_{TrS}^{TrS}_{d,R} = \begin{cases} P_{d,R}^{Tr} \times (DT_{R} - \hat{A}T_d) & \text{if } \hat{A}T_d <= SDT_R + HT_R, \text{ for } d=1,2,\ldots,D \\ 0 & \text{Otherwise} \end{cases}
\]

\[
T_{TrM}^{TrM}_{d,R} = \begin{cases} P_{d,R}^{Tr} \times (SDT_{R+1} - \hat{A}T_d) & \text{if } \hat{A}T_d <= SDT_R + HT_R, \text{ for } d=1,2,\ldots,D \\ 0 & \text{Otherwise} \end{cases}
\]

\[
T_{R}^{InV} = P_{R}^{InV} \times (DT_{R} - SDT_R)
\]

\[
T_{R}^{Dwn} = \text{fun}_{d,n} (\left[ DT_{R} \restriction HT_R = DT_{R} - SDT_R \right]) - \text{fun}_{d,n} (\left[ DT_{R} \restriction HT_R = 0 \right])
\]

\[
DT_R = SDT_R + HT_R
\]

**FIGURE 4-24 Generic formulation of the deterministic CP Model**

### 4.2 Probabilistic CP Model without Early Dispatching (PM-2)

This approach is similar to the model of this study in that it considers the predicted arrival times of the delayed runs as random variables (Lee, 1993; Ting, 1997; Chowdhury and Chien, 2001). However, it ignores the early dispatching of a receiving run; it assumes that run $R$ stays at the transfer point until the expiry of the pre-determined holding time even though all late runs may
arrive earlier. Strictly speaking, it is the real-time extension of the approach employing scheduled departure time.

When the number of delayed runs is one (i.e. one feeder route only), the costs of the transfer passengers are:

\[
E(T_{i,R}^{TrS}) = P_{i,R}^{Tr} \times \int_{SDT_R}^{SDT_R+HT_R} (SDT_R + HT_R - x) f_{A_{i,R}}(x) dx
\]  
(EQ. 4-42)

\[
E(T_{i,R}^{TrM}) = P_{i,R}^{Tr} \times \int_{SDT_R+HT_R}^{\infty} (SDT_{R+1} + HT_R - x) f_{A_{i,R}}(x) dx
\]  
(EQ. 4-43)

Since run \( R \) is supposed to stay at the transfer point until the given holding time even with the arrival of run 1, the transfer passengers of run 1 experience some delay until run \( R \) departs. The successful connection cost of Equation 4-42 is based on the difference between the arrival time of run 1 and the departure time of run \( R (DT_R = SDT_R + HT_R) \). The missed connection cost of Equation 4-43 is the same as that of this study. Accordingly, the expected costs of in-vehicle passengers and downstream passengers are defined as Equations 4-44 and 4-45, respectively.

\[
E(T_R^{Inv}) = P_R^{Inv} \times [HT_R]
\]  
(EQ 4-44)

\[
E(T_R^{Down}) = fun_{dnt} (SDT_R + HT_R) - fun_{dnt} (SDT_R)
\]  
(EQ 4-45)

The total cost is the sum of the four costs presented above. The studies of Lee, Ting, and Chowdhury and Chien added the operational cost to the transit service provider as well. The cost was estimated through a simple linear function that proportionally increases depending on the holding time. However, all the studies commonly build the cost functions based on the above...
equations. For the case where two or more feeder runs are delayed, PM-2 has a general form, as presented in FIGURE 4-25.

If \( D \) (the number of delayed run) \( \geq 1 \), the optimal holding time, \( HT^*_R \), is:

\[
HT^*_R = \{ t : \min \{ E( T_{R}^{\text{total}} | HT_R = t) \} \text{ for } 0 \leq t \leq HT_R^{\text{Max}} \}
\]

Where:

\[
E( T_{R}^{\text{Total}} ) = \sum_{d=1}^{D} \{ E( T_{d,R}^{\text{TrS}} ) + E( T_{d,R}^{\text{TrM}} ) \} + E( T_{R}^{\text{Inv}} ) + E( T_{R}^{\text{Down}} ),
\]

\[
E( T_{d,R}^{\text{TrS}} ) = \int_{SDT_R + HT_R}^{\infty} (SDT_R + HT_R - x) f_{\tilde{A}_d'}(x) dx, \text{ for } d = 1,2,\ldots,D
\]

\[
E( T_{d,R}^{\text{TrM}} ) = \int_{SDT_R + HT_R}^{\infty} (SDT_{R+1} - x) f_{\tilde{A}_d'}(x) dx, \text{ for } d = 2,\ldots,D
\]

\[
E( T_{R}^{\text{Inv}} ) = \int_{SDT_R + HT_R}^{\infty} \{ HT_R \},
\]

\[
E( T_{R}^{\text{Down}} ) = \int_{SDT_R + HT_R}^{\infty} \{ HT_R \} - \int_{SDT_R + HT_R}^{\infty} \{ HT_R = 0 \}
\]

\[
DT_R = SDT_R + HT_R.
\]

FIGURE 4-25 Generic formulation of the probabilistic CP Model without early dispatching

The next figure shows how the cost function of PM-2 is different from that of this study.

FIGURE 4-26 (a) shows the total cost functions of the two models against different \( E( A\tilde{T}_i ) \)'s.

FIGURE 4-26 (b) plots the cost components of each model in detail, when \( E( A\tilde{T}_i ) = 3:03 \). The downstream model of this study is integrated with PM-2 in order to estimate the downstream cost of the model.
Holding Time (min.)
Total Cost (pass. min.)

(a) Total cost functions

(b) Cost components

\[ c_v = 0.2, \; P_{1,R}^{Tr} = 10, \; P_{R}^{hv} = 20, \; SDT_R = 3:00, \; SDT_{R+1} = 3:30 \]

FIGURE 4-26 Comparison of cost functions: with early dispatching vs. without early dispatching

TABLE 4-3 Comparison of holding decisions: with early dispatching vs. without early dispatching

<table>
<thead>
<tr>
<th>( E (\hat{A}T_1) )</th>
<th>PM-2 (Without early dispatching)</th>
<th>This Study (With early dispatching)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( HT^*_R )</td>
<td>Total cost</td>
</tr>
<tr>
<td>3:01:00</td>
<td>1.5</td>
<td>69.9</td>
</tr>
<tr>
<td>3:02:00</td>
<td>2.67</td>
<td>139</td>
</tr>
<tr>
<td>3:03:00</td>
<td>3.67</td>
<td>221.1</td>
</tr>
<tr>
<td>3:04:00</td>
<td>0</td>
<td>290</td>
</tr>
<tr>
<td>3:05:00</td>
<td>0</td>
<td>290</td>
</tr>
</tbody>
</table>
In FIGURE 4-26 (a), the total cost function of PM-2 does not converge but continuously rises as the holding time increases. This results from the condition that the receiving run stays until the expiry of the fixed holding time, even though the passengers of the feeder run may have arrived and boarded the receiving run before that time. FIGURE 4-26 (b) shows the influences of the assumptions on the cost components. As the holding time increases, \( E(T^\gamma_{l_R}) \), \( E(T^{InV}_{l_R}) \) and \( E(T^D_{l_R}) \) of PM-2 continuously increase due to the assumption.

TABLE 4-3 shows the decision of the two models and the expected total costs, based on FIGURE 4-26 (a). When \( E(A^\hat{T}_1) = 3:01:00 \), the decisions of the two model are very similar to each other. As \( E(A^\hat{T}_1) \) increases to \( E(A^\hat{T}_1) = 3:03 \), however, the optimal holding time of this study is later than that of PM-2, because of the different assumptions. Allowing for early dispatching, it is possible to increase the optimal holding time and achieve more benefit from the holding control. When \( E(A^\hat{T}_1) = 3:04 \) or 3:05, the two models commonly conclude that run R should be dispatched without any holding.

### 4.3 Comparison Through Simulation

In this section, the model of this study is compared with the deterministic model (PM-1) and the probabilistic model without early dispatching (PM-2) through a simulation. The simulation is based on the testbed previously explained: at 3:00:00, a receiving run is held in order to protect the target connection for the incoming transfer passengers from delayed feeders. FIGURE 4-27 presents the planned simulation scenarios. It is assumed that the delayed feeders randomly arrive at the transfer point from 3:00:00 to 3:06:00. There exists an hypothetical prediction model offering the predicted arrival time which includes an intentional margin of prediction error with 0.1, 0.2 and 0.3 of \( c_v \)’s. As seen in the figure, some errors are intentionally added to the predicted arrival times provided to the holding model. The holding model makes a holding decision using the prediction value, and the simulation model simulates the transfer process with the holding decision. It is assumed that the range of the holding times is unconstrained. However, the simulation also considers cases that the range of the holding times is restricted from one to
five with increments of one min. The receiving transit departs as soon as it processes the transfer passengers from the delayed feeders. The number of transfer passengers and the number of in-vehicle passengers are considered as deterministic values, which are input to the models. The CP models share the downstream model of Section 4 of Chapter 3, and the downstream conditions of TABLE 4-1 are input to the models. The simulation of the downstream operation of the receiving run is replaced with the downstream model: for each run of simulation, once the simulator computes the departure time of the receiving run based on the given holding decision, the downstream model estimates the waiting times of the downstream passengers.

FIGURE 4-27 Framework of simulation

FIGURE 4-28 shows the simulated results for the case of one delayed feeder. The number of transfer passengers and the number of in-vehicle passengers are given as 10 and 20, respectively. The Y-axis represents the waiting times of all passengers (transfer passengers, in-vehicle passengers and downstream passengers) after 5,000 runs of the simulation. FIGURE 4-28 (a), (b) and (c) represent the results with 0.1, 0.2 and 0.3 of \( c_v \)'s, respectively. CP refers to the model of this study, while No Hold is the departure of the receiving run without any holding. The simulated results indicate that PM-1 is much less efficient than PM-2 and CP, since it ignores the errors of the predicted values. The gap between PM-2 and CP is caused by the condition of PM-2
that does not allow early dispatching. The results imply that, taking early dispatching into account, we can achieve more of a benefit from the holding control.

These results also indicate that the prediction errors reduce the benefit of the holding control. The average waiting times of $c_v = 0.3$ (FIGURE 4-28 (c)) are higher than those of $c_v = 0.1$ (FIGURE 4-28 (a)). As the error increases, the average waiting time of the passengers increases, and the efficiency of the holding control decreases. Also, the maximum holding time has an influence on the efficiency of the holding model. The results also show that if the values of the maximum holding times are too small, the benefit of the holding can be restricted.
FIGURE 4-29 shows the results of the case where two feeders are delayed. It should be noted that one model is additionally considered in the comparison of this case. The model (referred to as CP.S1 in the figure) is based on the approach of this study, but finds the optimal holding time without the scenario that the receiving run can be held for one of the two delayed feeders. In the figures, it can be seen that there exist some gaps between CP.S1 and CP. Those additional benefits are caused by the inclusion of the scenario that the receiving run can be held for one of the two delayed feeders. The other results are not different from those of one delayed feeder: the performances of the probabilistic models are better than that of deterministic approach, and the models accounting for the early dispatching, CP.S1 and CP, are better than the other, PM-2.

(a) Prediction error of $c_v = 0.1$

(b) Prediction error of $c_v = 0.2$

(c) Prediction error of $c_v = 0.3$

$P_{1,R}^{T_r} = 5$, $P_{2,R}^{T_r} = 10$, $P_{R}^{Inv} = 20$, $SDT_R = 3:00$, $SDT_{R+1} = 3:30$

FIGURE 4-29 Comparison via simulation (case of two delayed feeders)
Those simulated results show that there exist discrepancies on the holding decisions among the CP models, depending on: whether a model considers the prediction errors; whether a model considers the early dispatching strategy or not; and whether it considers the flexibility of holding the receiving run for one of the two delayed feeders or not. Taking all those factors into consideration, therefore, the CP model of this study can achieve a more efficient decision on holding the receiving run.

5 Summary

This chapter suggests a strategy for the CP problem: to hold a transit unit in order to wait for another transit unit that was planned to provide a timed transfer but had been delayed. While the holding of Chapter 3 is a static dispatching policy (i.e., non real-time) based on historical patterns, this chapter focuses on a real-time CP control based on real-time data, accounting for possible daily variances of transit operations.

The developed CP model was designed to protect the planned connections based on the transfer passenger waiting time, the in-vehicle passenger waiting time, and the downstream passenger waiting time. The optimal holding decision of the model is a trade-off between such waiting time components. Furthermore, it explicitly considers the errors on the predicted arrival time of a delayed run. The evaluation of the model shows that it is able to capture adequately those trade-offs and make the appropriate decisions according to the given delays.

The model of this study incorporates the early dispatching of a receiving run, which was ignored in almost all of the previous studies. It also includes a holding scenario that a receiving run can be held for one of the two delayed feeders as well as for the both of them, when two delayed feeders exist. It can be seen that those properties are effective in reducing the waiting times of all passengers affected by the coordination among the lines.
Chapter 5
Conclusions and Recommendations

1 Introduction

This study has focused on the transfer management problem, and two approaches have been developed for the provision of efficient transfers. In Chapter 3, a transfer optimization model is proposed in order to coordinate inter-model transfer between commuter transit services and local transit services at a suburban transfer point. Chapter 4 presents a CP (Connection Protection) control model to protect a coordinated transfer against unexpected delays of the transit feeder run. In the two chapters, cost functions are formulated to account for not only the delay times of the transfer passengers, but also to account for the impact of the transfer optimization and CP control on the delay times of other non-transfer passengers. The developed models explicitly incorporate the probabilistic nature of transit operations, and formulate the given problems as realistically as possible. It is employs a relatively sophisticated approach with high potential to yield a good solution. This study excludes the operator costs, since the models produced new schedules by modifying the existing schedules and controlled a transit unit without the requirement of additional transit units. However, costs that are potentially incurred by the operator can be added to the objective functions quite easily and without loss of generality using appropriate conversion factors to account for the difference in weights between the system cost attributes and those of the waiting times of passengers. The developed models would be beneficial for transit service providers to revise their existing schedules and operate their transit systems in a coordinated way.

2 Transfer Optimization for Inter-modal Transfer

The transfer optimization model minimizes the transfer-associated costs by modifying the existing schedule of the local transit service, while at the same time keeping the schedule of the commuter transit service fixed. In order to provide better transfers, the model allows for not only the shifting of the whole schedule, but also the changing of the scheduled headway.
Simultaneously, the model considers an additional dispatching policy for further reduction of transfer-related costs: a transit unit can extend its departure beyond its scheduled departure time up to a given holding time so as to wait for a delayed run, but will be dispatched immediately if the transfer is successfully completed before the maximum holding time has expired. A “downstream model” is also developed to evaluate the influences of schedule shifting and the dispatching policy on the waiting times of non-transfer passengers at downstream stops. It is a probabilistic model that explains the arrival and departure process of the local transit run along its downstream stops, and estimates the numbers of the stop passengers and their waiting times at the stops. A genetic algorithm is adopted as a solution approach for the given optimization problem.

The case study, which is based on actual schedule data, shows that the developed optimization model is promising and favourable to improving timetables in terms of transfer related costs. The transfer related delays can be reduced significantly by simply shifting a whole schedule. The allowance of changing the scheduled headway offers potential for additional improvements, as the optimization model could tweak the schedule by independently shifting transit trip chains. Consequently, it improves the probabilities of successful connections and decreases the transfer costs, compared to the scenario of shifting a whole schedule. The additional dispatching policy could bring further reduction of the total cost. The results indicate that the developed model can be applied to revise the existing schedule with various optimization scenarios, and reasonably improve the overall transfer related costs.

From a theoretical point of view, the developed model shows that transfer optimization encompasses a set of trade-offs that affect various groups of transfer passengers differently. That is, the revised schedule may not be beneficial for all transfer passengers across all scheduled time connections, but instead some originally intended transfers may be sacrificed, since the objective function is to minimize the overall cost. To deal with the randomness in the transit operations, a transfer slack time should be provided to each transfer to render it efficient, while balancing the successful connection cost and the missed connection cost. The decision of the optimal transfer slack time is also a trade-off between the two costs: too much slack time increases the successful connection cost, while not enough slack time raises the missed connection cost. By introducing the additional dispatching policy, the model can bring further reduction in the total cost by making use of the holding time with a small margin of the transfer slack time, instead of
increasing the slack time. The developed function to estimate the expected transfer time of transfer passengers adequately evaluates the transfer costs according to the given operational scenarios, and, therefore, offers a vital mechanism for the transfer optimization.

3 Real-Time Control Strategy

A real-time CP control model is formulated to maintain a scheduled coordinated transfer, considering the daily variance of transit operations. The control involves holding a transit unit to wait for another transit unit that is planned to provide a coordinated transfer but has been delayed using real-time information from AVL (Automatic Vehicle Location) and APC (Automatic Passenger Counting) systems.

The optimal CP control is basically a trade-off between benefits and costs of applying the holding control in terms of passenger waiting times. The holding control reduces the waiting times of transfer passengers, but it delays the in-vehicle passengers who are already aboard the receiving run as well as the passengers who are waiting for the run at the downstream stops. Accordingly, the developed CP model was designed to make a holding decision based on the waiting times of the three groups of related passengers: the transfer passenger waiting time; the in-vehicle passenger waiting time; and the downstream passenger waiting time. The analysis of the model shows that it is able to capture adequately those trade-offs and make appropriate decisions according to the given delays.

Even though it assumes that an external model provides the predicted arrival times of delayed feeders, this study recognizes that the predicted arrival times by any prediction model are random variables having some probability distribution, and it explicitly considers the predicted arrival time of a delayed run as a random variable normally distributed. Therefore, the developed CP model makes a more robust decision than that using a deterministic approach.

Comparing it to previous studies, the model of this study has the following features: the estimation of the downstream passengers’ cost through a sophisticated downstream model; the allowance of early dispatching of a receiving run; and the holding of a receiving run for one of the two delayed feeders as well as for the both of them when the number of delayed feeders is
two. The evaluation shows that those properties are more beneficial to the waiting times of related passengers in applying the CP control.

4 Major Contributions

This study contributes to the transfer management problem in that it takes two approaches to the provision of efficient transfers: the planning-stage and operation-stage approaches. The planning-stage approach, based on the transfer optimization, schedules the departure times at a transfer point so as to minimize the overall waiting times of the related transit users. Although the timed transfer approach -- synchronizing the timetables through a pulse headway -- is a viable option, it would be a formidable task when being applied to inter-modal transfer coordination, because of institutional barriers between two transit agencies and the complexity of simultaneous rescheduling of commuter transit and local transit services. Accordingly, this study modifies only the existing schedule of the local transit service, while fixing the schedule of the commuter transit service. For the provision of better transfers, the developed optimization model allows for various schedule modification schemes. In contrast to other studies that focused on revising schedules by shifting whole timetables, this study also considers the changing of existing headways and the additional dispatching policy that further reduce transfer-related costs. Therefore, the developed model provides schedulers of transit services with various alternatives to revise their existing schedules. Although the design of the optimization schemes must be dependent on the decision and the policy of transit agencies, this flexibility of the model is very beneficial, facilitating the development of a timetable that maximizes convenience for transit users and simultaneously adheres to their policy.

The operation-stage approach – the real-time CP control – is a backup strategy intended to maintain coordinated transfers against unexpected delays. The developed CP model makes a holding decision based on the related passengers’ waiting times. Unlike previous studies, the model of this study incorporates the early dispatching of a receiving run, and estimates the downstream passengers’ cost through the downstream model, explaining the downstream operation of the local transit. The downstream model is also applicable to other transit-related problems, such as timetable evaluation and real-time transit controls (e.g., holding, stop skipping and short turning).
5 Recommendations for Future Studies

Future research should involve investigations of possible areas where the models developed in this study can be further improved and extended. Notwithstanding the sophistication of these models, the study needs to be enhanced in terms of the following points:

- Transit demand is assumed to be independent of the revision of a transit schedule. However, the revision of the existing schedule changes the level of the transit service and, consequently, may have an influence on the mode choice of overall travel demand. This dynamic aspect of the demand side according to schedule revision needs to be integrated in optimizing transfers.

- This study assumes that transfer passengers always wait for the next available transit if they miss their connecting run. However, the passengers may find other options to get to their final destinations. In order to incorporate these alternate modes into the transfer optimization model and the CP control model, another model is required explaining the behaviour of passengers and estimating the costs of the alternate modes. A separate future study is warranted to address this question.

- It is necessary to look more closely at different weights to reflect the relative disutility of the cost components of the cost function. Different weights can be assigned to the various components in order to improve the model. For example, Liu et al. found that the unit delay caused by missed transfers weighs more than twice the unit delay involved in regular stop waiting (Liu et al. 1998). It is also expected that one minute of delay incurred by oblivious passengers at downstream stops may be more onerous for such passengers than one minute of delay endured by the in-vehicle passengers at the transfer stop. Furthermore, the operators’ costs need to be added to the objective function with some conversion factors to unify the difference in attributes between the system costs and the waiting times of passengers.

- An extra cost associated with the missed connection of transfer passengers must be considered. It represents a time independent penalty resulting from the missed connection itself (Wirasinghe, 1993). If transit users fail to transfer to their intended run, for example, they will experience not only the delay of waiting for the next available
transit run, but also personal inconveniences such as missed appointments and disrupted schedules. Accordingly, the time independent penalty should be added to the cost functions of the transfer optimization model and the CP model. An additional study is required to decide the penalty and fuse it to the cost functions.

- In metropolitan areas, travel demand patterns have become more complex. Not only do suburban residents have to get back home from their job in the CBD, but an increasing number of urban residents have to get home from their jobs in the suburbs. Furthermore, there exist passengers transferring among local transit routes that this study does not take into account. The transfer optimization model should incorporate the growing trend of reverse commuting and local transfers, and needs to be modified in order to deal with two-directional transfer optimization. A hybrid model can be considered by adopting the timed transfer approach: a local transit can be coordinated by sharing a pulse headway, and then all the schedules of the local transit can be simultaneously shifted to be coordinated to commuter transit routes.

- The cost function of the developed transfer optimization model takes into account the transfer time at a transfer point located on a suburban area, and the stop waiting time of the non-transfer passenger. However, various delays exist throughout one transit trip, such as delays at the origin, one or more transfer points, and the destination. In extending the transfer optimization model, therefore, the overall delays of transit users should be modeled. As mentioned earlier, different weights should be applied to each delay to reflect the relative disutility of each delay.

- This study employs a genetic algorithm as the solution approach for its global search capability, and it demonstrates sufficient ability in solving the given optimization problem. However, the genetic algorithm finds solutions close to an optimal solution, but lacks the sophistication to find the optimal solution itself. Also, it is computationally expensive. Thus, improvement of the solution method should be considered in future research.

- The CP model focuses only on the case where the number of delayed feeder runs is one or two. The limitation of the number of delayed feeders should be relaxed in future studies.
There is a need for an additional study that utilizes a scheme where the CP model re-evaluates the holding decision as prediction values are updated. It should evaluate how much this re-evaluation can improve the benefit of the holding control.
References


Shih, M-C. (1994). “A design Methodology for Bus Transit Route Networks with Coordinated Operations.” Ph.D Dissertation, University of Texas, Austin, TX.


Appendices

Appendix 1  Notation (Chapter 3)

$AT_{i,j,0}$ : Arrival time of run $j$ on commuter transit route $i$ at transfer point

$AT_{L,n,k}$ : Arrival time of run $n$ on local transit route $L$ at stop $k$

$B_L$ : Number of transit units serving local transit route $L$

$C_i$ : Number of runs on commuter transit route $i$

$DT_{L,n,k}$ : Departure time of run $n$ on local transit route $L$ at stop $k$

$E(X)$ : Expected value (mean) of $X$

$F_X(X)$ : Cumulative density function (Normal distribution) of random variable $X$

$f_X(X)$ : Probability density function (Normal distribution) of random variable $X$

$G_X(X)$ : Cumulative density function (Lognormal distribution) of random variable $X$

$g_X(X)$ : Probability density function (Lognormal distribution) of random variable $X$

$HT_{L,n}$ : Holding time (in minutes) for run $n$ on local transit route $L$, (integer value)

$HT_{L,n}^{max}$ : Maximum of $HT_{L,n}$

$I$ : Number of commuter transit routes

$i$ : Index of commuter transit routes ($i=1,2,3,\ldots,I$)

$j$ : Index of runs on a commuter transit route ($j=1,2,3,\ldots,C_i$)

$K_L$ : Number of downstream stops on local transit route $L$
Index of stops on a local transit route ($k = 0, 1, 2, \ldots, K_L$): transfer point, if $k = 0$; downstream stops, if $k \geq 1$.

Index of a local transit route ($n(L) = 1$).

Number of runs on local transit route $L$.

Index of runs on a local transit route ($n = 1, 2, 3, \ldots, M_L$).

Number of aware stop passengers for run $n$ on route $L$ at stop $k$.

Number of in-vehicle passengers on run $n$ on route $L$ at stop $k$.

Number of stop passengers for run $n$ on local transit route $L$ at stop $k$.

Number of transfer passengers from run $j$ on commuter transit route $i$ to local transit route $L$ (PM-peak).

Number of transfer passengers from run $n$ on local transit route $L$ to commuter transit route $i$ (AM-peak).

Number of transfer passengers who aim for transferring from local transit route $L$ to run $j$ on commuter transit route $i$ (i.e., Real demand for run $j$ on route $i$ from route $L$) (AM-peak).

Number of transfer passengers from stop $k$ of local transit route $L$ aiming to transfer to run $j$ on commuter transit route $i$ (i.e., real demand for run $j$ on route $i$ from each stop of route $L$).

Number of transfer passengers from run $j$ on commuter transit route $i$ to the primary receiver on local transit route $L$ (PM-peak).

Number of transfer passengers from run $j$ on commuter transit route $i$ to the secondary receiver on local transit route $L$ (PM-peak).
\( P_{L,n,k}^U \) : Number of unaware stop passengers for run \( n \) on route \( L \) at stop \( k \)

\( SAT_{i,j,0} \) : Scheduled arrival time of run \( j \) on commuter transit route \( i \) at transfer point

\( SAT_{L,n,k} \) : Shifted scheduled arrival time of run \( n \) on local transit route \( L \) at stop \( k \)

\( SAT_{L,n,k}^O \) : Original scheduled arrival time of run \( n \) on local transit route \( L \) at stop \( k \)

\( SDT_{i,j,0} \) : Scheduled departure time of run \( j \) on commuter transit route \( i \) at transfer point

\( SDT_{L,n,k} \) : Shifted scheduled departure time of run \( n \) on local transit route \( L \) at stop \( k \)

\( SDT_{L,n,k}^O \) : Original scheduled departure time of run \( n \) on local transit route \( L \) at stop \( k \)

\( Skw(X) \) : Skewness value of \( X \)

\( Stdv(X) \) : Standard deviation of \( X \)

\( T_{L,n,k}^{A} \) : Waiting time of an aware stop passenger for run \( n \) on route \( L \) at stop \( k \)

\( T_{L,i,j}^{H} \) : Hidden waiting cost of transfer passengers who aim to transfer from local transit route \( L \) to run \( j \) on commuter transit route \( i \) (AM-peak)

\( T_{L,n,k}^{St} \) : Waiting cost of a stop passenger who starts trip at stop \( k \) through run \( n \) on local transit route \( L \)

\( T_{i,j,L}^{Tr} \) : Transfer cost of a transfer passenger from run \( j \) on commuter transit route \( i \) to local transit route \( L \) (PM-peak)

\( T_{L,n,i}^{Tr} \) : Transfer cost of a transfer passenger from run \( n \) on local transit route \( L \) to commuter transit route \( i \) (AM-peak)

\( T_{i,j,L}^{TrM} \) : Missed connection cost of a transfer passenger from run \( j \) on commuter transit route \( i \) to local transit route \( L \) (PM-peak)
$T_{L,n,i}^{TrM}$ : Missed connection cost of a transfer passenger from run $n$ on local transit route $L$ to commuter transit route $i$ (AM-peak)

$T_{i,j,L}^{TrS}$ : Successful connection cost of a transfer passenger from run $j$ on commuter transit route $i$ to local transit route $L$ (PM-peak)

$T_{L,n,i}^{TrS}$ : Successful connection cost of a transfer passenger from run $n$ on local transit route $L$ to commuter transit route $i$ (AM-peak)

$T_{L,n,k}^{U}$ : Waiting time of an unaware stop passenger for run $n$ on route $L$ at stop $k$

$TT_{L,n,k}$ : Travel time from Stop $k$ to Stop $k+1$ of run $n$ on local transit route $L$

$Var (X)$ : Variance of $X$

$\alpha_L$ : Offset value (in minutes) to shift all runs on local transit route $L$, (integer value)

$\alpha_L^{min}, \alpha_L^{max}$ : Minimum and maximum of $\alpha_L$

$\beta_{L,n}$ : Offset value (in minutes) to shift run $n$ on local transit route $L$, (integer value)

$\beta_{L,n}^{min}, \beta_{L,n}^{max}$ : Minimum and maximum of $\beta_{L,n}$

$\Delta$ : Boarding time per one passenger (3.5 sec. $\approx$ 0.0583 min.)

$\hat{\Delta}$ : Alighting time per one passenger (2.1 sec. $\approx$ 0.035 min.)

$\Delta_{L,n,k}^{St}$ : Boarding time of stop passengers to run $n$ on route $L$ at stop $k$

$\hat{\Delta}_{L,n,k}^{St}$ : Alighting time of run $n$ on route $L$ at stop $k$

$\Delta_{L,n,0}^{Tr}$ : Boarding time of transfer passengers to run $n$ at transfer point
\( \delta_{i,j,L,n} \): Primary connection flag (binary value): 1, if run \( n \) on local transit route \( L \) is the primary connected run of run \( j \) on route \( i \) (PM-peak)

\( \delta_{L,n,i,j} \): Primary connecting flag (binary value): 1, if run \( j \) on commuter transit route \( i \) is the primary receiver of feeder run \( n \) on local transit route \( L \) (AM-peak)

\( \delta_{S_{i,j,L,n}} \): Probability that transfer from run \( j \) on commuter transit route \( i \) to run \( n \) on local transit route \( L \) is successful (PM-peak)

\( \eta_{L,n,i,j} \): Local transit selection flag (binary value): 1, if transfer passengers destined to run \( j \) on commuter transit route \( i \) take run \( n \) on local transit route \( L \) (AM-peak)

\( \Lambda_{L,n,k} \): Dwell time of run \( n \) on local transit route \( L \) at stop \( k \)

\( \hat{\lambda}_{U_{L,n,k}} \): Arrival rate of unaware stop passengers for run \( n \) on route \( L \) at stop \( k \)

\( \hat{\lambda}_{A_{L,n,k}} \): Arrival rate of aware stop passengers for run \( n \) on route \( L \) at stop \( k \)

\( \nu_{L,i,j,k} \): Distribution rate of transfer passengers, \( P_{L,i,j}^{T} \), at stop \( k \). (AM-peak)

\( \tau_{L,n,k} \): Average arrival time in advance of \( SDT_{L,n,k} \) of aware passengers for run \( n \) on route \( L \) at stop \( k \)

\( \varphi_{L,i,j}^{*} \): Scheduled departure time of the commuter transit run aimed for by the transfer passengers who transfer from local transit route \( L \) to run \( j \) on commuter transit route \( i \) (i.e., \( \varphi_{L,i,j}^{*} = SDT_{L,i,j}^{T} \)) (AM-peak)

\( \varphi_{L,i,j} \): Scheduled departure time of the commuter transit run actually taken by the transfer passengers who transfer from local transit route \( L \) to run \( j \) on commuter transit route \( i \) (i.e., \( \varphi_{L,i,j} \leq \varphi_{L,i,j}^{*} \)) (AM-peak)

\( \omega_{L,n,k} \): Alighting rate of in-vehicle passengers on run \( n \) on route \( L \) at stop \( k \)
Appendix 2 Notation (Chapter 4)

$\hat{AT}_d$ : Predicted arrival time of delayed feeder run $d$

c_v : Coefficient of variation (the ratio of standard deviation to mean)

d : Index of delayed feeder runs

$DT_R$ : Departure time of run $R$

$D\hat{T}_R$ : Departure time of run $R$ when the CP control is provided for run 1 ($n(d)=2$)

$D\tilde{T}_R$ : Departure time of run $R$ when the CP control is provided for run 2 ($n(d)=2$)

$E(X)$ : Expected value (mean) of $X$

$f_X(X)$ : Probability density function (Normal distribution) of random variable $X$

$fun_{dns}$ : Downstream model (Section 3.4)

$HT_R$ : Holding time of run $R$

$HT^*_R$ : Optimal holding time of run $R$

$HT^{B^*}_R$ : Optimal holding time of run $R$ if the CP control is applied to run 1 and 2 ($n(d)=2$)

$HT^{1^*}_R$ : Optimal holding time of run $R$ if the CP control is applied to run 1 ($n(d)=2$)

$HT^{2^*}_R$ : Optimal holding time of run $R$ if the CP control is applied to run 2 ($n(d)=2$)

$P_{Tr}^{d,R}$ : Number of transfer passengers from run $d$ to run $R$

$P_{inv}^{R}$ : Number of in-vehicle passengers in run $R$ at $SDT_R$

$R$ : Index of receiving run

$SDT_R$ : Scheduled departure time of connecting run $R$ at a transfer point
$SDT_{R+1}$ : Scheduled departure time of the next available receiving run, run $R+1$

$Stdv(X)$ : Standard deviation of $X$

$T_{R}^{Dwn}$ : Waiting cost of downstream passengers waiting for run $R$ at the downstream stops of the transfer point

$T_{R}^{Dwn,1}$ : Waiting cost of downstream passengers waiting for run $R$ at the downstream stops of the transfer point when the CP control is provided for run 1 ($n(d)=2$)

$T_{R}^{Dwn,2}$ : Waiting cost of downstream passengers waiting for run $R$ at the downstream stops of the transfer point when the CP control is provided for run 2 ($n(d)=2$)

$T_{R}^{InV}$ : Waiting cost of in-vehicle passengers in run $R$

$T_{R}^{InV,1}$ : Waiting cost of in-vehicle passengers in run $R$ when the CP control is provided for run 1 ($n(d)=2$)

$T_{R}^{InV,2}$ : Waiting cost of in-vehicle passengers in run $R$ when the CP control is provided for run 2 ($n(d)=2$)

$T_{R}^{Total}$ : Total cost of all passengers related to the CP control

$T_{R}^{Total,1}$ : Total cost of all related passengers when the CP control is provided for run 1 ($n(d)=2$)

$T_{R}^{Total,2}$ : Total cost of all related passengers when the CP control is provided for run 2 ($n(d)=2$)

$T_{d,R}^{Tr}$ : Transfer cost of transfer passengers from run $d$ to run $R$

$T_{d,R}^{Tr,1}$ : Transfer cost of transfer passengers from run $d$ to run $R$ when the CP control is provided for run 1 ($n(d)=2$)

$T_{d,R}^{Tr,2}$ : Transfer cost of transfer passengers from run $d$ to run $R$ when the CP control is provided for run 2 ($n(d)=2$)
\[ T^{TrS}_{d,R} \] : Successful connection cost of transfer passengers from run \( d \) to run \( R \)

\[ T^{TrM}_{d,R} \] : Missed connection cost of transfer passengers from run \( d \) to run \( R \)

\( \mathcal{E}_X \) : Prediction error in predicted value \( X \)

\( \Delta^{Tr}_{d,R} \) : Boarding time of the transfer passengers from run \( d \) to run \( R \)

\( \Delta \) : Boarding time per one passenger (3.5 sec. \( \approx 0.0583 \) min.)

\( \sigma_X \) : Standard deviation of predicted value \( X \)