HEADWAY CONTROL SCHEMES TO RESIST BUS BUNCHING

A Thesis
Presented to
The Academic Faculty

by

Zhihao Ding

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Industrial and Systems Engineering

Georgia Institute of Technology
May 2016

HEADWAY CONTROL SCHEMES TO RESIST BUS BUNCHING

Approved by:

Professor John J. BARTHOLDI, III, Advisor
School of Industrial and Systems Engineering
Georgia Institute of Technology

Professor Donald D. EISENSTEIN
Booth School of Business
The University of Chicago

Professor George Nemhauser
School of Industrial and Systems Engineering
Georgia Institute of Technology

Professor Martin Savelsbergh
School of Industrial and Systems Engineering
Georgia Institute of Technology

Professor Alan Erera
School of Industrial and Systems Engineering
Georgia Institute of Technology

Date Approved: February 25, 2016

I dedicate this thesis to

my family,

my advisors, and

those who might find it interesting or useful.

ACKNOWLEDGEMENTS

I would like to thank my father and my mother for their love. Although they cannot understand what I have studied during these years, they have been trusting me and supporting me unconditionally.

I would like to thank my advisors, Prof. Bartholdi and Prof. Eisenstein. They have been giving me maximum freedom, enough funding and great encouragement for my research. They showed me ways to carve excellent work. They shared with me special ideas to attack new problems. Without them, I cannot finish my PhD study so fast.

I would like to thank the other committee members, Prof. Erera, Prof. Nemhauser and Prof. Savelsbergh for providing me valuable feedback on my research and my performance in their classes.

I would like to thank Paul Vazquez, Brian Mann, Brian Herman, Chenxi Zeng, Carl Morris and Ryan Kaltsukis for providing guidance and help at Amazon. For the same reason, I would like to thank Joe Salvo, Walter Yund, Peter Koudal, John Carbone, Dan Yang, Benjamin Beckmann, Xing Wang and Jin Xia at GE Global Research. They showed me how to better apply what I have learned at school to the work in industry.

I would like to thank all my friends, roommetes and classmates at Georgia Tech for bringing me fun, joy and help during these years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Bus bunching occurs when two or more buses travel head to tail. It is an annoying problem in public transportation because it increases passengers' average waiting time and traveling time, wastes bus capacity, reduces the frequency of bus service and increases the pressure on bus drivers. So eliminating bus bunching is important in public transportation.

Eliminating bus bunching is highly challenging due to the complexity and variability of the bus dynamics. Bus bunching results from a positive feedback mechanism of headway evolution, which is a flaw born with the bus system. In this thesis, we quantify the intensity of the tendency to bus bunching and propose a headway control modeling framework to reverse tendency. Our framework subsumes many headway control schemes to coordinate buses and so enables batch analysis. Given different headway information, our framework produces different control schemes under which headways self-equalize. The stability of the bus system under control is characterized by a single measure and it can be optimized. Besides, the bus system under control is robust against traffic conditions and the level of ridership.

The framework is based on a snapshot model capturing the bus dynamics including the tendency to bunch by taking traffic conditions and the level of ridership into account. It is linear and time-invariant, which makes the bus dynamics tractable. This model considers a single control point and constant bus velocity in a deterministic manner, but it can be extended to handle many control points, inhomogeneous velocity along the route, and randomness.

Using our framework, we further study two simple control schemes—Threshold control and "Prefol". Threshold control drives headways to self-equalize the fastest

but the corresponding bus system needs large slack time for robustness. "Prefol" needs small slack time but headways self-equalize more slowly. We hybridize them and find the hybrid control scheme balances robustness and fast headway equalization. We also show that it outperforms several state-of-the-art control schemes in tests on a simulated bus route in Chicago.

CHAPTER I

INTRODUCTION

Buses are scheduled to be evenly spaced along a bus route, but in fact two or more of them often travel in bunches. This is the notorious bus bunching phenomenon.

The mechanism of bus bunching has been long understood [49]. Even though some buses will inevitably become late or early due to disturbance caused by variability of traffic conditions and level of ridership, the root cause is a positive feedback mechanism born with the bus system: A late bus picks up the passengers that should have been picked up by the following bus and thus gets further delayed; Meanwhile, the following bus spends less boarding time so it tends to catch up until the two buses pair up. It is the positive feedback mechanism that makes the bus system unstable.

Bus bunching increases passengers' average waiting time and traveling time. When buses travel in bunches, some of the buses become late. Since more passengers get on late buses, more passengers spend more time waiting. Besides, these buses travel more slowly because of more boarding and alighting time for the increased number of passengers. So more passengers spend more time traveling.

Bus bunching wastes bus capacity. While late buses pick up most passengers, trailing buses are almost empty. Operating these buses wastes fuel and drivers' work hours.

Bus bunching reduces the frequency of bus service. When buses travels in pairs, the frequency halves and thus the bus service will be degraded for passengers.

Bus bunching increases pressure on bus drivers. When their buses are late, they may experience hostility from passengers. Pressure to recover the schedule may create safety issues. Drivers in some bus systems are even monitored and told to speed up and slow down from time to time while they are driving. Some drivers complain that this heavily-mediated style brings extra pressure and triggers anger from passengers [54].

If buses can maintain equal headways, passengers will receive the best service and all the defects mentioned above are maximally mitigated. Therefore, our goal is to maintain equal headways effectively while minimally interfering with the drivers.

1.1 Literature Review

The problem of bus bunching remains a challenge in public transportation [53]. The basic idea to separate buses that are too close is to speed up the leading bus or slow down the trailing bus. However, buses that want to speed up may be blocked by the traffic and buses that want to slow down will annoy the following traffic and on-board passengers. There are other indirect ways to achieve this. A bus can, in effect, speed up by skipping stops [20, 30, 63, 64], restricting boarding [23, 24], deadheading [28] and short-turning [38, 50, 58] and a bus can, in effect, slow down by delaying at designated stops.

Boarding restriction, stop-skipping, short-turning, and deadheading are not desired by many transit agencies because they inconvenience passengers.

Boarding restriction is used to limit the number of boarding people to reduce dwell time, but it increases wait time for passengers that cannot get on the bus.

Stop-skipping refers to a bus not making all designated stops along a route. It serves to reduce the traveling time for passengers on board the vehicle, but it increases waiting time for passengers that have been passed by and passengers who are forced to alight early. Lin et al. [40] recommended against stop-skipping based on simulation results.

Short-turning refers to a bus turning around before it reaches the route terminal. It may reduce headway variance in the opposite direction by filling in a big gap in service. However, it affects passengers on the bus who are forced to alight and transfer to the subsequent bus.

Deadheading refers to a bus traveling to a certain position without accepting passengers. If the positioning is optimized, it reduces headway variance. But deadheading is inefficient. It incurs costs for the operator in terms of non-revenue earning fuel use, wages, and a reduction in the utilization of the driver's legal hours of driving.

In contrast, delaying buses at certain bus stops (referred to as "control points") is widely used in practice. The control points are typically located at terminals or transfer points where there is large passenger demand, few on-board passengers and enough spare spaces (e.g. bus bay, bus lane, parking lane, etc.). Delaying buses at these control points has little effect on either the traffic or passengers. The implementation is easy compared with other methods mentioned above. In this thesis, we will focus on bus-delaying control schemes. The core of these schemes is the determination of the delay time.

There are roughly four categories in bus delaying control schemes: target schedule control, target headway control, optimization-based headway control and self-organizing headway control.

Target schedule control aims at keeping buses on a pre-determined schedule. A bus is delayed longer at a control point if it is ahead of the schedule or shorter (or even not delayed) if it is behind the schedule.

Target headway control does not have a schedule. Instead, it focuses on regulating headways to achieve a pre-specified static value. A bus is delayed longer at a control point if its headway is smaller than the target headway or shorter (or even not delayed) if its headway is larger than the target headway.

Optimization-based headway control optimizes the headways typically by minimizing some cost function related to passengers' waiting time and variance of headways based on real-time or forecasting information.

Self-organizing headway control specifies only a simple rule for delaying buses. Under self-organizing headway control headways self-equalize and converge to the ideal achievable headway, which is unknown in advance.

In the following, we review the four categories of bus delaying control schemes and identify the best application scenario for each of them.

1.1.1 Target Schedule Control

The goal of target schedule control is to reach and maintain a target schedule. The arrival times and/or departure times of buses at all control points are planned to the minute. Slack time is budgeted in the schedule to help late buses catch up. A bus is delayed longer at a control point if it is ahead of the schedule or shorter (or even not delayed) if it is behind the schedule.

A schedule is particularly useful in low-frequency transit: If the schedule can be maintained, passengers can make travel plans based on the schedule and reduce waiting time. Target schedule control is simple to implement.

The limitation of target schedule control is its rigidity. Schedules are hard to maintain, hard to recover, and expensive to change. A target schedule is static, while the environment can be dynamic. The environment may change so much that the buses cannot be on time. Then the schedule exists in name only and the service discourages the passengers. Also, if a bus breaks down, the "hole" in the schedule tends to grow, leading to bunching.

There are two types of target schedule control schemes: binary target-schedule control, which decides only whether to delay a bus at a control point until the scheduled departure time, and continuous target-schedule control, which always delays the bus but the delay time is a continuous function of the arrival times and the scheduled departure times. The studies of the former focus on slack time determination and control point selection while those of the latter explore the relation between the delay time and the arrival deviation from the schedule. The former is simpler and widely used but the latter requires less slack time to provide the same level of schedule adherence [74].

Binary Target-schedule Control

Binary target-schedule control works as follows: When a bus arrives at a control point before the departure time, it pauses until that time; otherwise it departs immediately. The binary target control is commonly used in practice for its simplicity.

"Slack time" is the difference between the scheduled departure time and the expected arrival time at a control point. Slack time is added to the schedule to help late buses catch up. Too large a slack time reduces the service frequency unnecessarily when most buses are ahead of schedule while an insufficient one increases the risk of bus falling behind schedule.

Due to the complicated nature of the bus dynamics, studies on the determination of the slack time for the binary target-schedule control have been limited in what they can achieve. They either offer insights based on strong assumptions or else model the complicated dynamics but provide no insights.

Newell [47] is an example of the first type. He built a model based on five assumptions:

1. The travel times between adjacent control points are identically distributed random variables,

2. The bus headways are scheduled to be uniform,

3. The passenger arrivals at each control point are identically distributed Poisson variables,

4. The lateness of buses at a control point has a probability density that satisfies a Fokker-Planck (diffusion) type equation,

5. The preceding bus is always on time.

He showed that the bus departure times are unlikely to deviate significantly from the schedule if the slack time exceeds $\sigma\sqrt{2a_r(\tau' + \tau'' e^{-a_r h})}$, where

$\sigma =$ standard deviation of bus travel time between adjacent control points,

$a_r =$ mean arrival rate of passengers at a control point,

$\tau' =$ marginal time taken to board one passenger,

$\tau'' =$ time spent on stopping and starting a bus, and

$h =$ headway between buses.

This lower bound on slack time increases with standard deviation of bus travel time, passengers' arrival rate, and boarding time. One limitation of this study is that the argument holds only in the highly specialized setting satisfying all the assumptions above. Another limitation is that the study considers only schedule adherence. The slack time may inflate the headway and thus discourage passengers.

Wirasinghe [72] derived optimality conditions for slack time in a model with the objective of minimizing the sum of the cost of the travel time for the trip, the cost of expected delay and a penalty associated with a delayed arrival. The result is simple and clear but the assumptions in the model are very strong. The model considers a route where passengers board at a terminal and buses run non-stop to the destination. Furthermore the analysis was for a single bus whose preceding bus runs exactly as scheduled. It assumes that the travel times of the bus, including dwell times, are independent and identically distributed with a unimodal probability density function. Such strong assumptions might be satisfied only in a one-control-point low-frequency route with stable traffic conditions, but it seems doubtful whether the result can be extended to a more general case.

By elegantly utilizing queueing theory, Zhao et al. [77] first showed the stability of the bus system under binary target-schedule control if the scheduled round-trip travel time is greater than the expectation of real round-trip travel time. This result is based on the assumption that the round-trip travel times are independent, which is more likely to hold in low-frequency route that experiences stable traffic intensities (for example, no rush hour). They also proposed an analytical method to derive the optimal slack time for a one-bus route when the travel time is exponentially distributed, as well as several approximation approaches for a multi-bus route and other travel time distributions to minimize passengers' waiting time. However, the objective function is derived from the assumption that passengers arrive randomly at the control point, which is not true for low-frequency route. This assumption is not consistent with the previous assumption to some extent and thus they are unlikely to hold simultaneously.

Instead of generating analytical insights, Carey [18] focused on demonstrating the feasibility of a modeling approach. He performed a comprehensive study on a transit system with multiple vehicles and multiple control points. His model took the randomness of the traveling times and waiting times into consideration. The bus dynamics are described by a set of integral equations. As a result, the model is general but too complicated to generate deep insights or even to solve for the optimal slack time efficiently.

Some studies tried to determine both the slack time and the control point simultaneously by optimizing certain objectives, but this problem is not easy to solve. Researchers either introduced simplified assumptions or tried to solve it using heuristics. For example, Wirasinghe and Liu [73] built an analytical model which minimizes the sum of waiting time cost, traveling time cost, delay penalty and operating cost. They used dynamic programming to solve simultaneously for the optimal location of control points and the amounts of slack time. However, the model considers only a very special case where all boarding passengers coordinate their arrivals at each stop in such a way that they never miss their intended bus. This can hardly be true. Also, it is tricky to determining the weights of different cost terms.

Liu and Wirasinghe [41] adopted the same objective as that in [73] but considered a general passenger arrival pattern in a simulation model. This model is more realistic and may be useful for schedule design. They tried to solve it using a combination of heuristic search, enumeration, and population ranking and selection techniques but cannot guarantee optimality. Mazloumi et al. [43] expressed this problem as a knapsack problem for which input data was derived from simulation, and applied an ant colony algorithm and a genetic algorithm to solve it. Their approach is more efficient in computation.

Continuous Target-schedule Control

While binary target-schedule control decides only whether to delay a bus at a control point until the departure time, continuous target-schedule control always delays the bus but the delay time is a continuous function of the arrival times and the scheduled departure times.

Xuan et al. proposed a parametric family of continuous target-schedule control schemes. They also showed a sufficient condition to guarantee bounded deviations from the schedule over time [74]. This study analytically addressed a broad range of problems: transit systems with many buses, many control points and stochastic traveling time. They also singled out a one-parameter control scheme that relies only on the scheduled departure times and the arrival times of the current bus and its preceding bus at the current control point. If an upper bound on the deviation from the schedule is pre-specified, they are able to calculate the optimal parameter by minimizing the slack time. They showed that this method requires less slack time than the binary target-schedule control to provide the same level of schedule adherence in a numerical study. A similar control scheme proposed by He [33] utilized the arrival time of the previous bus at the next control point additionally and showed that the bus system needs less slack time to maintain the schedule.

However, these schemes all try to achieve a target schedule, and so they inherit the limitations of a static schedule. A static schedule cannot adapt to the dynamic environment, and so all target schedule control schemes are vulnerable to disturbances, especially system-wide disruptions. Also, they cannot respond to bus break-down.

1.1.2 Target Headway Control

When realized headways are below 12 minutes, passengers tend to not care about a schedule and arrive randomly at the bus stops [51]. Some researchers abandoned the notion of a schedule and aim at maintaining a pre-determined static target headway. A bus is delayed longer at the control points if its headway is smaller than the target headway or shorter (or even not delayed) if its headway is larger than the target headway.

The information required by target headway control is mild—only the arrival times at the control points. And generally target headway control outperforms schedule-based control in terms of service regularity. Target headway control is less ambitious than target schedule control and therefore more likely to achieve its goals. Every bus may be "late" but the headway may nevertheless be on target.

Target headway control shares weaknesses similar to those of target schedule control. The tricky part in applying target headway control is the determination of target headway. The ideal headway changes with traffic conditions, weather, level of ridership and habits of the driver. Consequently, any system that coordinates buses based on static target headway must sometimes underestimate achievable headway, and so fail to meet the target, and sometimes overestimate it, and so reduce service frequency unnecessarily.

Another similar weakness of target headway control is that it cannot respond adequately to serious disruption. For example, when a bus breaks down, it leaves a gap until a replacement can be inserted. When the gap is large enough, it will overwhelm any planned slack. Also when there is a system-wide disruption, such as a snow storm, that reduces the bus velocity, all headways will necessarily increase and so the system will fail to meet its targets. When the disruption is large, most headways are greater than the target headway, so there is no delay at all for most buses, leaving the target headway control existing only in name.

Similar to target schedule control, two kinds of target headway control have been proposed: binary target-headway control and continuous target-headway control. Binary target-headway control, also called “Threshold Control” or “Threshold Strategy” in the literature, decides only whether to delay a bus at a control point until its headway equals to a threshold while the continuous headway control always delays the bus but the delay time is a continuous function of the headways. The latter has more flexibility in control. It may use more headway information than the former and thus has the potential to produce smaller headway variance.

Binary Target-headway Control

In binary target-headway control, the bus at the control point is not delayed if its forward headway is larger than a predetermined static threshold which may be different from the target headway, otherwise it is delayed until its forward headway equals the threshold. There are a number of studies about the determination of the threshold. The following studies use analytical models that assume the threshold equals the target headway.

Osuna and Newell [52] considered a bus route with only one or two buses. Traveling times are assumed independent and identically distributed. They derived optimal thresholds that depend on the distribution of the traveling time to minimize passengers’ waiting time.

Barnett [9] concentrated on the case that all passengers board at the same stop, alight at the same stop and traveling times are independent and identically distributed. He derived the optimal threshold to minimize service irregularity.

Turnquist and Bowman [68] considered a bus route with multiple buses and multiple control points and sought to minimize average waiting time of passengers. They derived a threshold value based on the ratio of onboard passengers to the boarding demand at the control point according to daily statistics. This solution is sub-optimal in the model.

Due to the complicated nature of the real bus dynamics, simulation models [15, 31, 38, 40, 70] were developed to test different thresholds. Some of these models tested thresholds that are different from the target headway. All of these studies concluded that an increase in the threshold can reduce passengers' waiting time at the expense of longer traveling time.

Continuous Target-headway Control

Continuous target-headway control always delays the bus and the delay time is a continuous function of the headways. Daganzo proposed a parametric family of continuous target-headway control schemes and a sufficient condition on the parameters to bound variance of the headways over time [21]. This study analytically addressed a broad range of problems: transit systems with many buses, many control points and stochastic traveling time. Daganzo also singled out a control scheme that only relies on the forward headway of the bus at the control point (referred to as "forward headway control scheme" later) and showed that this control scheme requires less slack time than binary target schedule control to provide the same level of schedule adherence.

1.1.3 Optimization-based Headway Control

Optimization-based headway control determines delay time by optimizing a certain objective based on a forecast (deterministic or stochastic) of headways and relevant information for a certain number of periods in the future. Table 1 presents different optimization-based headway control schemes according to the following characteristics:

1. Passenger demand (PD) and bus traveling times (TT)
2. Number of control points
3. Number of buses considered in the optimization process
4. Bus capacity
5. Objective function, that could include waiting time of passengers at stops (W_s), waiting time for passengers aboard a bus being delayed at a control point (W_o), extra waiting time of passengers that cannot board the first bus arrived (W_e) or the total variance of headways between buses (V)

Optimization-based headway control is adaptive to the environment. Since it assumes knowledge of the future, it is able to make anticipatory adjustment no matter how the environment changes. However, this approach is problematic in its impractical requirements, unconvincing artificial objective functions, and nonintuitive solutions.

All the schemes listed in Table 1 require accurate forecast of bus arrival times and rates of passenger arrival. Such requirements cannot be guaranteed, and it is doubtful that these schemes can remain effective with inaccurate information.

Many of the schemes (e.g. [35, 55, 65, 76] and [78]) consider objective functions as the sum of several weighted terms, but it is unclear which term should be more valued and assigned larger weight, nor is there any agreement or convincing evidence.

Table 1: Optimization-based headway control

Studies	PD and TT	Control points	Buses	Cap.	Obj.
Eberlein et al., 2001 [29]	Deterministic	One	Multiple	$=\infty$	W_s
Ding and Chien, 2001 [27]	Deterministic	Multiple	One	$=\infty$	V
Hickman, 2001 [35]	Stochastic	One	One	$=\infty$	$W_s + W_o$
Zhao et al., 2003 [76]	Stochastic	Multiple	One	$=\infty$	$W_s + W_o$
Sun and Hickman, 2008 [65]	Deterministic	Multiple	Multiple	$=\infty$	$W_s + W_o$
Zolfaghari et al, 2004 [78]	Deterministic	One	Multiple	$<\infty$	$W_s + W_e$
Puong and Wilson, 2008 [55]	Deterministic	Multiple	Multiple	$<\infty$	$W_s + W_0 + W_e$
Hernández et al., 2015 [34]	Deterministic	Multiple	Multiple	$<\infty$	V

1.1.4 Self-organizing Headway Control

Self-organizing headway control is simple and adaptive. It uses a simple rule to determine delay times at the control points. Under the rule the headways are able to self-equalize and so adapt to the environment. This rule makes an unstable bus system stable.

Bartholdi and Eisenstein [11] first proposed a self-organizing headway control scheme, under which the delay time of a bus at the control point is a proportion of its backward headway (the headway of its trailing bus). We call it "backward headway control scheme" in this thesis. They showed in an idealized model that headways spontaneously converge to an ideal headway, even though it is unknown in advance. Bus drivers may even be unaware of this process. When traffic conditions change, or even under system-wide disruption, the bus system under this control scheme is still able to direct headways to the new ideal headway. This method has been successfully implemented on a bus route in Atlanta [10].

Turnquist et al. [68] proposed a bus-holding scheme called "Prefol", in which the delay time of the bus at a control point depends linearly on its forward and backward headways. This control scheme was originally used to minimize a weighted sum of on-board and off-board passengers' average waiting times at each control point. Turnquist et al. did not study the overall behavior of the bus system, or how the headways evolve. Daganzo and Pilachowski proposed a strategy that uses both headways, but they concentrate on adjusting bus velocity continuously instead of delaying buses at control points [22]. Xuan et al. studied the delaying strategy and showed that the "Prefol" control scheme leads to bounded deviation from a target headway if bounded randomness of traveling time of buses is considered [74]. Later in this thesis we shall prove that "Prefol" is self-organizing.

Berrebi et al. [14] suggested a scheme that is designed to allow achievable headways to emerge, like a self-organizing scheme, but they did not prove convergence. This scheme requires the joint distribution of all the headways. Its performance largely depends on the accuracy of the forecasting accuracy of headways.

1.1.5 Summary

Target schedule control, target headway control, optimization-based headway control and self-organizing headway control all have their pros and cons. Each of them may be useful in certain situations.

1. If the route frequency is low, target schedule control is favored. Such schemes are widespread use because they provide passengers with a timetable for planning. If the frequency is high (e.g. average headways are less than 12 minutes), buses become more susceptible to bunching and the usefulness of a timetable fades.

2. If the route frequency is high and the environment is stable during a day, target headway control may be practical. Indeed it has been successfully implemented [4] because is easy to understand and cheap to implement. More importantly, the ideal target headway can be estimated accurately due to the stable environment.

3. Optimization-based headway control requires accurate forecasts and produces nonintuitive solutions. To the best of the author's knowledge, none of the optimization-based headway control schemes have been implemented.

4. If the route frequency is high and the environment is changeable, self-organizing headway control is promising for its simplicity, adaptivity and cost-effectiveness.

They also have different data requirements. Both target schedule control and target headway control requires only the arrival times of the buses at the control point, which can be obtained by a clock. Self-organizing headway control requires some or all of the real-time headways, while optimization-based headway control may need extra information like passengers' arrival rate.

Essentially, these control schemes attack the problem at different depths. The root cause of bunching is the positive feedback mechanism in the evolution of headways. Optimization-based headway control aims to relieve the symptoms by global monitoring and adjustment. Target schedule control and target headway control attempt to suppress the positive feedback mechanism by imposing constraints on headways, so they lack flexibility; and even then the positive feedback mechanism may be too strong to be suppressed. Self-organizing headway control creates a negative feedback loop to counter the positive one so that the bus system becomes stable. The effect may not be immediate, but it addresses the root cause of bus bunching.

1.2 Our Contributions

In this thesis, we extend the study of Bartholdi and Eisenstein [11] on self-organizing headway control. We propose a unified headway control framework based on a snapshot model that captures the bus dynamics including the tendency to bunching (See Chapter 2 for the model). Our framework subsumes many headway control schemes to coordinate buses including most of the target headway control schemes and self-organizing control schemes. We summarize the properties of our framework in the following which will be shown in Chapter 3:

1. The schemes produced by our framework are simple. They use only headway information. They delay buses only at control points. Drivers are not distracted to check the time, bus position or velocity. They just concentrate on driving buses and serving passengers (See Section 3.1).

2. Our framework is general and flexible. The form of the delay time can adapt to different amounts of headway information (See Section 3.1).

3. Under any control scheme produced by our framework, headways self-equalize

spontaneously. This process does not require direction of management or intention or even awareness of the drivers (See Section 3.2).

4. The bus system under control is robust against the level of ridership and traffic conditions (See Section 3.3 and Section 3.4).

The demonstration of these properties is based on a measure of systemic stability we identify (See Section 3.2.2). It essentially determines whether bus bunching would occur and how severe it would be. It can be used to compare different self-organizing bus systems.

In Chapter 4, we use our framework to further study two simple control schemes—Threshold control and "Prefol". Threshold control drives headways to self-equalize the fastest but the corresponding bus system needs large slack time for robustness. "Prefol" needs small slack time but headways self-equalize more slowly. We hybridize them and find the hybrid control scheme balances robustness and fast headway equalization. We also show that it outperforms several state-of-the-art control schemes in tests on a simulated bus route in Chicago.

CHAPTER II

BUS DYNAMICS

In this chapter, we develop a snapshot model that captures bus dynamics. It is specially designed for headway control schemes that delay buses at control points. It captures the most important factors that influence the bus dynamics: traffic conditions and the level of ridership. It has the mechanism that produces bus bunching

> A delayed bus picks up the passengers that should have been picked up by the following bus and thus gets further delayed. Meanwhile, the following bus spends less time on boarding passengers so it tends to catch up until the two buses pair up.

This model is linear, so it is simple and tractable. The form of the model shows how headways evolve in a clear manner and how they are influenced by traffic conditions, level of ridership and bus delays at the control points directly.

The majority of studies regarding bus bunching are empirically based or focus only on minimizing some objective for a bus at a control point. Recently, there are several models that capture how a bus system evolves under headway control schemes. Daganzo proposed a model in 2009 that captures the main characteristics of bus systems [21]. But his model is based on a schedule which is not necessary for headway control. Daganzo and Pilachowski came up with a continuum model that assumes infinitely divisible passengers and continuous passenger boarding along every point of the bus route [22]. This model simplifies boarding behavior reasonably and thus is easier for analyzing performance of bus systems. However, it is specially designed for velocity adjusting strategies instead of bus delaying. Bartholdi and Eisenstein developed an idealized model [11] that assumes constant commercial speed which includes the consideration of bus traveling and boarding. It is so simple for further systemic analysis. However, it does not contain the tendency of bus bunching even if there is no control at all. It cannot be used to identify what contributes to bus bunching and how well a bus system under control resist bunching.

We extend the model of Bartholdi and Eisenstein by taking the bus bunching mechanism into consideration, so the new model not only inherits its simplicity, but enables the analysis of bunching resistance of bus systems under control as well. Besides, our model is compatible with control schemes that either pre-determine a target headway or not. Although the model considers only one control point, it can be extent to multiple control points when we analyze the properties of headway control schemes.

In the following, we describe an idealized model without dwell time first (Section 2.1). This model has the same form of that of Bartholdi and Eisenstein but has different assumptions. Then we incorporate deterministic dwell time into the idealized model (Section 2.2). This new model can explain bus bunching. Then we discuss the extensions of this model.

2.1 Bus Dynamics without Dwell Time Delay

In this section, we build a model that assumes constant traveling speed and ignores dwell time. This model shows the basic dynamics of a bus system under any bus delaying control.

Modeling

Consider a bus route with n buses and one control point. Temporarily assume that buses have identical constant average velocity v along the route and dwell time is ignored.

We normalize the length of the route to be 1. When a bus arrives at the control point, a new epoch starts. Denote the index of the epoches t, $t \in \{1, 2, 3, \ldots\}$. At epoch t, index 1 the bus arriving at the control point and the others sequentially in the traveling direction. Since the bus route is a close loop, the bus following Bus 1 is Bus n. For epoch t, denote by the vector $\mathbf{x}^t = (x_1^t, x_2^t, \ldots, x_n^t)$ the positions of the buses where $x_1^t = 0$ and $x_i^t < 1$, $i = 2, \ldots, n$. For convenience of formula derivation to appear later in this section, let $x_{n+1}^t = 1$. The trajectory of bus positions $\{\mathbf{x}^0, \mathbf{x}^1, \ldots\}$ can be regarded as a series of snapshots of the bus route at the beginning of each epoch. This "snapshot" setup is most suitable for headway control, because delaying a bus at the control point is essentially an impulse control occurring only at the beginning of each epoch.

The forward headway of a bus is the time it moves from its current position to the current position of its preceding bus. It consists of three parts: traveling time, dwell time and delay time at the control point which is 0 except for Bus 1. Denote the traveling time of Bus i, $i = 1, 2, \ldots, n$, traveling from x_i^t to x_{i+1}^t by s_i^t. We call it travelling time in this thesis. It holds that

$$s_i^t = \frac{(x_{i+1}^t - x_i^t)}{v}, \quad \forall i = 1, 2, \ldots, n.$$

Essentially, they are the distances between buses normalized by the traveling velocity. Denote the headway of Bus i, $i = 2, 3, \ldots, n$ at epoch t by h_i^t. Also denote by h_1^t the headeway of Bus 1 before delay time is assigned. Since dwell time is ignored in this case, headways are simply traveling times. So we have

$$h_i^t = s_i^t, \quad \forall i = 1, 2, \ldots, n.$$

Let D^t be the delay time of the bus at the control point. And let the actual headway of Bus 1 after delay time is assigned be $\widehat{h}_1^t$. Then $\widehat{h}_1^t = h_1^t + D^t$. h_1^t may be used to determine the delay time and $\widehat{h}_1^t$ is the real headway.

The length of an epoch is the amount of time between successive arrivals at the control point. So the length of Epoch t, $t = 1, 2, \ldots$, is the headway of Bus n at Epoch

t, h_n^t. During Epoch t, Bus i, $i = 2, 3, \ldots, n-1$, spends all the time for traveling due to the ignorance of dwell time. So at the beginning of Epoch $t+1$, Bus i at Epoch t becomes Bus $i+1$ and its position becomes

$$x_{i+1}^{t+1} = x_i^t + h_n^t v.$$

Bus n at Epoch t arrives at the control point at Epoch $t+1$. It then becomes Bus 1 at Epoch $t+1$ and its position becomes

$$x_1^{t+1} = 0.$$

Since Bus 1 at Epoch t is delayed at the control point for D^t amount of time, it only spends $h_n^t - D^t$ amount of time on traveling. At the beginning of Epoch $t+1$, it becomes Bus 2 and its position becomes

$$x_2^{t+1} = x_1^t + (h_n^t - D^t)v.$$

So we have the following bus dynamics in terms of bus positions:

$$x_1^{t+1} = 0 \tag{1}$$

$$x_2^{t+1} = x_1^t + (s_n^t - D^t)v \tag{2}$$

$$x_i^{t+1} = x_{i-1}^t + s_n^t v \quad \forall i = 3, \ldots, n. \tag{3}$$

From these equations, we can obtain the dynamics equations expressed in terms of s_i^t and h_i^t, $\forall i = 1, 2, \ldots, n$, $\forall t = 1, 2, \ldots$. Using Equation (3) we can write for each $i = 3, \ldots, n$,

$$\begin{aligned} s_i^{t+1} &= \frac{x_{i+1}^{t+1} - x_i^{t+1}}{v} \\ &= \frac{(x_i^t + s_n^t v) - (x_{i-1}^t + s_n^t v)}{v} \\ &= \frac{x_i^t - x_{i-1}^t}{v} \\ &= s_{i-1}^t \end{aligned}$$

and

$$
\begin{aligned}
s_2^{t+1} &= \frac{x_3^{t+1} - x_2^{t+1}}{v} \\
&= \frac{(x_2^t + s_n^t v) - (x_1^t + (s_n^t - D^t)v)}{v} \\
&= \frac{x_2^t - x_1^t + D^t v}{v} \\
&= s_1^t + D^t
\end{aligned}
$$

and finally,

$$
\begin{aligned}
s_1^{t+1} &= \frac{x_2^t - x_1^t}{v} \\
&= \frac{(s_n^t - D^t)v}{v} \\
&= s_n^t - D^t
\end{aligned}
$$

Thus we have the system

$$
\begin{aligned}
s_1^{t+1} &= s_n^t - D^t && (4) \\
s_2^{t+1} &= s_1^t + D^t && (5) \\
s_i^{t+1} &= s_{i-1}^t \quad \forall i = 3, \ldots, n. && (6) \\
h_i^t &= s_i^t, \quad \forall i = 1, 2, \ldots, n. && (7)
\end{aligned}
$$

Equation (4) (5) (6) describe the underlying dynamics. Equation (7) shows the relation between headways and traveling times. Let I be the index set of the headways we know. Then $h_i^t, \forall i \in I$ is the information we can use for control.

Comments

The model we develop in this section has the same form of that of Bartholdi and Eisenstein [11]. The difference is that they assume constant commercial speed which includes the consideration of bus traveling and boarding while we separate traveling and boarding and ignore boarding at this moment.

This model weakens the concept of bus stop. The model is linear and time-invariant so that the headways are tractable.

However, this model does not produce bus bunching. Even when there is no control at all, namely $D^t = 0$ for all t, the headways keep permutating and do not diverge. This is because the bus headway evolution is only triggered by bus traveling, but it is passengers' boarding behavior that keeps augmenting the variation of headways and produces bus bunching. Passengers' boarding behavior can be captured in their dwell time.

2.2 Bus Dynamics with Deterministic Dwell Time Delay

In this section, we will incorporate dwell time into the previous model so that the new model has the potential to produce bunching. It shows how the level of ridership affects the bus dynamics.

Modeling

When dwell time is not ignored, headway consists of both traveling time and dwell time. Note that the lengths of an epoch for all buses are the same and dwell time is typically much smaller than traveling time in the US, so the traveling distances of all buses except the one that may be delayed at the control point are similar at a epoch. But the accumulative numbers of passengers at different places along the route are slightly different, which depend on the amount of time since the preceding bus leaves that place.

When buses are in equilibrium state, the headways of all the buses are the same. They spend the same amount of time on traveling, so the dwell times are the same. Denote the equilibrium traveling time by s^* and denote the equilibrium dwell time by Q^*, which are unknown in advance.

When buses are not in equilibrium state, a bus with larger distance between it and its preceding bus spends more time on boarding. We assume that its dwell time spent at Epoch t, which we denote Q_i^t, is approximately affine in s_i^t, i.e. that:

$$Q_i^t \approx Q^* + b(s_i^t - s^*) = bs_i^t + (Q^* - bs^*). \tag{8}$$

The dwell times of all buses have the constant term $Q^* - bs^*$. This constant amount of time includes the time spent on slowing down when a bus arrives at a bus stop, opening the doors, closing the doors, and speeding up when the bus leaves the bus stop. This constant part can be incorporated into the bus velocity, implying that the velocity becomes smaller due to the constant delay described above and the dwell time becomes

$$Q_i^t \approx bs_i^t. \tag{9}$$

b is the marginal increase in dwell time arising from a unit increase in traveling time — becasue the larger travelling time implies longer time the preceding bus has left which is also the amount of time passengers accumulates at the bus stops. The value of b reflects the level of ridership. It ranges from 0 to 0.2 for Georgia Tech campus bus system during a day: when there are no passengers early in the mornings, b is close to zero; when it comes to morning peaks and evening peaks, the value becomes much larger.

So the length of Epoch t, which is the headway of Bus n, equals the sum of traveling time and dwell time, i.e. that

$$\begin{aligned} h_n^t &= s_n^t + Q_n^t \\ &= s_n^t + bs_n^t. \\ &= (1+b)s_n^t \end{aligned}$$

Similarly, the headway of Bus i, $i = 1, 2, \ldots, n$, at Epoch t is approximately

$$h_i^t = (1+b)s_i^t$$

The time used for Bus i, $i = 2, 3, \ldots, n$ to travel at Epoch t is the length of Epoch t subtracted by the dwell time, i.e. that

$$h_n^t - Q_i^t = (1+b)s_n^t - bs_i^t$$

The traveling time of Bus 1 at Epoch t is similar except additionally subtracting delay time at the control point, i.e. that

$$h_n^t - Q_1^t - D^t = (1+b)s_n^t - bs_1^t - D^t$$

Thus, we have the following dynamics equations in terms of bus positions:

$$x_1^{t+1} = 0 \tag{10}$$

$$x_2^{t+1} = x_1^t + [(1+b)s_n^t - bs_1^t - D^t]v \tag{11}$$

$$x_i^{t+1} = x_{i-1}^t + [(1+b)s_n^t - bs_{i-1}^t]v \quad \forall i = 3, \ldots, n. \tag{12}$$

The mechanism that produces bus bunching is embedded in this model. Look at Equation (12). For Bus $i-1$ at Epoch t (or Bus i at Epoch $t+1$), when the distance between it and its preceding bus is larger, namely s_{i-1}^t is large, then bs_{i-1}^t is larger. That means it spends more time on boarding passengers. And then $(1+b)s_n^t - bs_{i-1}^t$ becomes smaller. That means the time spent on traveling becomes smaller. So the distance between it and its preceding bus tend to become even more larger later.

The dynamics above can be expressed only in terms of s_i^t. We can write for each $i = 3, 4, \ldots, n$,

$$\begin{aligned}
s_i^{t+1} &= \frac{x_{i+1}^{t+1} - x_i^{t+1}}{v} \\
&= \frac{[x_i^t + [(1+b)s_n^t - bs_i^t]v] - [x_{i-1}^t + [(1+b)s_n^t - bs_{i-1}^t]v]}{v} \quad \text{(by Equation (12))} \\
&= s_{i-1}^t + b(s_{i-1}^t - s_i^t)
\end{aligned}$$

and

$$
\begin{aligned}
s_1^{t+1} &= \frac{x_2^{t+1} - x_1^{t+1}}{v} \\
&= \frac{[(1+b)s_n^t - bs_1^t - D^t]v}{v} \quad \text{(by Equation (11))} \\
&= (1+b)s_n^t - bs_1^t - D^t
\end{aligned}
$$

and finally,

$$
\begin{aligned}
s_2^{t+1} &= \frac{x_3^{t+1} - x_2^{t+1}}{v} \\
&= \frac{[x_2^t + [(1+b)s_n^t - bs_2^t]v] - [x_1^t + [(1+b)s_n^t - bs_1^t - D^t]v]}{v} \\
&\quad \text{(by Equation (11) and (12))} \\
&= s_1^t + D^t - b(s_2^t - s_1^t)
\end{aligned}
$$

Thus, we have the following system

$$
\begin{aligned}
s_1^{t+1} &= s_n^t - b(s_1^t - s_n^t) - D^t \\
s_2^{t+1} &= s_1^t - b(s_2^t - s_1^t) + D^t \qquad (13) \\
s_i^{t+1} &= s_{i-1}^t - b(s_i^t - s_{i-1}^t) \quad \forall i = 3, \ldots, n. \\
h_i^t &= (1+b)s_i^t \quad \forall i = 1, 2, \ldots, n. \qquad (14)
\end{aligned}
$$

$h_i^t, \forall i \in I$ is the information we can use for control. We use h_1^t here, if $1 \in I$, instead of $\widehat{h}_1^t$ because $\widehat{h}_1^t$ relies on the actual delay time but we do not know the delay time before we determine it.

There is a parameter, v, that does not explicitly appear in the model. In fact, it appears in the underlying constraint that the sum of all the traveling times equals the loop time, i.e. that

$$\sum_{i=1}^{n} s_i^t = \frac{1}{v}.$$

The average velocity, v, captures the overall traffic conditions which are influenced by weather, time period, population density, traffic rules, etc.

This model is an extension of that in Section 2.1. In fact, when $b = 0$, they have the same form.

This model remains linear and time-invariant but captures basic bus dynamics with the tendency to bus bunching. It is a simple and useful platform to analyze how well a bus system under control resists bunching.

Generating Bus bunching

This model directly displays the essence of the cause of bus bunching. When no control is applied, namely $D^t = 0$ for all t, it holds that $s_i^{t+1} = (1+b)s_{i-1}^t - bs_i^t$ for all i. The hand right side of this equality consists of two parts, the original traveling time s_{i-1}^t and the dwell time difference of the bus and its preceding bus $b(s_{i-1}^t - s_i^t)$. If the bus has larger dwell time than its preceding bus, then $s_i^{t+1} > s_{i-1}^t$. It follows that $b(s_{i-1}^t - s_i^t) > 0$ which implies that the bus falls further behind. And the amount of time falling behind is larger as b becomes larger.

Here is an instance of a bus route with 4 buses. traveling times are generated using dynamics (13) without any delay, namely $D^t = 0, \forall t > 0$. The simulation starts with almost equal traveling times: $s_1^0 = 0.25 + 0.001, s_2^0 = 0.25 - 0.001, s_3^0 = s_4^0 = 0.25$. The boarding intensity parameter b in (13) is set 0.1. The result shows in Figure 1. The curves represent headways generated by Dynamics Equations (13).

This typical example demonstrates the intensive positive feedback mechanism in the bus system that leads to strong bunching. The rate of headway changes is super-linear. Only after 20 epoches which means all buses completed 5 loops of the route, the traveling time of Bus D hits 0. That means Bus D catches up with Bus A. The other two buses are also close to each other.

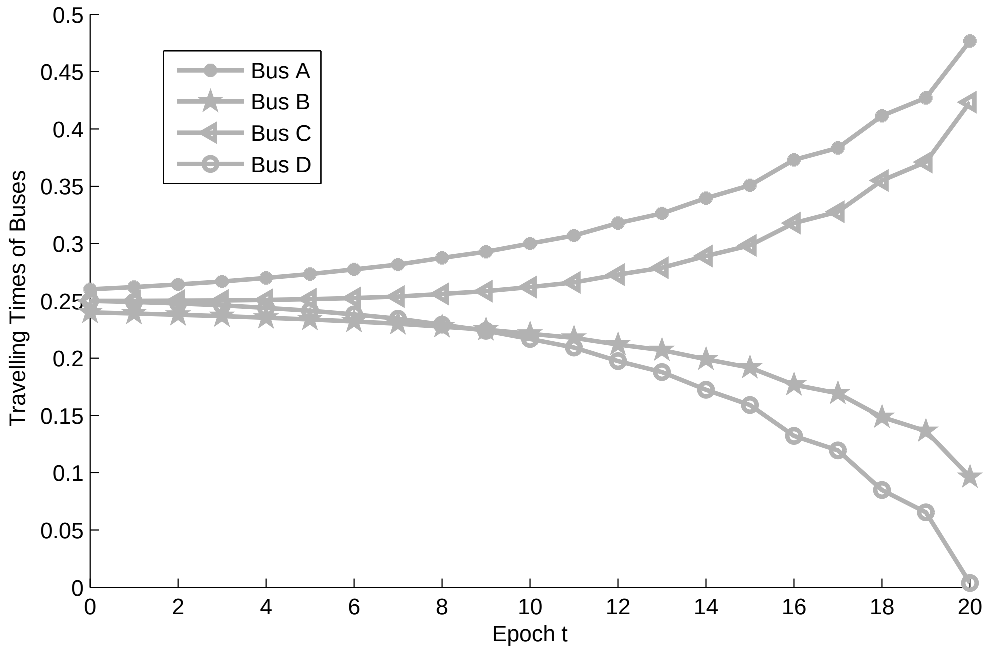

Figure 1: Evolution of traveling Times without Control

2.2.1 Bus bunching intensity and the level of ridership

When there is no control, the vector form of the bus dynamic equations is $\mathbf{s^{t+1}} = A\mathbf{s^t}$ where $\mathbf{s^t} = [s_1^t, s_2^t, \ldots, s_n^t]^T$ and

$$A = \begin{pmatrix} -b & & \cdots & & 1+b \\ 1+b & -b & \cdots & & \\ & 1+b & \cdots & & \\ & & \ddots & -b & \\ & & & 1+b & -b \end{pmatrix}_{n\times n}.$$

According to linear algebra, there exists an invertible matrix $V \in \mathbb{R}^{n\times n}$ such that $A = V^{-1}\Lambda V$ where Λ is a diagonal matrix with A's eigenvalues on the diagonal. So

$$\mathbf{s^t} = V^{-1}\Lambda V\mathbf{s^{t-1}} = (V^{-1}\Lambda V)^2\mathbf{s^{t-2}} = V^{-1}\Lambda^2 V\mathbf{s^{t-2}} = \cdots = V^{-1}\Lambda^t V\mathbf{s^0}.$$

If the modulus of the largest eigenvalue of A, denoted $|\lambda_1(A)|$, is greater than 1, then the divergence rate of traveling times is $|\lambda_1(A)|$. So $|\lambda_1(A)|$ measures the intensity of bus bunching. $|\lambda_1(A)|$ depends only on the boarding parameter b which reflects the level of ridership. We derives an exact formula for $|\lambda_1(A)|$:

Theorem 2.2.1 *When there is no control, it holds that*

$$|\lambda_1(A)| = 1 + 2b.$$

Proof. When there is no control, the characteristic polynomial of A is

$$p(\lambda) = (-b - \lambda)^n - (1 + b)^n.$$

Hence the roots of $p(\lambda) = 0$, which are also the eigenvalues of A, equal

$$-b - (1 + b)e^{\frac{2k\pi}{n}i}, \quad k = 0, 1, 2, \dots, n - 1.$$

So the eigenvalue of the largest modulus is

$$\lambda_1 = -b - (1 + b) = -1 - 2b$$

and finally

$$|\lambda_1| = 1 + 2b.$$

□

Since $b > 0$, it holds that $|\lambda_1(A)| > 1$. Therefore, the traveling times always diverge if there is no control. Bus bunching was born with the bus system. Since $|\lambda_1(A)|$ is an increasing function of b, the tendency to bus bunching becomes stronger as the number of passengers increase. It is ironical that the bus system is created to serve the passengers and passengers are the culprit of bus bunching.

Figure 2 shows the evolution of the traveling times of the same bus in a 4-bus route when $b_1 = 0.1$ and $b_2 = 0.05$. Since b_1 is twice as large as b_2, the divergence of the headways in the former case is much faster.

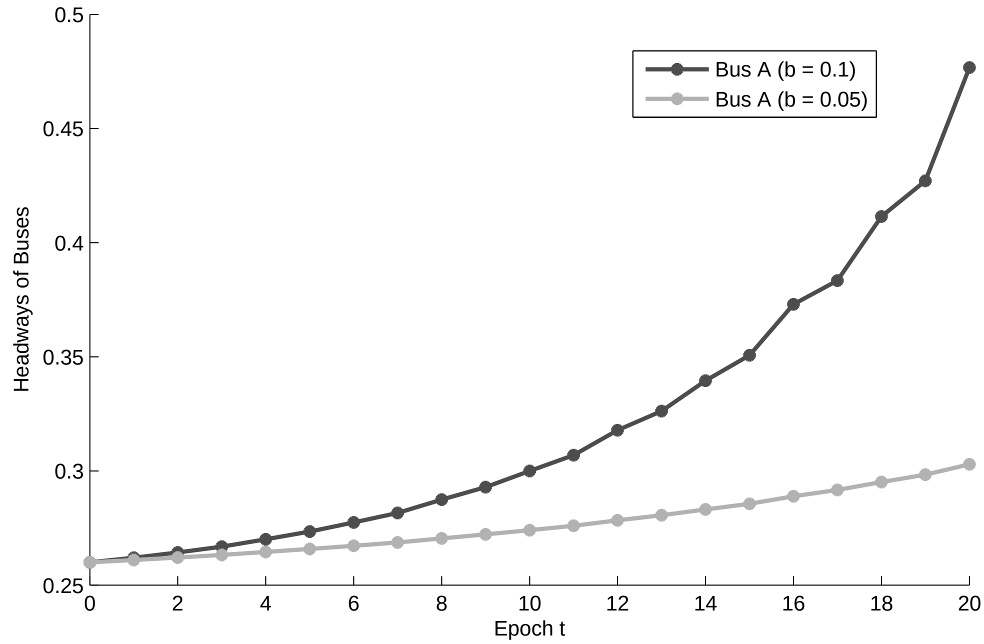

Figure 2: Comparison of Headways Generated by Dynamics Equations with Different b without Control

Model Extensions

The model developed in this section is based on several assumptions such as constant traveling velocity, deterministic dwell time and single control point. Some of these assumptions can be relaxed to be more general.

The model can be extended, by scaling, to consider any common bus traveling velocity function $v(x)$ that gives the instantaneous velocity of a bus at each point x along the route, as long as $v(x)$ is bounded above and below at every point (see [12], for example). This weakens the assumption that all buses have a constant common average traveling velocity.

The model also can be extended to account for sufficiently smooth and small noise as in [13] or [21], but the subsequent stability analysis, which is presented in Chapter 3, is similar. We will incorporate headway forecasting errors into the model in Chapter 3 to study the influence of different amounts of headway information on the stability of the bus system.

The model cannot be extended to accommodate more than one control points in general. But when we study control schemes that depend only on local information, the stability analysis can be extended to multi-control-point case, which will be discussed in Chapter 3.

CHAPTER III

A GENERAL CLASS OF HEADWAY CONTROL SCHEMES

In this chapter, we study a general class of headway control schemes in the model developed in Section 2.2. We focus on the stability and robustness of the bus system under control.

In Section 3.2, we derive a fundamental measure of system stability that characterizes how fast headways converge or diverge. We present a sufficient and necessary condition under which headways self-equalize to a unique common headway.

In Section 3.3, we demonstrate that the stability of the bus system is robust against the level of ridership. In Section 3.4, we study what is the sufficient amount of slack.

In Section 3.5, we show bounded deviations from ideal headway under the appearance of headway forecasting errors. We also quantify how the forecasting errors affects the bound.

3.1 General Form of Delay Time

Thanks to the automatic vehicle location (AVL) system such as global positioning system (GPS) and signpost based system, real-time headways can be estimated. We can control the bus system based on real-time headway information. We restrict the expression of D^t in the following form affinely depending on real-time headways:

$$D^t = \gamma_0 + \sum_{i=1}^{n} \gamma_i h_i^t \tag{15}$$

where all the coefficients $\gamma_i,\ \forall i = 0, 1, \ldots, n$, are constants independent of time. γ_0 is a large enough constant so that D^t is positive almost all the time.

In practice, this form is easy for managers, supervisors, and drivers to understand. The corresponding bus system is easy to implement and monitored.

In theory, it maintains the linear structure of the bus dynamic equations. Besides, the dynamic equations are time-invariant and thus it is easy to track the bus dynamics.

This form is flexible to represent a broad class of control schemes. If only partial headway information is known, we let the coefficients corresponding to the unknown headways be 0. It subsumes target headway control and current self-organizing headway control (See literature review in Section 1.1). Target headway control has delay time

$$D^t = g + \gamma_1(h_1^t - h) = (g - \gamma_1 h) + \gamma_1 h_1^t$$

where g is the constant slack and h is the target headway.

Backward headway control [11] has delay time

$$D^t = \gamma_n h_n^t.$$

"Prefol" has delay time

$$D^t = \gamma_0 + \gamma_1 h_1^t + \gamma_n h_n^t$$

and generally $\gamma_0 = 0$, $\gamma_1 = -0.5$, and $\gamma_n = 0.5$.

Instead of choosing such a simple form, one may directly solve some sort of optimal control problems to obtain the corresponding optimal delay time. However, it is optimal only given that the parameters b and v are known precisely. But in fact, they fluctuates during a day. So the optimal solution to the model generally is not an optimal solution in practice. Also, the solution may appear to be more complicated (time-variant or non-linear). In contrast, we assume that the delay time is of an affine form with time-invariant coefficients to keep simplicity and then try to guarantee that the bus system under control is robust against the variation of the parameters.

3.2 Self-Equalization of Headways

The most important property of a bus system is that it has unique equal equilibrium headways. Equilibrium headways refer to the headways that can be stabilized at after disturbance. There are two conditions for headways to be equilibrium headways:

1. Headways are stationary. When they start from themselves, they remain unchanged in the future.

2. Headways are stable. When they are disturbed, they are able to resile.

We study these conditions and propose a measure of stability. This measure characterizes how fast headways converge or diverge.

3.2.1 Stationarity of Equilibrium Headways

Headways are stationary if they remain the same after one epoch in the snapshot model. Hence equilibrium headways are always stationary.

Headways are stationary if and only if traveling times are stationary. Indeed, when traveling times are equal, then so are the amounts of time spent on boarding passengers, and thus so are headways. When traveling times are not equal, the amounts of time spent on boarding passengers are different, and thus headways will change. Now we show the sufficient and necessary condition for the existence of stationary traveling times.

Denote by $\mathbf{s}^*$ the stationary traveling times.

Proposition 3.2.1 *The bus system with delay time D^t at Epoch t, $t = 1, 2, \ldots,$ has unique stationary traveling times if and only if $n(1+\gamma_1) - \sum_{i=1}^n \gamma_i \neq 0$. If it has, the stationary traveling times are*

$$\begin{aligned} s_1^* &= \frac{(1-\sum_{i=2}^n \gamma_i)\frac{1}{v} - (n-1)\frac{\gamma_0}{1+b}}{n(1+\gamma_1) - \sum_{i=1}^n \gamma_i} \\ s_i^* &= \frac{(1+\gamma_1)\frac{1}{v} + \frac{\gamma_0}{1+b}}{n(1+\gamma_1) - \sum_{i=1}^n \gamma_i}, \quad \forall i = 2, \ldots, n. \end{aligned} \tag{16}$$

Proof. Since

$$h_i^t = (1+b)s_i^t,$$

it holds that

$$D^t = \gamma_0 + (1+b)\sum_{i=1}^{n} \gamma_i s_i^t. \tag{17}$$

Plugging in (17) into the bus dynamics (13), we have that

$$\begin{aligned} s_1^{t+1} &= s_n^t + b(s_n^t - s_1^t) - (1+b)\sum_{i=1}^{n} \gamma_i s_i^t - \gamma_0 \\ s_2^{t+1} &= s_1^t + b(s_1^t - s_2^t) + (1+b)\sum_{i=1}^{n} \gamma_i s_i^t + \gamma_0 \\ s_i^{t+1} &= s_{i-1}^t + b(s_{i-1}^t - s_i^t) \quad \forall i = 3, \dots, n. \end{aligned}$$

We can represent the dynamics in vector form: $\mathbf{s^{t+1}} = A\mathbf{s^t} + \mathbf{r}$ where $A =$

$$\begin{pmatrix} -b-(1+b)\gamma_1 & -(1+b)\gamma_2 & -(1+b)\gamma_3 & \cdots & -(1+b)\gamma_{n-1} & (1+b)(1-\gamma_n) \\ (1+b)(1+\gamma_1) & -b+(1+b)\gamma_2 & (1+b)\gamma_3 & \cdots & (1+b)\gamma_{n-1} & (1+b)\gamma_n \\ & 1+b & -b & & & \\ & & 1+b & \cdots & & \\ & & & \ddots & -b & \\ & & & & 1+b & -b \end{pmatrix}_{n\times n},$$

$\mathbf{s^t} = [s_1^t, s_2^t, \dots, s_n^t]^T$ and $\mathbf{r} = [-\gamma_0, \gamma_0, 0, 0, \dots, 0]^T$.

By the definition of stationarity, $\mathbf{s}^*$ satisfies that

$$\mathbf{s}^* = A\mathbf{s}^* + \mathbf{r}$$

So we have that

$$s_i^* = s_{i-1}^* + b(s_{i-1}^* - s_i^*), \quad \forall i = 3, \dots, n,$$

and thus

$$s_i^* = s_{i-1}^*, \quad \forall i = 3, \dots, n.$$

Then there exists a constant s^* such that

$$s_i^* = s^*, \quad \forall i = 2, 3, \dots, n. \tag{18}$$

The first equation of the algebraic equations is

$$s_1^* = s_n^* + b(s_n^* - s_1^*) - (1+b)\sum_{i=1}^{n} \gamma_i s_i^* - \gamma_0$$

Plugging in Equation (18), we have

$$(1+\gamma_1)s_1^* = (1 - \sum_{i=2}^{n} \gamma_i)s^* - \frac{\gamma_0}{1+b}.$$

Combined with the route length constraint

$$\sum_{i=1}^{n} s_i^* = \frac{1}{v}$$

if $n(1+\gamma_1) - \sum_{i=1}^{n} \gamma_i \neq 0$, we have

$$s_1^* = \frac{(1 - \sum_{i=2}^{n} \gamma_i)\frac{1}{v} - (n-1)\frac{\gamma_0}{1+b}}{n(1+\gamma_1) - \sum_{i=1}^{n} \gamma_i}$$
$$s_i^* = s^* = \frac{(1+\gamma_1)\frac{1}{v} + \frac{\gamma_0}{1+b}}{n(1+\gamma_1) - \sum_{i=1}^{n} \gamma_i}, \quad \forall i = 2, \ldots, n.$$

If $n(1+\gamma_1) - \sum_{i=1}^{n} \gamma_i = 0$, the algebraic equations have infinite solutions when $-\frac{\gamma_0}{1+b} = (1+\gamma_1)\frac{1}{v}$ and have no solution when $-\frac{\gamma_0}{1+b} \neq (1+\gamma_1)\frac{1}{v}$. □

The stationary traveling time of the bus at the control point (s_1^*) may be different from others because it may be delayed for a constant amount of time at the stationary situation.

By the relation between traveling time and headway, we have the sufficient and necessary condition for the existence of stationary headways.

Corollary 3.2.1 *The bus system with delay time D^t has unique stationary headways if and only if $n(1+\gamma_1) - \sum_{i=1}^{n} \gamma_i \neq 0$. If it has, the stationary headways are*

$$\begin{aligned} h_1^* &= \frac{(1+b)(1 - \sum_{i=2}^{n} \gamma_i)\frac{1}{v} - (n-1)\gamma_0}{n(1+\gamma_1) - \sum_{i=1}^{n} \gamma_i} \\ h_i^* &= \frac{(1+b)(1+\gamma_1)\frac{1}{v} + \gamma_0}{n(1+\gamma_1) - \sum_{i=1}^{n} \gamma_i}, \quad \forall i = 2, \ldots, n. \end{aligned} \tag{19}$$

This condition relates to a spectral property of the transition matrix–when there is only one eigenvalue that equals 1, this condition is satisfied. We will identify and use this property in the stability analysis.

3.2.2 Stability of Equilibrium Headways

Equilibrium headways are stationary, but stationary headways are not necessarily equilibrium headways–They may be unstable. In this part, we study the stability property of the headways. We find that the systemic stability is completely reflected by the modulus of the second largest eigenvalue of the transition matrix. It can be used to determine whether bus bunching will occur. If bus bunching occurs, it measures how fast the headways diverge. If bus bunching does not occur, it reflects how fast headways self-equalize from any starting situation. It characterizes the resistance of the bus system to bunching and the recoverability from disruptions.

Note that the column sums of A all equal 1. So 1 is an eigenvalue of A. Denote the eigenvalues of A by $\lambda_1(A), \lambda_2(A), \ldots, \lambda_n(A)$ such that $|\lambda_1(A)| \geq |\lambda_2(A)| \geq \cdots \geq |\lambda_n(A)|$ and denote their corresponding (generalized) eigenvectors by $v_1, v_2, \ldots, v_n$.

Theorem 3.2.2 [Self-Equalizing Theorem for traveling Time]

1. *If $|\lambda_2(A)| < 1$, then $\lambda_1(A) = 1$ and*

   (a) *there exists unique stationary traveling times.*

   (b) *Given any initial traveling times, the traveling times converge to the stationary traveling times $\mathbf{s}^*$ in Proposition 3.2.1.*

   (c) *There exists a constant C such that $\|A^t\mathbf{s^0} - \mathbf{s}^*\|_2 \leq C|\lambda_2(A)|^t, \forall t > 0$ when A is diagonalizable.*

2. *If $|\lambda_1(A)| \geq 1$ and $|\lambda_2(A)| \geq 1$, then there are no unique stationary traveling times.*

Remark 3.2.1 *When A is not diagonalizable, the inequality in* 1.(*c*) *holds for all $t > n$. The idea in the proof is similar, but the argument is slightly different (See Appendix A).*

Proof.

1. The proof consists of three steps:

   (a) Show that there exists unique stationary traveling times when $|\lambda_2(A)| < 1$.

   (b) Show that the inhomogeneous linear system $\mathbf{s^{t+1}} = A\mathbf{s^t} + \mathbf{r}$ can be transformed to a homogeneous linear system $\mathbf{y^{t+1}} = A\mathbf{y^t}$ by an affine transform of $\mathbf{s^t}$.

   (c) Prove the convergence of the linear system $\mathbf{y^{t+1}} = A\mathbf{y^t}$.

   (**Step 1a**)

   The characteristic polynomial of A is

$$\begin{aligned} p(\lambda) &= (-b-1)^n(z-1)\sum_{i=1}^{n-1}(1+\gamma_1-\gamma_i)z^{n-i} \\ &= (z-1)q(\lambda) \end{aligned}$$

   where $z = \frac{-b-\lambda}{-b-1}$. All the eigenvalues are roots of $p(\lambda)$. When $\lambda = 1$, we have $z = 1$ and thus $p(1) = 0$. So all the eigenvalues except 1 are roots of $q(\lambda) = 0$. Since $|\lambda_2(A)| < 1$, 1 is not a root of $q(\lambda)$. That is

$$n(1+\gamma_1) - \sum_{i=1}^{n}\gamma_i = q(1) \neq 0$$

   According to Proposition 3.2.1, there exists unique stationary traveling times. This finishes the proof of 1.(a) in the theorem.

   (**Step 1b**)

   Let $\mathbf{y^t} = \mathbf{s^t} + \mathbf{q}$, where $\mathbf{q}$ is a constant vector. Forcing $\mathbf{y^{t+1}} = A\mathbf{y^t}$ to hold, we have that $\mathbf{r} = (A-I)q$, or

$$\begin{aligned} & q_2 = q_3 = \cdots = q_n \\ & (-1-\gamma_1)q_1 - (1-\sum_{i=2}^{n}\gamma_i)q_2 = -\gamma_0 \end{aligned}$$

When $\gamma_1 \neq -1$, we can set $q_2 = q_3 = \cdots = q_n = 0$ and $q_1 = \frac{\gamma_0}{1+\gamma_1}$.

When $\gamma_1 = -1$, we claim that $1 - \sum_{i=1}^{n} \gamma_i \neq 0$ and we can set $q_1 = 0$ and $q_i = -\frac{\gamma_0}{1-\sum_{i=1}^{n} \gamma_i}$, $\forall i = 2, 3, \ldots, n$.

The claim comes from the fact that unique stationary headways exist. According to Corollary 3.2.1, it holds that $n(1 + \gamma_1) - \sum_{i=1}^{n} \gamma_i \neq 0$. Since $\gamma_1 = -1$, it follows that $1 - \sum_{i=1}^{n} \gamma_i \neq 0$.

Therefore, the inhomogeneous linear system $\mathbf{s^{t+1}} = A\mathbf{s^t} + \mathbf{r}$ can be transformed to a homogeneous linear system $\mathbf{y^{t+1}} = A\mathbf{y^t}$.

(**Step 1c**)

Let $\mathbf{y}^* = \mathbf{s}^* + \mathbf{q}$. Then $\mathbf{y}^*$ satisfy the stationary condition:

$$\mathbf{y^{t+1}} = A\mathbf{y^t}.$$

So $\mathbf{y}^*$ is a multiple of v_1 which is an eigenvector of A corresponding to the eigenvalue 1.

Also, since the column sums of A all equal 1, it holds that $\overrightarrow{1}A = \overrightarrow{1}$. So

$$\overrightarrow{1}v_i = \overrightarrow{1}Av_i = \overrightarrow{1}\lambda_i(A)v_i = \lambda_i(A)\overrightarrow{1}v_i, \quad \forall i = 2, 3, \ldots, n.$$

Since $\lambda_i(A) \neq 1$, it holds that $\overrightarrow{1}v_i = 0$, $\forall i = 2, 3, \ldots, n$.

Since A is diagonalizable, we can write $\mathbf{y^0}$ in terms of $\mathbf{y}^*, v_2, v_3, \ldots, v_n$ as

$$\mathbf{y^0} = a_1\mathbf{y}^* + a_2v_2 + a_3v_3 + \cdots + a_nv_n.$$

Left-multiplying by $\overrightarrow{1}$ on both sides, we have

$$\overrightarrow{1}\mathbf{y^0} = \overrightarrow{1}a_1\mathbf{y}^* + \overrightarrow{1}(a_2v_2 + a_3v_3 + \cdots + a_nv_n) = a_1\overrightarrow{1}\mathbf{y}^*$$

Since

$$\overrightarrow{1}\mathbf{y^0} = \overrightarrow{1}(\mathbf{s^0} + \mathbf{q}) = \overrightarrow{1}(\mathbf{s}^* + \mathbf{q}) = \overrightarrow{1}\mathbf{y}^*$$

it holds that $a_1 = 1$. Therefore, $A\mathbf{y^0} = \mathbf{y}^* + a_2\lambda_2(A)v_2 + a_3\lambda_3(A)v_3 + \cdots + a_n\lambda_n(A)v_n$, and hence

$$\begin{aligned}
\|A^t\mathbf{y^0} - \mathbf{y}^*\|_2 &= \|\sum_{i=2}^{n} a_i\lambda_i^t(A)v_i\|_2 \\
&\leq \sum_{i=2}^{n} |\lambda_i(A)|^t \|a_iv_i\|_2 \\
&\leq |\lambda_2(A)|^t \sum_{i=2}^{n} \|a_iv_i\|_2 \\
&\triangleq C|\lambda_2(A)|^t
\end{aligned}$$

Since $A^t\mathbf{y^0} - \mathbf{y}^* = A^t(\mathbf{y^0} - \mathbf{y}^*) = A^t(\mathbf{s^0} - \mathbf{s}^*) = A^t\mathbf{s^0} - \mathbf{s}^*$, it holds that

$$\|A^t\mathbf{s^0} - \mathbf{s}^*\| \leq C|\lambda_2(A)|^t.$$

So given any initial traveling times, the traveling times converge to the equilibrium traveling times which are $\mathbf{s}^*$.

2. If $|\lambda_1(A)| \geq 1$ and $|\lambda_2(A)| \geq 1$, when there are two eigenvalues equal to 1, then there exist an eigenvalue $\lambda_j(A)$ such that $n(1+\gamma_1) - \sum_{i=1}^{n}\gamma_i = q(\lambda_j(A)) = 0$ according to the argument in Step 1a. By Proposition 3.2.1, there are no unique stationary traveling times.

   Suppose only one of the eigenvalues equals 1. Without loss of generality, let $\lambda_k(A) = 1$. Since there are at least two eigenvalues of modulus not less than 1, there exists some $j \geq 2$ such that $\lambda_i(A) \geq 1$ for $i = 1, 2, \ldots, j$. Let U be the set consisting of the indices of eigenvalues whose value does not equal 1 and modulus not less than 1.

   Suppose that traveling times converge for all initial traveling times. We can pick initial traveling times such that $a_i \neq 0$, for $i = 1, 2, \ldots, n$. Arguments in Step 1a and 1b still hold here. We can again decompose $\mathbf{y^0}$:

$$\mathbf{y^0} = a_1v_1 + a_2v_2 + a_3v_3 + \cdots + a_nv_n.$$

So

$$\begin{aligned}
\mathbf{y^t} &= A^t \mathbf{y^0} \\
&= A^t(a_1 v_1 + a_2 v_2 + a_3 v_3 + \cdots + a_n v_n) \\
&= \sum_{i=1}^{n} a_i A^t v_i \\
&= \sum_{i=1}^{n} a_i \lambda_i(A)^t v_i
\end{aligned}$$

Note that $a_k \lambda_k(A)^t v_k = a_k v_k$. For $i > j$, $a_i \lambda_i(A)^t v_i$ goes to 0 when t goes to infinity. But for $i \in U$, the limit of $\lambda_i(A)^t$ does not exist. So $\sum_{i \in U_2} a_i \lambda_i(A)^t v_i = 0$ for t greater than some constant, which implies $a_i = 0$ and thus contradicts our assumption that $a_i \neq 0$. So the traveling times do not converge for any initial traveling times.

$\square$

Because of the fixed proportional relation between headways and traveling times, we have the following corollary

Corollary 3.2.2 [Self-Equalizing Theorem for Headways]

1. *If $|\lambda_2(A)| < 1$, then*

   (a) *there exists unique stationary headways.*

   (b) *Given any initial headways, the headways converge.*

   (c) *The equilibrium headways are the stationary headways $\mathbf{h}^*$ in Corollary 3.2.1, and*

   (d) *There exists a constant C such that $\|A^t \mathbf{h^0} - \mathbf{h}^*\|_2 \leq C|\lambda_2(A)|^t, \forall t > 0$ when A is diagonalizable.*

2. *If $|\lambda_1(A)| \geq 1$ and $|\lambda_2(A)| \geq 1$, then headways do not converge for all initial headways.*

$|\lambda_2(A)|$ not only reflects the stability of equilibrium headways, it is also a fundamental measure of systemic stability because convergence occurs for any initial headways. The smaller $|\lambda_2(A)|$ is, the more stable the bus system is. When $\lambda_1(A) \neq 1$, bus bunching will occur. When $|\lambda_2(A)| < 1$, headways self-equalize and $|\lambda_2(A)|$ is the convergence rate of headways in the sense that $\|A^t\mathbf{h^0} - \mathbf{h}^*\|_2 \leq C|\lambda_2(A)|^t$.

If there are m control points and we use the same form of delay time at all of them, then the convergence rate becomes $|\lambda_2(A)|^m$ which is smaller than $|\lambda_2(A)|$. So more control points lead to a more stable bus system.

With this theorem, we can easily check the stability of a bus system. For example, $|\lambda_1(A)| = 1.02$ for a bus system with four buses under no control and the boarding parameter b equals 0.01. So the headways diverge as expected. If we apply backward headway control by setting $\gamma_0 = \gamma_1 = \gamma_2 = \cdots = \gamma_{n-1} = 0$ and $\gamma_n = 0.5$, $|\lambda_2(A)| = 0.8894$ and thus headways self-equalize as shown in [11].

With this theorem, we can easily compare different control schemes. For example, if we change γ_n to be 0.1 in the backward headway control scheme mentioned above, $|\lambda_2(A)|$ becomes 0.9848. So this new scheme should lead to a slower headway self-equalization, which can be checked by plotting the evolution of the headway standard deviations in the model (See Figure 3).

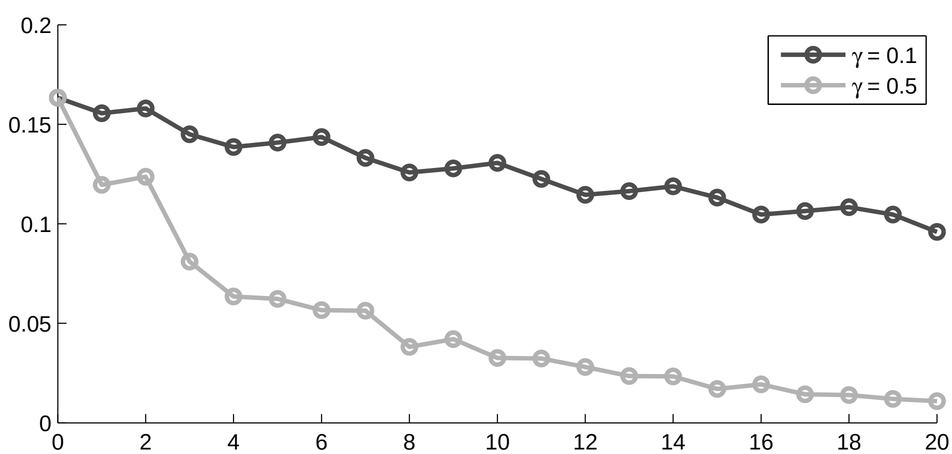

Figure 3: Evolution of headway standard deviation when $b = 0.01$

Are there any control schemes with $|\lambda_2(A)| < 1$? The backward headway control scheme [11] which sets $\gamma_0 = \gamma_1 = \gamma_2 = \cdots = \gamma_{n-1} = 0$ and $\gamma_n = 0.5$ when $b < 0.05$ is an example.

Now that there exists such control schemes where $|\lambda_2(A)| < 1$, there should be an optimum among them in theory in the sense of having the smallest $|\lambda_2(A)|$. Mathematically, we can identify it by solving the following optimization problem:

$$\min_{\gamma_i, i \in I} |\lambda_2(A)| \tag{20}$$

This problem is not convex. But according to the author's numerical experiences, the objective function is trumpet-shaped—conical but flaring broader and broader. We conjecture that there is only one local optimal which is thus global optimal. We also solve this problem when $1 \in I$ theoretically in Chapter 4.

Generally, it is needless to solve this problem to optimal. The optimal solution in this problem is probably not optimal in practice because the model is idealized. For practice, we suggest identifying parameters corresponding to a small enough $|\lambda_2(A)|$, e.g. less than 0.75, and then adjusting them in field tests.

3.3 Robustness of Systemic Stability

Since $|\lambda_2(A)|$ is a function of the transition matrix A and A has the boarding parameter b, the systemic stability depends on b which reflects the level of ridership. Because we do not know b in advance, it is necessary to study the robustness of systemic stability against the level of ridership.

In this section, we will show that $|\lambda_2(A(b))|$ is insensitive to b. So generally the systemic stability varies little when the level of ridership changes. However, bus system with $|\lambda_2(A)|$ close to 1 may turn unstable when the level of ridership changes.

Let $\mu_i(b)$'s, $i = 1, 2, \ldots, n$, be eigenvalues of $A(b)$, not necessary in the same order of λ_i. In other words, $\mu_i(b)$'s is a permutation of λ_i's.

Proposition 3.3.1 [Sensitivity of systemic stability]

1. For $i = 1, 2, \ldots, n$, it holds that

$$\frac{d|\mu_i(A(b))|}{db} = \frac{(1+b)(z_{ix}^2 + z_{iy}^2) - (1+b)z_{ix} + b(1 - z_{ix})}{|\mu_i(A(b))|}.$$

where $z_{ix}, z_{iy} \in \mathbb{R}$ and $z_{ix} + iz_{iy}$ is the $i-th$ eigenvalue of $A(b))$ when $b = 0$.

2. For any positive number Δb, it holds that

$$|\ |\mu_i(A(b))| - |\mu_i(A(b + \Delta b))|\ | \leq 2\Delta b.$$

Proof. 1. Given γ_i, $\forall i = 1, 2, \ldots, n$, the characteristic polynomial of A is

$$\begin{aligned} p(\mu) &= p(z(\mu)) \\ &= (-1-b)^n (z(\mu) - 1) \sum_{i=2}^{n} (1 + \gamma_1 - \gamma_i) z(\mu)^{n-i} \\ &\triangleq (-1-b)^n (z(\mu) - 1) q(z(\mu)) \end{aligned} \tag{21}$$

where $z(\mu) = \frac{\mu+b}{1+b}$. When $\mu = 1$, then $z = 1$ and hence $p(1) = 0$. So 1 is an eigenvalue of $A(b)$ no matter what b is.

Note that the roots of $p(z) = 0$ are independent of b. Suppose that they are z_i, $i = 1, 2, \ldots, n$ and $z_i = \frac{\mu_i(A(b))+b}{1+b}$ where $\mu_i(A(b))$ is the eigenvalue corresponding to z_i. Then

$$\mu_i = (z_i - 1)b + z_i$$

Let the real part and imaginary part of z_i be z_{ix} and z_{iy}, respectively. Then

$$\mu_i = (z_{ix} - 1)b + z_{ix} + i(z_{iy}b + z_{iy}).$$

So

$$|\mu_i| = \sqrt{[(z_{ix} - 1)b + z_{ix}]^2 + (z_{iy}b + z_{iy})^2}.$$

and

$$\frac{d|\mu_i|}{db} = \frac{(1+b)(z_{ix}^2 + z_{iy}^2) - (1+b)z_{ix} + b(1 - z_{ix})}{|\mu_i(A(b))|}.$$

Note that when $b = 0$, $z_i = \mu_i$. So z_i can be calculated as an eigenvalue of A when $b = 0$.

2. Let $b' = b + \Delta b$. It holds that

$$z_i = \frac{\mu_i(A(b)) + b}{1+b} = \frac{\mu_i(A(b')) + b'}{1+b'}. \tag{22}$$

This is equivalent to

$$\begin{aligned}
\mu_i(A(b')) &= \frac{1+b'}{1+b}\mu_i(A(b)) + \frac{\Delta b}{1+b'} \\
&= \mu_i(A(b)) + \Delta b \left(\frac{\mu_i(A(b))}{1+b} + \frac{1}{1+b'} \right)
\end{aligned}$$

So

$$|\mu_i(A(b'))| = \left| \mu_i(A(b)) + \Delta b \left(\frac{\mu_i(A(b))}{1+b} + \frac{1}{1+b'} \right) \right|$$

Then

$$\begin{aligned}
|\mu_i(A(b'))| &= \left| \mu_i(A(b)) + \Delta b \left(\frac{\mu_i(A(b))}{1+b} + \frac{1}{1+b'} \right) \right| \\
&\leq |\mu_i(A(b))| + \Delta b \left(\frac{|\mu_i(A(b))|}{1+b} + \frac{1}{1+b'} \right) \\
&\leq |\mu_i(A(b))| + \Delta b \left(\frac{1}{1+b} + \frac{1}{1+b'} \right) \\
&\leq |\mu_i(A(b))| + \frac{2\Delta b}{1+b} \\
&\leq |\mu_i(A(b))| + 2\Delta b
\end{aligned}$$

Similarly,

$$|\mu_i(A(b))| \leq |\mu_i(A(b'))| + 2\Delta b$$

So

$$|\ |\mu_i(A(b))| - |\mu_i(A(b'))|\ | \leq 2\Delta b$$

□

Corollary 3.3.1 *For any positive number Δb, it holds that*

$$|\ |\lambda_i(A(b))| - |\lambda_i(A(b + \Delta b))|\ | \leq 2\Delta b.$$

Proof. During the boarding parameter changes from b to $b + \Delta b$, the i^{th} largest modulus eigenvalue $\lambda_i(A)$ may switch to different μ_j's for finite times. Suppose that $\lambda_i(A)$ switches when the boarding parameter is b_k, $k = 1, 2, \ldots, M$, with $b = b_1 < b_2 < \cdots < b_M = b + \Delta b$. Then

$$\begin{aligned}
&| \, |\lambda_i(A(b))| - |\lambda_i(A(b+\Delta b))| \, | \\
&\qquad \leq \sum_{k=1}^{M-1} | \, |\lambda_i(A(b_k))| - |\lambda_i(A(b_{k+1}))| \, | \\
&\qquad \leq \sum_{k=1}^{M-1} 2(b_{k+1} - b_k). \quad \text{by 2 in Proposition 3.3.1} \\
&\qquad \leq 2\Delta b.
\end{aligned}$$

□

Corollary 3.3.1 shows that the change of $\lambda_2(A(b))$ with b is small. And since b has a small range, the variation of b has little influence on the value of $\lambda_2(A(b))$.

According to the first statement in Proposition 3.3.1, when $|\lambda_2(A(0))| < 1$, it holds that $|z_2| < 1$ and thus $1 - z_{2x} > 0$. If $z_{2x}^2 + z_{2y}^2 - z_{2x} > 0$, then $|\lambda_2(A(b))|$ is an increasing function of b. This is generally true. This is consistant with the intuition that as passengers arrival rate is larger, the tendency to bunching is stronger.

Since the parameter b ranges from 0 to 0.2 for the campus bus system at Georgia Tech, the magnitude of variation of the convergence rate could only be as large as 0.2. If $|\lambda_2(A(0))|$ is less than 0.7, then there is no risk that it exceeds 1 even if b climbs up to 0.2. But when $|\lambda_2(A(b))|$ is close to 1 when b is small, the headways may diverge when there are more passengers. For example, the backward headway control with delay time equal to $D^t = 0.5 s_n^t$ yields that $|\lambda_2(A(0))| = 0.9406$ and $|\lambda_2(A(0.1))| = 1.0239$ with 5 buses in the bus route. So when $b = 0.1$, the headways diverge.

3.4 Slack and Existence of Equilibrium Headways

The essence of slack is a mean to speed up a late bus by delaying all the other buses longer. If there is no slack, late buses have to speed up to avoid bunching, which is not allowed. So slack serves as a base of the delay time—If a bus should be sped up, it is delayed for a shorter time as long as the delay time is nonnegative; If a bus should be slowed down, it is delayed for longer than that when slack is not added.

Since a larger slack implies longer loop time, slack should be as small as possible as long as its function of speeding up remains effective, or in another word, the desired delay time is almost always non-negative. In this section, we study how large a slack should be even if traffic conditions and the level of ridership change.

In target schedule control, slack is the difference between scheduled departure time and expected arrival time at a control point. In target headway control, slack is the delay time of the bus at the control point when its current headway equals the target headway. To be consistent, the definition of slack in our general form is the bus's delay time when the headways are in the equilibrium state. Mathematically, the slack equals

$$D^* = \gamma_0 + \sum_{i=1^n} \gamma_i s_i^* = \frac{n(\gamma_0 + \frac{1+b}{nv}\sum_{i=1}^n \gamma_i)}{n(1+\gamma_1) - \sum_{i=1}^n \gamma_i}$$

Only when this delay time is non-negative, the equilibrium headways exist. It leads to the following conditions.

Lemma 3.4.1 *Equilibrium headways existence constraints:*

$$\gamma_0 + \frac{1+b}{nv}\sum_{i=1}^n \gamma_i \geq 0 \tag{23}$$

$$n(1+\gamma_1) - \sum_{i=1}^n \gamma_i > 0$$

Proof. Suppose that the numerator is non-positive and the denominator is negative. Since the stability of the system is invariant of γ_0, we can keep increasing γ_0 and fix other γ_i's so that D^* is still non-negative and the stability of the system does not change. Since the numerator is non-positive and is an increasing function of γ_0, it will become positive for some γ_0. But since the denominator is still negative, D^* becomes negative, contradicting the fact that D^* is non-negative. So the numerator is non-negative and the denominator is positive. Inequality (23) follows. $\square$

The second constraint is a stronger condition than that in Proposition 3.2.1. If this constraint is not satisfied, the equilibrium headways do not exist in the real world.

The first constraint gives a lower bound of γ_0. The lower bound is a function of b and v. In practice, we do not know the exact values of b and v. We can only obtain some confidence intervals of them using daily data. Assume the intervals are $[v_l, v_u]$ and $[b_l, b_u]$. To guarantee that the equilibrium headways almost always exist, Inequality (23) becomes

$$\gamma_0 \geq -\frac{1+b}{nv}\sum_{i=1}^{n}\gamma_i, \quad \forall v \in [v_l, v_u], b \in [b_l, b_u]. \tag{24}$$

This implies that

$$\gamma_0 \geq \max_{v\in[v_l,v_u],b\in[b_l,b_u]}\left\{-\frac{1+b}{nv}\sum_{i=1}^{n}\gamma_i\right\} \tag{25}$$

In implementation, we suggest setting γ_0 to be

$$m_u + m_c \tag{26}$$

where m_u is an upper bound for

$$\max_{v\in[v_l,v_u],b\in[b_l,b_u]}\left\{-\frac{1+b}{nv}\sum_{i=1}^{n}\gamma_i\right\}$$

and m_c is a constant. The m_c amount of time is not only used for compensating the time for management issues such as break of drivers and communication between supervisors and drivers, but enables extra delay time incurred by irregular headways as well.

Since we determine the maximum term according to the worst case, we are adding more delay time than necessary in general. Larger slack implies larger equilibrium headways. Consequently, the service frequency decreases.

The maximum term has different robustness when $\sum_{i=1}^{n} \gamma_i$ has different signs:

1. If $\sum_{i=1}^{n} \gamma_i > 0$,
$$\max_{v \in [v_l, v_u], b \in [b_l, b_u]} \left\{ -\frac{1+b}{nv} \sum_{i=1}^{n} \gamma_i \right\} = -\frac{1+b_l}{nv_u} \sum_{i=1}^{n} \gamma_i.$$
This term is always negative. So $m_u = 0$ is an upper bound of it no matter what b and v are even they are outside their confidence intervals. However, since we determine m_u according to the worst case, we are adding more delay time than necessary in general. Backward headway control scheme [11] is an example of this kind. $\sum_{i=1}^{n} \gamma_i > 0$ implies that the control scheme emphasizes more on "slowing down" an early bus to help its following bus catch up. "Slowing down" requires small slack.

2. If $\sum_{i=1}^{n} \gamma_i < 0$,
$$\max_{v \in [v_l, v_u], b \in [b_l, b_u]} \left\{ -\frac{1+b}{nv} \sum_{i=1}^{n} \gamma_i \right\} = -\frac{1+b_u}{nv_l} \sum_{i=1}^{n} \gamma_i.$$
This term increases as v decreases or b increases and it has no upper bound if v has no lower bound. We can determine m_u based only on some high percentage quantile of the parameters. Once v becomes too small to be out of the confidence interval, for example there is a snow storm, the equilibrium headways do not exist and the bus system becomes unstable. Also like the previous case, since we determine m_u according to almost the worst case, we are adding more delay time than necessary in general. Daganzo's control scheme [21] is an example of this kind. $\sum_{i=1}^{n} \gamma_i < 0$ implies that the control scheme emphasizes more on "speeding up" a late bus. But as discussed before, "speeding up" a bus in a setting that only allow delaying buses at the control points require large slack.

3. If $\sum_{i=1}^{n} \gamma_i = 0$,
$$\max_{v \in [v_l, v_u], b \in [b_l, b_u]} \left\{ -\frac{1+b}{nv} \sum_{i=1}^{n} \gamma_i \right\} = 0.$$

So we suggest setting $m_u = 0$. "Prefol" policy [68] is an example of this kind. Variation of bus velocity or the level of ridership does not influence the maximum term at all. The balance of "speeding up" and "slowing down" leads to efficient utilization of slack.

Summary

In the setting that only allows delaying buses at control points, slack is a mean to speed up a late bus by delaying all the other buses longer. The more a control scheme emphasizes on "speeding up", the larger slack it requires.

Considering the uncertainty of traffic conditions and the level of ridership, slack is determined according to the worst case. Control schemes that emphasize more on "speeding up" a late bus or "slowing down" an early bus have unnecessarily too large slack in general. And the former has the risk of leading to unstable headways. The balance of "speeding up" and "slowing down" leads to the most efficient utilization of slack.

3.5 Systemic Stability and Headway Forecasting Accuracy

In all the analysis above, we assume that all the headway information is known precisely. But in implementation, headways are forecasted so errors always exist. In this section, we study how the accuracy of headway forecasts affects the stability of the bus system. We will consider bounded unbiased forecasts of headways and conclude that headways will be stabilized within some intervals instead of converging to a fixed point. The length of the intervals depends on the number of headways forecasted and the bound of the forecasting errors.

Suppose that the forecast of h_i^t, denoted $\widehat{h_i^t}$, is unbiased with errors, denoted δ_i^t, bounded by a constant $\overline{\delta}$ for all i and t. Then the delay time becomes

$$
\begin{aligned}
D^t &= \gamma_0 + \sum_{i=1}^{n} \gamma_i \widehat{h}_i^t \\
&= \gamma_0 + \sum_{i=1}^{n} \gamma_i (h_i^t + \delta_i^t) \\
&= \gamma_0 + (1+b) \sum_{i=1}^{n} \gamma_i s_i^t + (1+b) \sum_{i=1}^{n} \gamma_i \delta_i^t
\end{aligned}
$$

Let

$$
\varepsilon^t = (1+b) \sum_{i=1}^{n} \gamma_i \delta_i^t.
$$

Then ε^t is bounded by $\overline{\varepsilon}$ where

$$
\overline{\varepsilon} = (1+b)\ \overline{\delta} \sum_{i=1}^{n} |\gamma_i|. \tag{27}
$$

The delay time equals

$$
D^t = \gamma_0 + (1+b) \sum_{i=1}^{n} \gamma_i s_i^t + \varepsilon^t
$$

and the bus dynamics become

$$
\begin{aligned}
s_1^{t+1} &= s_n^t + b(s_n^t - s_1^t) - (1+b) \sum_{i=1}^{n} \gamma_i s_i^t - \gamma_0 - \varepsilon^t \\
s_2^{t+1} &= s_1^t + b(s_1^t - s_2^t) + (1+b) \sum_{i=1}^{n} \gamma_i s_i^t + \gamma_0 + \varepsilon^t \\
s_i^{t+1} &= s_{i-1}^t + b(s_{i-1}^t - s_i^t) \quad \forall i = 3, \ldots, n.
\end{aligned}
$$

Its vector form is

$$
\mathbf{s^{t+1}} = A\mathbf{s^t} + \mathbf{r} + \boldsymbol{\varepsilon}^{\mathbf{t}} \tag{28}
$$

where $\boldsymbol{\varepsilon}^{\mathbf{t}} = [-\varepsilon^t, \varepsilon^t, 0, \ldots, 0]^T$.

We show in the following theorem that if the second largest eigenvalue of A is less than 1, the headways will be stablized in a bounded interval containing the fixed point $\mathbf{s}^*$ of the deterministic system

$$
\mathbf{s^{t+1}} = A\mathbf{s^t} + \mathbf{r}. \tag{29}
$$

and the bound is proportional to the magnitude of the errors.

Theorem 3.5.1 *In bus dynamics* (28)*, if* $|\lambda_2(A)| < 1$ *and* $|\varepsilon^{\mathbf{t}}| \leq \bar{\varepsilon}$*, then for any positive* η*, there exists an large enough integer* N *such that for any* $t > N$*, it holds that*

$$\|\mathbf{s^t} - \mathbf{s}^*\| \leq \eta + \frac{C\,\bar{\varepsilon}}{1 - |\lambda_2(A)|}$$

where C *is a function of* A *and* $\mathbf{s}^*$ *is the equilibrium headways of the bus dynamics* (29)*.*

Proof. When $|\lambda_2(A)| < 1$, we can transform the non-homogeneous equations into homogeneous ones like what was done in Step 2 in the proof of Theorem 3.2.2:

$$\mathbf{y^{t+1}} = A\mathbf{y^t} + \varepsilon^{\mathbf{t}}$$

We transform A to its Jordan form—There exists an invertible matrix V such that

$$A = V^{-1}JV.$$

Since $|\lambda_2(A)| < 1$, we assume that without loss of generality the first block of J is 1 and the other blocks have diagonal entries with magnitude less than 1:

$$J = \begin{pmatrix} 1 & 0 \\ 0 & J_1 \end{pmatrix}$$

Let V be

$$\begin{pmatrix} - & v_1 & - \\ & \vdots & \\ - & v_n & - \end{pmatrix}$$

Since the first row of V is the row eigenvector corresponding to the eigenvalue 1, v_1 is in $span([1, 1, \ldots, 1])$:

$$v_1 = c[1, 1, \ldots, 1].$$

Since $\sum_{i=1}^{n} \varepsilon_{\mathbf{i}}^{\mathbf{t}} = 0$, the first entry of $V\varepsilon^{\mathbf{t}}$ is zero. Let $\mathbf{z^t}$ be the $n-1$ vector by deleting the first entry from $V\varepsilon^{\mathbf{t}}$. Then

$$JV\varepsilon^{\mathbf{t}} = J_1\mathbf{z^t}$$

So

$$\|JV\varepsilon^{\mathbf{t}}\|_2 = \|J_1\mathbf{z}^{\mathbf{t}}\|_2 \leq \|J_1\|_2\,|\mathbf{z}^{\mathbf{t}}| = \|J_1\|_2\,|V\varepsilon^{\mathbf{t}}| \leq \|J_1\|_2\,\|V\|_2\,|\varepsilon^{\mathbf{t}}| \leq \|J_1\|_2\,\|V\|_2\,\overline{\varepsilon}$$

The explicit expression of $\mathbf{y}^{\mathbf{t}}$ is

$$\mathbf{y}^{\mathbf{t}} = A^t\mathbf{y}^{\mathbf{0}} + \sum_{\tau=0}^{t-1} A^\tau \varepsilon^{\mathbf{t}-\tau}$$

So

$$\mathbf{y}^{\mathbf{t}} - \mathbf{y}^* = A^t\mathbf{y}^{\mathbf{0}} - \mathbf{y}^* + \sum_{\tau=0}^{t-1} \overline{A}^\tau \varepsilon^{\mathbf{t}-\tau}$$

Replacing A with $V^{-1}JV$, we have

$$\mathbf{y}^{\mathbf{t}} - \mathbf{y}^* = A^t\mathbf{y}^{\mathbf{0}} - \mathbf{y}^* + \sum_{\tau=0}^{t-1} V^{-1}J^\tau V\varepsilon^{\mathbf{t}-\tau}$$

Since $|\lambda_2(A)| < 1$, it holds that

$$\|J_1\|_2 = |\lambda_2(A)| < 1.$$

Then we have

$$\begin{aligned}
\|\mathbf{y}^{\mathbf{t}} - \mathbf{y}^*\|_2 &= \|A^t\mathbf{y}^{\mathbf{0}} - \mathbf{y}^* + \sum_{\tau=0}^{t-1} V^{-1}J^\tau V\varepsilon^{\mathbf{t}-\tau}\|_2 \\
&\leq \|A^t\mathbf{y}^{\mathbf{0}} - \mathbf{y}^*\|_2 + \|\sum_{\tau=0}^{t-1} V^{-1}J^\tau V\varepsilon^{\mathbf{t}-\tau}\|_2 \\
&\leq \|A^t\mathbf{y}^{\mathbf{0}} - \mathbf{y}^*\|_2 + \sum_{\tau=0}^{t-1} \|V^{-1}\|_2\,\|J^\tau V\varepsilon^{\mathbf{t}-\tau}\|_2 \\
&\leq \|A^t\mathbf{y}^{\mathbf{0}} - \mathbf{y}^*\|_2 + \sum_{\tau=0}^{t-1} \|V^{-1}\|_2\,\|J_1^\tau\|_2\,\|V\|_2\,|\varepsilon^{\mathbf{t}-\tau}| \\
&= \|A^t\mathbf{y}^{\mathbf{0}} - \mathbf{y}^*\|_2 + \sum_{\tau=0}^{t-1} \|V^{-1}|_2\,|\lambda_2(A)|^\tau\,\|V\|_2\,|\varepsilon^{\mathbf{t}-\tau}| \\
&\leq \|A^t\mathbf{y}^{\mathbf{0}} - \mathbf{y}^*\|_2 + \left(\sum_{\tau=0}^{t-1} |\lambda_2(A)|^\tau\right) \|V^{-1}\|_2\,\|V\|_2\,\overline{\varepsilon} \\
&\leq \|A^t\mathbf{y}^{\mathbf{0}} - \mathbf{y}^*\|_2 + \frac{1}{1-|\lambda_2(A)|}\|V^{-1}\|_2\,\|V\|_2\,\overline{\varepsilon}
\end{aligned}$$

Since

$$\lim_{t\to\infty} \|A^t\mathbf{y^0} - \mathbf{y}^*\|_2 = 0,$$

letting $C = \|V^{-1}\|_2 \, \|V\|_2$ we finish the proof. $\square$

Combining Theorem 3.5.1 and Equation (27), we find that the final traveling time deviation is proportional to the bound of headway forecast. And it is likely that the deviation is larger if more headway forecasts are used.

3.6 Process to Determine a Control Scheme

In this chapter, we propose a class of control schemes that affinely depends on real-time headways. We discuss equilibrium headways, systemic stability and its sensitivity to level of ridership, how to determine slack, and the impact of headway forecasting accuracy on systemic stability. Here is the process to determine the parameters of a control scheme:

1. Determine which headways to use, say h_j^t, $j \in U \subseteq \{1, 2, \ldots, n\}$. Then set $\gamma_i = 0$, $\forall i \in \{1, 2, \ldots, n\} \backslash U$. [By flexibility of the general form, Section 3.1]

2. Determine whether the slack is independent of bus velocity and the level of ridership. If yes, set constraint $\sum_{i=1}^{n} \gamma_i = 0$. [By robustness of slack, Section 3.4]

3. Obtain the value of γ_i, $i = 1, 2, \ldots, n$ by optimizing the systemic stability. [By self-equalization of headways, Section 3.2]

4. Obtain the value of γ_0 using Equation (26). [By robustness of slack, Section 3.4]

3.6.1 An Example: Backward Headway Control

The backward headway of a bus refers to the headway of its following bus. Bartholdi and Eisenstein [11] proposed a self-coordinating control scheme that only depends on backward headway, but they did not provide a way to determine the coefficients and any quantitative robustness analysis. We use this backward headway control scheme as an example to show how to use our framework.

Step 1: Determine the form of delay time

Suppose there are 5 buses in the route. In this setting, only h_5^t is used. So we have $\gamma_1 = \gamma_2 = \cdots = \gamma_4 = 0$. The formula of delay time is

$$D^t = \gamma_0 + \gamma_5 h_5^t.$$

Step 2: Determine whether the slack depends on the parameters

A "Yes" leads to that $\gamma_5 = 0$. In this case, we are not using the backward headway information and the delay time is a constant γ_0. Headways under this control will diverge. So say "No".

Step 3: Determine γ_i, $i = 1, 2, \ldots, 5$

First, we estimate the boarding parameter b. Suppose $b = 0.02$. Figure 4 displays the relation between $|\lambda_2(A(b))|$ and γ_5. When $\gamma_5 < 0.1$ or $\gamma_5 > 0.92$, $|\lambda_2(A(b))| \geq 1$. It implies that the backward control scheme definitely fails and results in bunching. The smallest $|\lambda_2(A(b))|$ is achieved when $\gamma_n = 0.58$, so we set $\gamma_n = 0.58$ which optimizes the systemic stablity. This process shows the simplicity and strength of our framework.

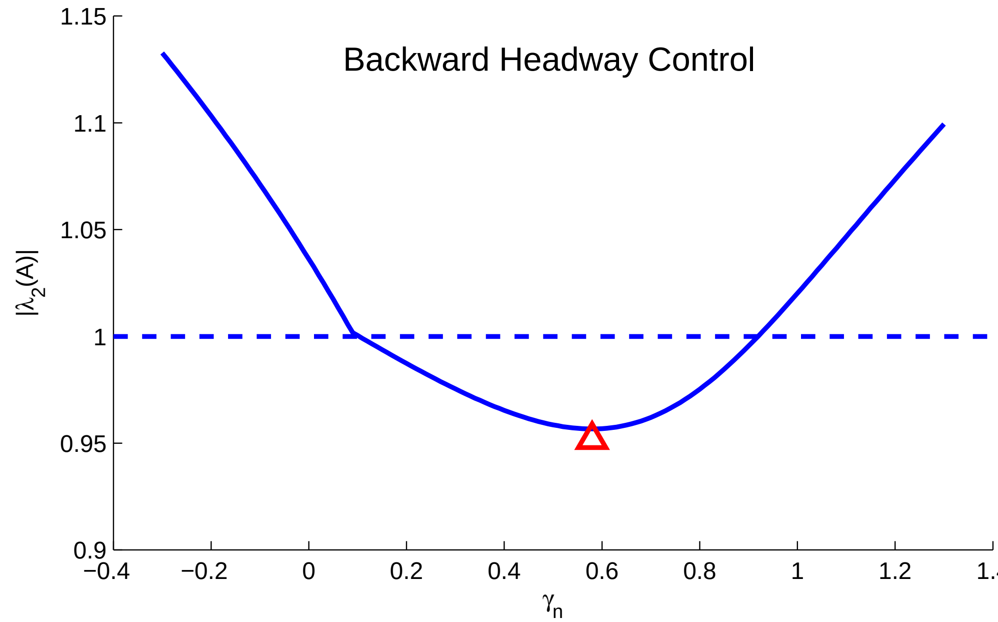

Figure 4: Relation between the convergence rate $|\lambda_2(A)|$ and γ_n when $b = 0.02$ and $n = 5$

The smallest $|\lambda_2(A(b))|$ is 0.9567 which is close to 1. It implies slow convergence of headways. When b increases to 0.07, $|\lambda_2(A)| > 1$, which implies that headways diverge and lead to bunching under any backward headway control.

Step 4: Determine γ_0

According to Equation (26),

$$\gamma_0 = m_u + m_c$$

Since $\sum_{i=1}^{5} \gamma_i > 0$, we set $m_u = 0$ according to Section 3.4. The delay time at the equilibrium state is

$$D^* = \frac{5m_c + \frac{1+b}{v}\gamma_5}{5 - \gamma_5}.$$

Service provider can set m_c to be an appropriate value so that this delay time is enough for bus drivers' break.

CHAPTER IV

A PRACTICAL CONTROL SCHEME

The analysis framework developed in Chapter 3 serves as a tool to select effective and robust control schemes. In this chapter, we propose a practical control scheme that has some desired properties. This control scheme is a hybrid of two simple control schemes: Threshold control and "Prefol".

We show that headways under the threshold control self-equalize fastest among all control schemes in the general class, but the price for robustness is high. In contrast, the stability of the bus system under "Prefol" is almost independent of traffic conditions and the level of ridership, but headways self-equalize slower. We combine both controls so that the hybrid control scheme inherits fast headway self-equalization from threshold control and small slack to maintain robustness from "Prefol".

4.1 Threshold Control

Threshold control is the binary target headway control described in Section 1.1.2 in the literature review. In binary target-headway control, the bus at the control point is not delayed if its forward headway is larger than a predetermined static threshold which may be different from the target headway, otherwise it is delayed until its forward headway equals the threshold.

We call "forward headway control" the control that depends only on the forward headway. We start with forward headway control and will conclude that threshold control has the fastest headway convergence among all forward headway control schemes.

We assume that the forward headway is the only knowledge we have, so the delay time has the following form:

$$D^t = \gamma_0 + \gamma_1 h_1^t$$

4.1.1 Headway Self-Equalization

In this part, we concentrate on studying the convergence rate of the headway self-equalization process. We determine γ_1 by minimizing the convergence rate, $|\lambda_2(A(b))|$.

It is interesting to find that γ_1 is always equal to -1 when it minimizes the convergence rate no matter what b and n are. Figure 5, 6 and 7 display the relations between $|\lambda_2(A(b))|$ and γ_1 for different b and n.

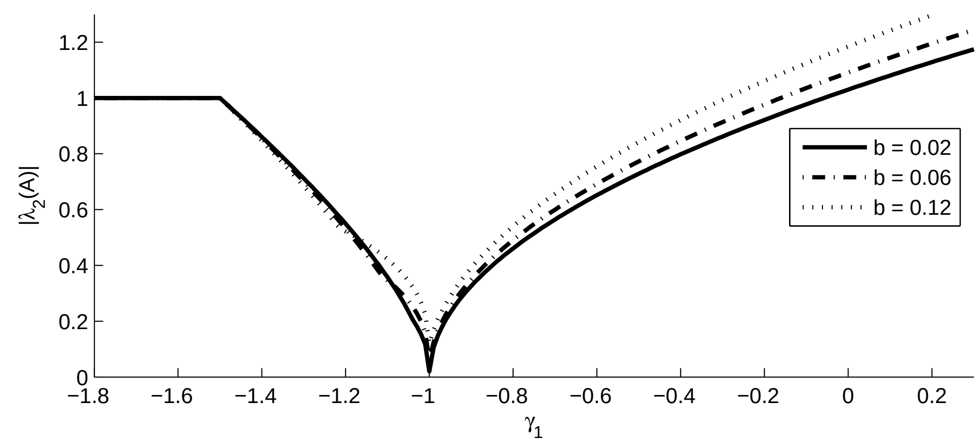

Figure 5: Relation between $|\lambda_2(A)|$ and γ_1 when $n = 3$

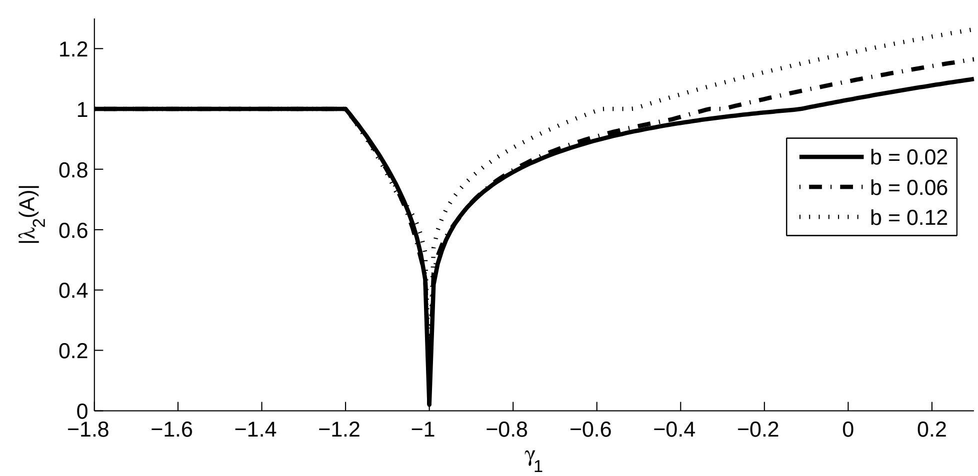

Figure 6: Relation between $|\lambda_2(A)|$ and γ_1 when $n = 6$

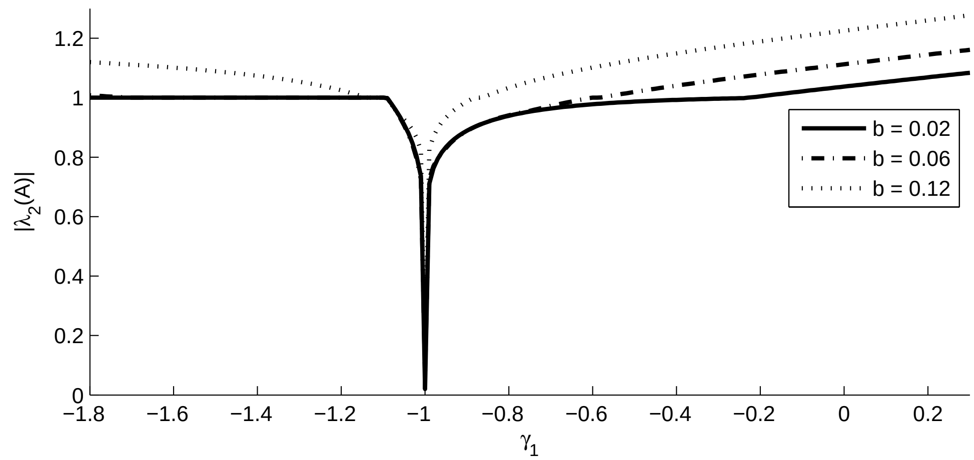

Figure 7: Relation between $|\lambda_2(A)|$ and γ_1 when $n = 12$

Mathematically, when $\gamma_1 = -1$, the transition matrix becomes

$$A = \begin{pmatrix} 1 & & & & & 1+b \\ 0 & -b & & & & \\ & 1+b & -b & & & \\ & & \ddots & \ddots & & \\ & & & & 1+b & -b \end{pmatrix}_{n\times n}$$

It is observed that the eigenvalues of the transition matrix are 1 of multiplicity 1 and $-b$ of multiplicity $n-1$. So the modulus of the second largest eigenvalue is

$$|\lambda_2(A(b))| = |-b| = b,$$

and thus the convergence rate is b which is very small, implying extremely fast convergence and strong resistance to bunching.

When $\gamma_1 = -1$, the delay formula becomes

$$D^t = \gamma_0 - h_1^t$$

Surprisingly, this is exactly the form of delay time of threshold control where γ_0 is the threshold. Intuitively, it implies that we delay every bus at the control point until the preceding bus has left the control point for γ_0 amount of time. Hence, the headway of the first bus always has the same value γ_0. This value tends to remain stable till the bus comes back to the control point again because buses with similar headways spend similar amount of time on boarding passengers.

4.1.2 Fastest Convergence Rate Overall

We conjecture that the forward headway control scheme with $\gamma_1 = -1$ converges fastest among all linear control schemes even if more headway information is available. Mathematically, this is equivalent to the statement:

Conjecture 4.1.1 $[\gamma_1, \gamma_2, \ldots, \gamma_n] = [-1, 0, \ldots, 0]$ *minimizes* $|\lambda_2(A)|$.

Two facts support this conjecture.

1. $|\lambda_2(A)| = b$. b is a small number in the possible range of $|\lambda_2(A)|$.

2. From the discussion above, we know that when $\gamma_i = 0$, $i = 2, 3, \ldots, n$, $|\lambda_2(A)|$ is minimized at $\gamma_1 = -1$. So $\frac{\partial|\lambda_2(A)|}{\partial\gamma_1}\Big|_{[\gamma_1,\gamma_2,\ldots,\gamma_n]=[-1,0,\ldots,0]} = 0$. We claim that $\frac{\partial|\lambda_2(A)|}{\partial\gamma_i}\Big|_{[\gamma_1,\gamma_2,\ldots,\gamma_n]=[-1,0,\ldots,0]} = 0$, $i = 2, 3, \ldots, n$. So $[\gamma_1, \gamma_2, \ldots, \gamma_n] = [-1, 0, \ldots, 0]$ satisfies the first order optimality condition.

Claim 4.1.1

$$\frac{\partial|\lambda_2(A)|}{\partial\gamma_i}\bigg|_{[\gamma_1,\gamma_2,\ldots,\gamma_n]=[-1,0,\ldots,0]} = 0, \quad i = 2, 3, \ldots, n.$$

Proof. The eigenvalues of A is the roots of Equation (21). Hence for any $i \in \{2, 3, \ldots, n\}$, when $\gamma_1 = -1$ and $\gamma_j = 0$, $\forall j \in 2, 3, \ldots, n \backslash \{i\}$, the eigenvalues except 1 satisfy

$$(-1-b)\left(\frac{-b-\lambda}{-1-b}\right) + \gamma_i \left(\frac{-b-\lambda}{-1-b}\right)^{n-i} = 0.$$

It follows that

$$\left[\left(\frac{-b-\lambda}{-1-b}\right)^{i-1} - \frac{\gamma_i}{1+b}\right]\left(\frac{-b-\lambda}{-1-b}\right)^{n-i} = 0$$

So there are $n-i$ solutions that are equal to $-b$. They are constant. The derivative of them is 0. For the other $i-1$ solutions that satisfy

$$\left(\frac{-b-\lambda}{-1-b}\right)^{i-1} - \frac{\gamma_i}{1+b} = 0,$$

it holds that

$$\frac{-b-\lambda}{-1-b} = \left(\frac{\gamma_i}{1+b}\right)^{\frac{1}{i-1}} e^{\frac{2k\pi}{i-1}\mathrm{i}}$$

So

$$\lambda = (1+b)\left(\frac{\gamma_i}{1+b}\right)^{\frac{1}{i-1}} e^{\frac{2k\pi}{i-1}\mathrm{i}} - b.$$

Therefore,

$$\left.\frac{d\lambda}{d\gamma_i}\right|_{\gamma_i=0} = \left.\frac{1}{i-1}\left(\frac{\gamma_i}{1+b}\right)^{\frac{1}{i-1}-1} e^{\frac{2k\pi}{i-1}\mathrm{i}}\right|_{\gamma_i=0} = 0.$$

This gives the wanted result.

$$\left.\frac{\partial|\lambda_2(A)|}{\partial\gamma_i}\right|_{[\gamma_1,\gamma_2,\ldots,\gamma_n]=[-1,0,\ldots,0]} = 0, \quad i = 2, 3, \ldots, n.$$

□

Even if the conjecture is not true, the convergence rate of the headways with the forward headway control scheme is fast enough.

4.1.3 Equilibrium Headways

According to Equation (19), when $\gamma_0 = -1$, the equilibrium headways are

$$h_1^* + D^* = h_i^* = \gamma_0, \quad \forall i = 2, 3, \ldots, n.$$

where the delay time in the equilibrium state is

$$D^* = n\gamma_0 - (1+b)\frac{1}{v}.$$

Since $D^* \geq 0$, it holds that

$$\gamma_0 \geq (1+b)\frac{1}{nv}$$

Since $\frac{1+b}{v}$ is the loop time of a bus, this addresses that γ_0 is not smaller than the loop time divided by the number of buses. Intuitively, once the opposite occurs, there is at least one bus falling behind and it can never catch up with the others. Then this forward headway control scheme fails and the bus service is not regular.

The term $(1+b)\frac{1}{nv}$ depends on the average velocity and the number of the buses. Smaller velocity makes it larger. So a lower bound of velocity is the key factor to determine γ_0. On the other hand, if buses breaking down is taken into account, then a lower bound of number of buses is also crucial. In order to avoid the risk of the system breaking down, γ_0 should be determined by lower bounds of the velocity and the number of buses. If so, however, γ_0 is too large for most of the time when the velocity and the number of buses is not that small, which makes delay time too large and wastes bus capacity.

For example, suppose that v ranges from v_{lb} to v_{ub} and n can be 5 or 4. Also suppose that $v_{lb} = 0.8v_{av}$ and $v_{ub} = 1.2v_{av}$ where v_{av} is the average. Then γ_0 should not be smaller than $(1+b)\frac{1}{4v_{lb}}$. Consider the ideal case that $b = 0$ and there is no disturbance. Let γ_0 be exactly $\frac{1}{4v_{lb}}$. In this way, the equilibrium delay time equals 0 when $v = v_{lb}$ and $n = 4$, but can be as large as $\frac{51}{4v_{lb}} - \frac{1}{v_{ub}} = \frac{351}{48v_{av}}$ when $v = v_{ub}$ and $n = 5$. The actual loop time in this case is only $\frac{1}{v_{ub}} = \frac{51}{12v_{av}}$, which is only 57% of $\frac{351}{48v_{av}}$. This is the ideal case. In fact, γ_0 is set larger to prevent the damage caused by disturbance by random events like traffic lights. Thus more time is wasted on lowering the risk of irregular service.

4.1.4 Simplicity

The threshold control is simple in three aspects:

1. From the prospective of information requirement, only the forward headway of the bus at the control point is needed. A watch or a clock is enough to catch this information. So the capital cost and maintenance cost is low.

2. From the prospective of control scheme design, there is only one coefficient to determine, namely γ_0. And this is exactly the equilibrium headway we want. The determination of γ_0 can be obtained from daily data. We can set different values for γ_0 for different time period throughout a day. Generally, morning

peaks and evening peaks require larger values than early morning and midday.

3. From the prospective of scalability, the forward headway control scheme can be easily extended to bus service with more than one control point. Just use the same formula for all control points. Control points work as processors in series, so the convergence rate becomes $|\lambda_2(A)|^m$ where m is the number of control points. And the equilibrium headways remain the same.

4.1.5 As a Component of Hybrid Control Schemes

In this part, we show that the fast convergence property and predictable equilibrium headway enable the threshold to be an component in powerful hybrid control schemes.

Suppose that there is another convergent control scheme $C1$ with delay time being D_1^t that produces equilibrium headway s^*. Let a hybrid control scheme have the following delay time:

$$D^t = \max\{D_1^t, \gamma_0 - h_1^t\}. \tag{30}$$

Recall that the equilibrium headway of threshold control is γ_0. Two possible scenarios would happen:

1. $\gamma_0 < s^*$. Then after n iterations most of the headways will be greater than γ_0 . The subsequent dynamics will mainly be controlled by D_1^t and drive headways to converge to s^*. So the convergence rate mainly depends on $C1$. The threshold control component plays a role in speeding up the convergence process at the beginning by imposing a planned headway.

2. $\gamma_0 \geq s^*$. Then for most of the time, $\gamma_0 - h_1^t \geq D_1^t$ and thus the forward headway control component is activated. The headways converge to γ_0 with convergence rate b.

This hybrid control schemes utilizes the fast convergence property of threshold control and remedy its defect of the large slack requirement to be robust. Suppose that v ranges from v_l to v_u and remains around v_m for most of the time. Then it is advisable to set γ_0 slightly greater than $\frac{1+b}{nv_m}$ so that for most of the time the headways converge fast to γ_0. And the other component of the hybrid scheme handles the situation when $\gamma_0 < \frac{1}{nv_m}$.

4.1.6 Conclusion

Threshold control has serval properties:

1. It is simple and economical in implementation. It requires only forecast of forward headways and statistics of bus loop time.
2. The headways with it converge exponentially fast with convergence rate b.
3. It can serve as a component in a hybrid control scheme.
4. It is not easy to determine the target headway. If the target headway is too large, it is a waste of slack. If it is too small, headways may not self-equalize.

4.2 Prefol Control

The name "Prefol" refers to the combination of "previous" and "following". "Prefol" was first proposed by Turnquist and Bowman as an approximate solution to a stochastic programming minimizing average waiting time [68]. They showed that this scheme is effective when consecutive headways have negative covariance and suggest applying this scheme at bus stops which have large arrival rate and few on-board passengers.

In this section, we further study this scheme. We pick out the one that drives headways equalize fast.

4.2.1 Convergence Rate

In our standard approach, we determine γ_1 and γ_n by minimizing $|\lambda(A)|$. However, the minimizer is $[\gamma_1, \gamma_n] = [-1, 0]$. The corresponding form of delay time is exactly that of threshold control. But threshold control suffers from the variation of bus velocity. When the bus velocity becomes small, the predetermined target headway may be too small to result in bunching. When the bus velocity is large, the target headway may be too large to result in waste of capacity. In Section 3.4, we show that when $\sum_{i=1}^{n} \gamma_i = 0$, the equilibrium delay time is independent of bus velocity. So we add this constraint to the optimization problem.

$$\begin{aligned} \min_{\gamma_1, \gamma_n} \quad & |\lambda_2(A)| \\ \text{s.t.} \quad & \gamma_1 + \gamma_n = 0 \end{aligned}$$

Here is an example. When $n = 5$ and $b = 0.02$, Figure 8 displays the relation between $|\lambda_2(A(b))|$ and γ_n. When $\gamma_n < 0.05$ or $\gamma_n > 0.86$, we have $|\lambda_2(A(b))| \geq 1$. It implies that the Prefol control scheme definitely fails and results in bunching. The smallest $|\lambda_2(A(b))|$ is achieved when $\gamma_n = 0.49$. When $\gamma_n < 0.49$, $|\lambda_2(A(b))|$ is an non-increasing function of γ_n; when $\gamma_n > 0.49$, $|\lambda_2(A(b))|$ is an non-decreasing function of γ_n. So when $n = 5$ and $b = 0.02$, we can just set $\gamma_n = 0.49$. This process shows the simplicity and strength of our analysis.

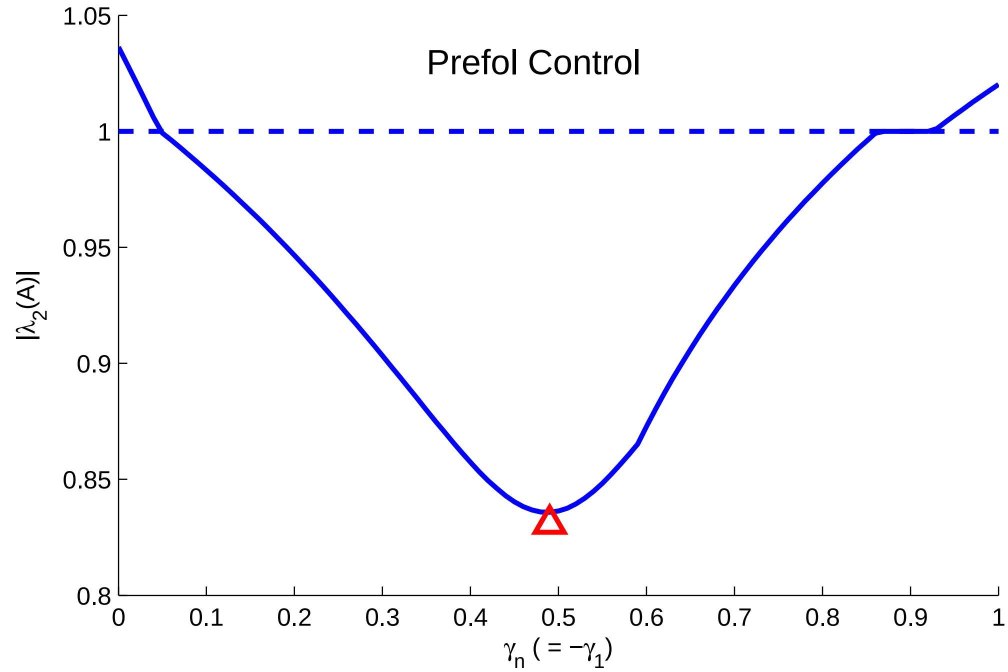

Figure 8: Relation between the convergence rate $|\lambda_2(A)|$ and γ_n when $b = 0.02$ and $n = 5$

The smallest $|\lambda_2(A(b))|$ is 0.8358 which is large compared to that of the forward headway control scheme and small compared to that of the backward headway control scheme.

4.2.1.1 When Level of Ridership is Unknown

When we determine the value of γ_n by optimizing $|\lambda_2(A)|$, we assume that we know the value of b exactly. However, there exists error when we estimate b. In practice, we only know the range of b, say from b_l to b_u. To determine the value of γ_n, we can minimize the maximum value of $|\lambda_2 A|$ over all $b \in [b_l, b_u]$.

$$\begin{aligned}
&\min_{\gamma_n} \quad \max_{b_l \le b \le b_u} |\lambda_2(A(b))| \\
&\text{s.t.} \quad \gamma_1 = -\gamma_n
\end{aligned}$$

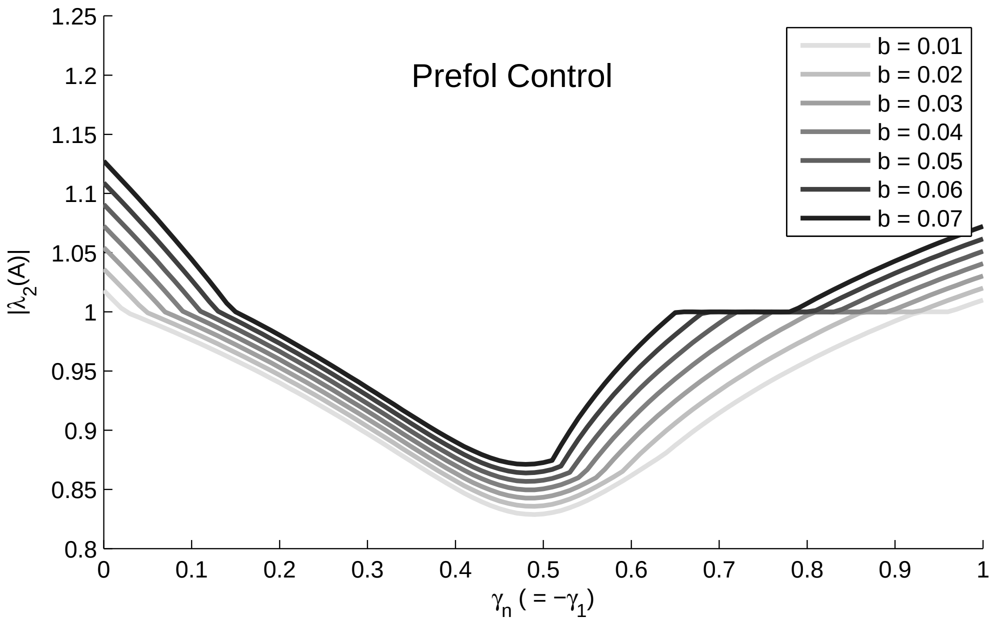

Figure 9: Relation between the convergence rate $|\lambda_2(A)|$ and γ_n when $n = 5$

Suppose that $n = 5$ and $0.01 \leq b \leq 0.07$. The solution to the corresponding robust problem is $\gamma_n = 0.48$. Figure 9 displays the relation between $|\lambda_2(A(b))|$ and γ_n with different values of b. In this case, $|\lambda_2(A(b))|$ is an increasing function of b for any fixed γ_n. So actually the robust optimization problem can be simplified to be

$$\begin{aligned} \min_{\gamma_n} \quad & |\lambda_2(A(b_u))| \\ \text{s.t.} \quad & \gamma_n = -\gamma_1 \end{aligned} \tag{31}$$

In fact, a larger b results in higher tendency to bunch, and so it is harder for a control scheme to handle. So generally we can solve the optimization problem (31) to obtain the value of γ_n.

4.2.2 Equilibrium Headways and Delay Time

According to Equation (19), the stationary headways are

$$h_1^* + D^* = h_i^* = \frac{(1+b)(1+\gamma_1)\frac{1}{v} + \gamma_0}{n(1+\gamma_1)}, \quad \forall i = 2, 3, \ldots, n.$$

where the delay time

$$D^* = \frac{\gamma_0}{1 + \gamma_1}$$

The equilibrium delay time is positive as long as γ_0 is positive.

More importantly, the equilibrium delay time is totally determined by the parameters we set. Both the bus velocity and the level of ridership have no influence on it. So there is never a waste of capacity. We don't need to consider changing γ_0 from time to time. The Prefol control scheme is unique among all control schemes that linearly depend on local headways in the sense of owning this celebrated independent property, because $\gamma_1 + \gamma_n = 0$ only holds here.

4.2.3 Simplicity

The "Prefol" is simple in three aspects:

1. From the prospective of information requirement, only the forward and backward headways of the bus at the control point is needed. A few beacons installed close to the control point along the bus route are enough to catch this information precisely.

2. From the prospective of control scheme design, γ_n can be easily solved by minimizing the convergence rate. The convergence rate with this parameter is robust against the variation of level of ridership. Also, once γ_0 is specified, there is no need to change it again. The equilibrium headway is invariant to the variation of both bus velocity and the level of ridership.

3. From the prospective of scalability, the "Prefol" can be easily extent to bus service with more than one control points. Just use the same formula for all control points. Control points work as processors in series. So the convergence rate becomes $|\lambda_2(A)|^m$ where m is the number of control points.

4.3 A Practical Control Scheme

"Prefol" and threshold are perfectly complementary. The bus system under threshold control has fastest convergence rate, but it is sensitive to the variation of bus velocity. The bus system under "Prefol" has comparatively slower convergence rate, but it is adaptive to the variation of bus velocity. The hybrid scheme made of them retains their advantages and discards their disadvantages.

The delay time of the hybrid scheme

$$D^t = \max\{\gamma_0 + \gamma_1 h_1^t + \gamma_n h_n^t, \gamma_0' - h_1^t\} \tag{32}$$

According to the analysis in Section 4.1.5, the headways with this hybrid scheme self-equalize from any initial values.

4.4 Comparison on a Simulated Route

In this section, we examine the performances of several control schemes in a realistic setting. We compare bus systems under four control methods—target schedule control, forward headway control [21], backward headway control [11] and the practical hybrid control we propose in Section 4.3—on a simulated bus route. We show that the practical hybrid control scheme outperforms the others.

Bus bunching is complained a lot at Chicago Transit Authority. Route 63 is a typical example. This route travels eastbound to Stony Island Avenue and back to Midway Airport along 63rd street. The entire loop is 17.75 miles long (28.57 kilometers). Our simulation is based on CTA data collected from GPS systems and automatic passenger counters on each bus. Route 63 has almost 80 stops, of which the CTA monitors GPS data from only 18, including the two control points, one each at the easternmost and westernmost ends of the route. The historical travel times between key stops is well-described as the sum of uniformly distributed times for each intervening city block (1/8 mile or 0.2 kilometers in length).

We matched the simulated passenger arrivals and departures with the historical daily patterns by proceeding as follows: From the data we set the total arrivals to the system every half hour over a 14 hour period, from 04:00 to 18:00. Arrivals and departures at key bus stops vary over the day according to four major time periods: AM Early, AM Peak, Midday, and Evening Peak (Figure 10). The mean arrival rate for each particular bus stop during a given time period was estimated by sampling from an exponential distribution with mean set to the mean number of boardings observed during that period. Dwell times at bus stops were computed based on the model of Milkovits (2008) which is nonlinear in the number of boarding, alighting and on-board passengers.

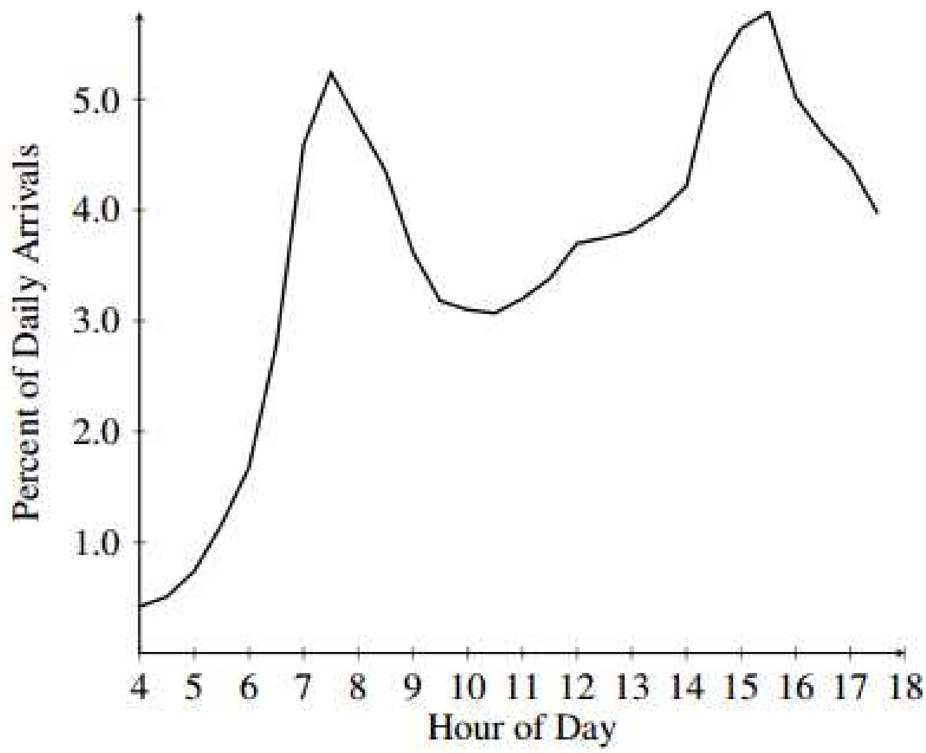

Figure 10: Arrival rates of passengers to CTA Route 63, showing morning and evening surges

We simulated the CTA route for a day under normal conditions and selected the best performing parameters under each of four control schemes: target schedule control, forward headway control, backward headway control and the practical hybrid control we propose in Section 4.3. All the control schemes utilize four control points.

In the early morning, buses run fast and there are few passengers. This period is ideally used to evenly spread out the buses along the route. We use the average and the coefficient of variation of the backward headway of the bus that just arrives at the first control point as performance measures.

When we choose the best parameters ex post facto for the control schemes, the practical hybrid control performed the best. Average headways under all the control schemes are similar (Figure 11) but the coefficient of variations of headways under the practical hybrid control is the smallest in all periods (Figure 12).

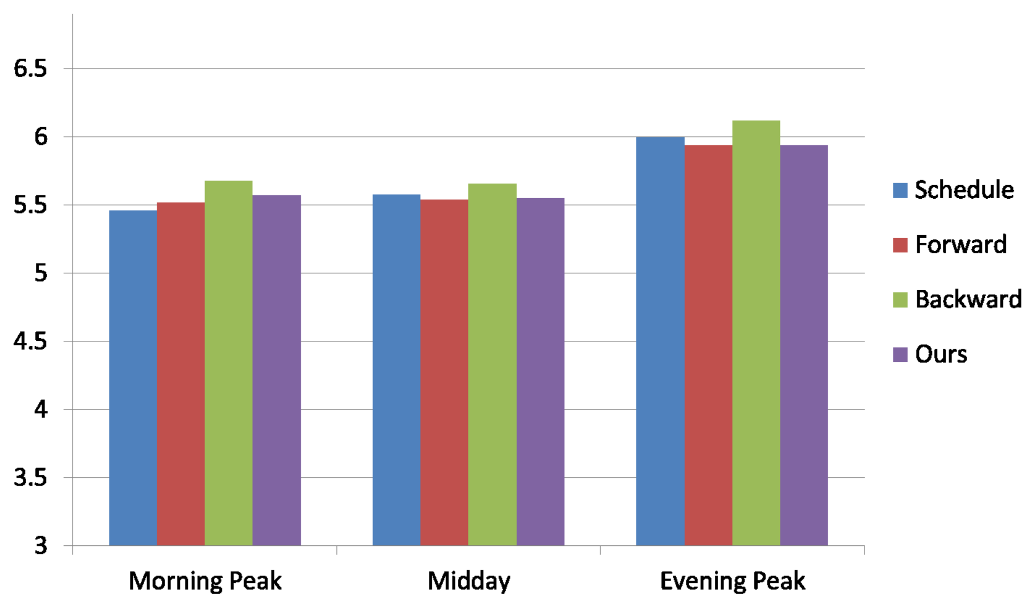

Figure 11: Average headways

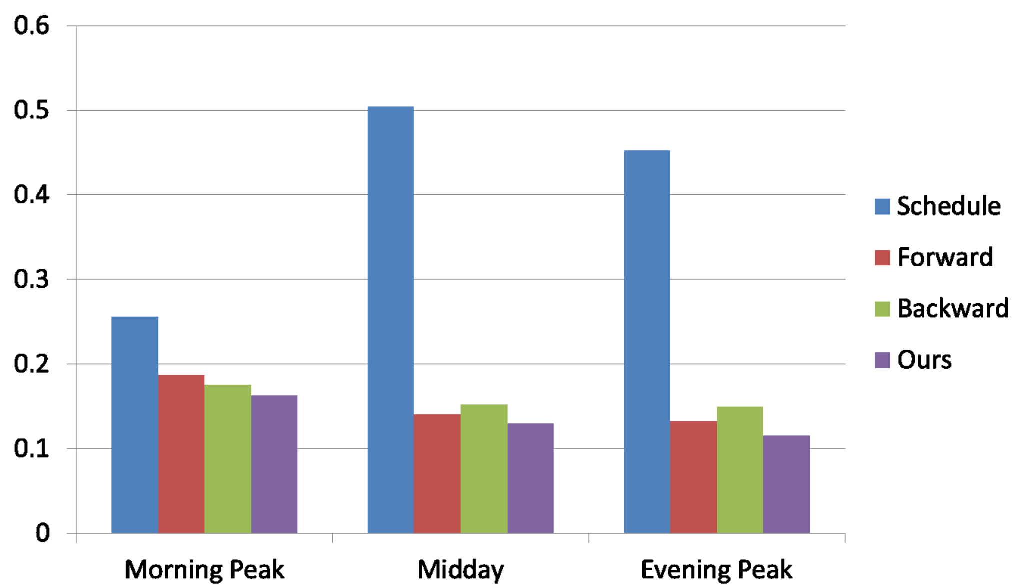

Figure 12: Coefficient of variations of headways

Of course, one cannot select parameters ex post facto in practice. The control scheme must be able to react to the shocks and variances on each day. To test this, we fixed the parameter settings, but re-ran the simulation to mimic a reduction in travel velocity by 10%, as might occur in rainy days.

Average headways under all the controls are similar (Figure 13) except that under forward headway control scheme the headways blew up in the evening peak. The coefficient of variations of headways under the practical hybrid control again is the smallest in all periods (Figure 14). This shows the robustness of the bus system under it.

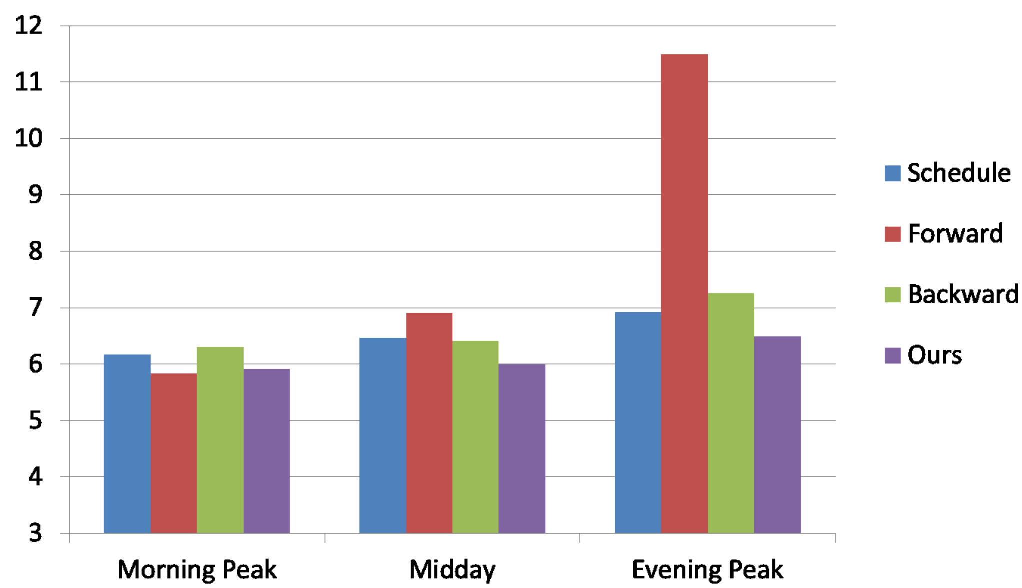

Figure 13: Average headways in rainy day

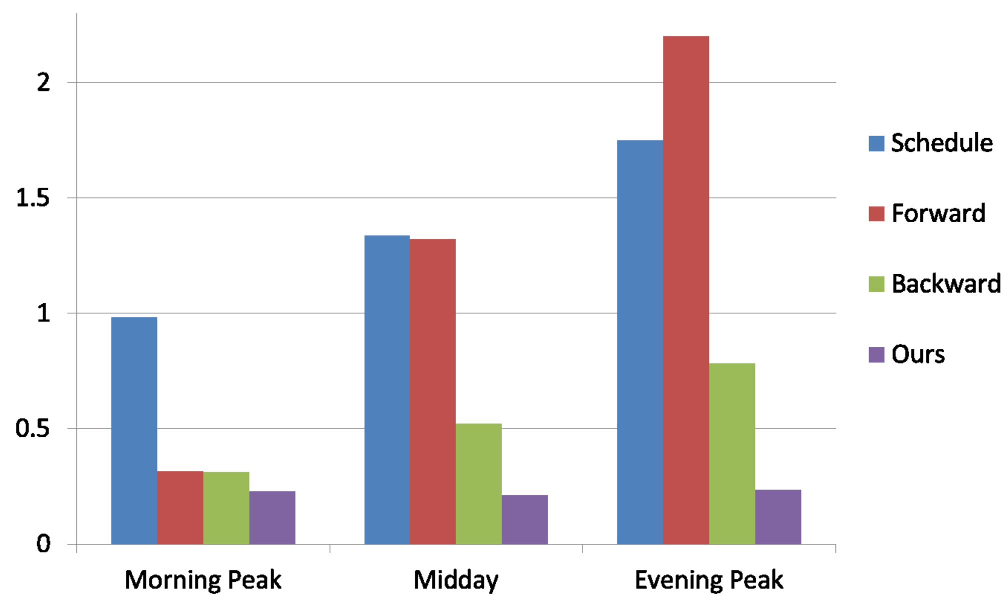

Figure 14: Coefficient of variations of headways in rainy day

APPENDIX A

PROOF

We show the proof of Theorem 3.2.2 when A is not diagonalizable here:

Theorem A.0.1 *If the modulus of the second largest eigenvalue of A, $|\lambda_2(A)|$, is less than* 1*, then*

1. *there exists unique stationary traveling times.*

2. *Given any initial traveling times, the traveling times converge.*

3. *The equilibrium traveling times are the stationary traveling times* $\mathbf{s}^*$ *in Proposition 3.2.1, and*

4. *There exists a constant C such that $\|A^t\mathbf{s^0} - \mathbf{s}^*\|_2 \leq C|\lambda_2(A)|^t, \forall t > n$ when A is not diagonalizable.*

Proof. Both Step 1 and 2 remain the same. Here is the step 3 in the proof of

Denote the stationary y^t by $\mathbf{y}^*$. Then $\mathbf{y}^* = \mathbf{s}^* + \mathbf{q}$ and $\mathbf{y}^* = A\mathbf{y}^*$. So $\mathbf{y}^*$ is an eigenvector of A corresponding to the eigenvalue 1.

Denote the distinct eigenvalues of A by $\lambda_2(A), \lambda_3(A), \ldots, \lambda_n(A)$ such that $|\lambda_2(A)| \geq |\lambda_3(A)| \geq \cdots \geq |\lambda_m(A)|$ corresponding to generalized eigenvectors $\{v_{21}, v_{22}, \ldots, v_{2k_2}\}$, $\{v_{31}, v_{32}, \ldots, v_{3k_3}\}$, $\ldots, \{v_{m1}, v_{m2}, \ldots, v_{mk_m}\}$ such that $Av_{i1} = \lambda_i(A)v_{i1}$ and $Av_{ij} = \lambda_i(A)v_{i,j-1}$, $\forall i = 2, 3, \ldots, m$ and $\forall j = 2, 3, \ldots, m_i$. k_i is the algebraic multiplicity of $\lambda_i(A)$, $\forall i = 2, 3, \ldots, m$.

Since $k_i < n$, $\forall i = 2, 3, \ldots, m$, it holds that $A^n v_{ij} = \lambda_i^n(A)v_{i1}$, $\forall i = 2, 3, \ldots, m$ and $\forall j = 1, 2, \ldots, k_i$.

We write $\mathbf{y^0}$ in terms of $\mathbf{y}^*$ and v_{ij}

$$\mathbf{y^0} = a_1\mathbf{y}^* + \sum_{i,j} a_{ij}v_{ij}$$

Left-multiplying A for n times on both sides, we have

$$A^n\mathbf{y^0} = A^n a_1\mathbf{y}^* + A^n\sum_{i,j} a_{ij}v_{ij} = a_1\mathbf{y}^* + \sum_{i} a_{ij}\lambda_i^n(A)v_{i1} \tag{33}$$

Again left-multiplying A for $t-n(>0)$ times on both sides of Equation (33), we have

$$A^t\mathbf{y^0} = a_1\mathbf{y}^* + \sum_{i,j} a_{ij}\lambda_i^t(A)v_{i1} \tag{34}$$

Since $|\lambda_i(A)| < 1$, $\forall i = 2, 3, \ldots, m$, it follows that $\lim_{t\to\infty}\lambda_i^t(A) = 0$, $\forall i = 2, 3, \ldots, m$. Hence the limit of the right hand side of Equation (34) exists. So does the left hand side. So $\lim_{\to\infty} A^t\mathbf{y^0} = \mathbf{y}^*$ by uniqueness of the stationary traveling times. Taking limit for both sides of Equation (34), we get $a_1 = 1$. So

$$A^t\mathbf{y^0} = \mathbf{y}^* + \sum_{i,j} a_{ij}\lambda_i^t(A)v_{i1}$$

It follows that

$$\begin{aligned}\|A^t\mathbf{y^0} - \mathbf{y}^*\|_2 &= \|\sum_{i,j} a_{ij}\lambda_i^t(A)v_{i1}\|_2 \\ &\leq |\lambda_2(A)|^t\sum_{i,j}\|a_{ij}v_{i1}\|_2 \\ &\leq C|\lambda_2(A)|^t\end{aligned}$$

for some C.

Since $A^t\mathbf{y^0} - \mathbf{y}^* = A^t(\mathbf{y^0} - \mathbf{y}^*) = A^t(\mathbf{s^0} - \mathbf{s}^*) = A^t\mathbf{s^0} - \mathbf{s}^*$, it holds that

$$\|A^t\mathbf{s^0} - \mathbf{s}^*\| \leq C|\lambda_2(A)|^t.$$

So given any initial traveling times, the traveling times converge to the equilibrium traveling times which are $\mathbf{s}^*$.

□

REFERENCES

[1] Abkowitz, M. and Tozzi, J., "Transit route characteristics and headway-based reliability control," Transportation Research Record, no. 1078, 1986.

[2] Abkowitz, M., Eiger, A., and Engelstein, I., "Optimal control of headway variation on transit routes," Journal of Advanced Transportation, vol. 20, no. 1, pp. 73–88, 1986.

[3] Abkowitz, M. and Engelstein, I., Methods for maintaining transit service regularity. No. 961, 1984.

[4] Abkowitz, M. D. and Lepofsky, M., "Implementing headway-based reliability control on transit routes," Journal of Transportation Engineering, vol. 116, no. 1, pp. 49–63, 1990.

[5] Adamski, A. and Turnau, A., "Simulation support tool for real-time dispatching control in public transport," Transportation Research Part A: Policy and Practice, vol. 32, no. 2, pp. 73–87, 1998.

[6] Andersson, P.-Å., Hermansson, Å., Tengvald, E., and Scalia-Tomba, G.-P., "Analysis and simulation of an urban bus route," Transportation Research Part A: General, vol. 13, no. 6, pp. 439–466, 1979.

[7] Andersson, P.-Å. and Scalia-Tomba, G.-P., "A mathematical model of an urban bus route," Transportation Research Part B: Methodological, vol. 15, no. 4, pp. 249–266, 1981.

[8] Barnett, A., "On controlling randomness in transit operations," Transportation Science, vol. 8, no. 2, pp. 102–116, 1974.

[9] Barnett, A. I., "Control strategies for transport systems with nonlinear waiting costs," Transportation Science, vol. 12, no. 2, pp. 119–136, 1978.

[10] Bartholdi, J. J., Clark, R. J., Williamson, D. W., Eisenstein, D. D., and Platzman, L., "Building a self-organizing urban bus route," in Self-Adaptive and Self-Organizing Systems Workshops (SASOW), 2012 IEEE Sixth International Conference on, pp. 66–70, IEEE, 2012.

[11] Bartholdi, J. J. I. and Eisenstein, D. D., "A self-coördinating bus route to resist bus bunching," Transportation Research Part B: Methodological, vol. 46, no. 4, pp. 481–491, 2012.

[12] Bartholdi III, J. J. and Eisenstein, D. D., "A production line that balances itself," Operations Research, vol. 44, no. 1, pp. 21–34, 1996.

[13] Bartholdi III, J. J., Eisenstein, D. D., and Foley, R. D., "Performance of bucket brigades when work is stochastic," Operations research, vol. 49, no. 5, pp. 710–719, 2001.

[14] Berrebi, S. J., Watkins, K. E., and Laval, J. A., "A real-time bus dispatching policy to minimize passenger wait on a high frequency route," Transportation Research Part B: Methodological, vol. 81, pp. 377–389, 2015.

[15] Bly, P. and Jackson, R. L., "Evaluation of bus control strategies by simulation," 1974.

[16] Bowman, L. A. and Turnquist, M. A., "Service frequency, schedule reliability and passenger wait times at transit stops," Transportation Research Part A: General, vol. 15, no. 6, pp. 465–471, 1981.

[17] Bullock, P., Jiang, Q., and Stopher, P. R., "Using gps technology to measure on-time running of scheduled bus services," Journal of Public Transportation, vol. 8, no. 1, pp. 21–40, 2005.

[18] Carey, M., "Reliability of interconnected scheduled services," European Journal of Operational Research, vol. 79, no. 1, pp. 51–72, 1994.

[19] Chen, Q., Adida, E., and Lin, J., "Implementation of an iterative headway-based bus holding strategy with real-time information," Public Transport, vol. 4, no. 3, pp. 165–186, 2013.

[20] Cortés, C. E., Sáez, D., Milla, F., Núñez, A., and Riquelme, M., "Hybrid predictive control for real-time optimization of public transport systems operations based on evolutionary multi-objective optimization," Transportation Research Part C: Emerging Technologies, vol. 18, no. 5, pp. 757–769, 2010.

[21] Daganzo, C. F., "A headway-based approach to eliminate bus bunching: Systematic analysis and comparisons," Transportation Research Part B: Methodological, vol. 43, no. 10, pp. 913–921, 2009.

[22] Daganzo, C. F. and Pilachowski, J., "Reducing bunching with bus-to-bus cooperation," Transportation Research Part B: Methodological, vol. 45, no. 1, pp. 267–277, 2011.

[23] Delgado, F., Muñoz, J., Giesen, R., and Cipriano, A., "Real-time control of buses in a transit corridor based on vehicle holding and boarding limits," Transportation Research Record: Journal of the Transportation Research Board, no. 2090, pp. 59–67, 2009.

[24] Delgado, F., Munoz, J. C., and Giesen, R., "How much can holding and/or limiting boarding improve transit performance?," Transportation Research Part B: Methodological, vol. 46, no. 9, pp. 1202–1217, 2012.

[25] Dessouky, M., Hall, R., Nowroozi, A., and Mourikas, K., "Bus dispatching at timed transfer transit stations using bus tracking technology," Transportation Research Part C: Emerging Technologies, vol. 7, no. 4, pp. 187–208, 1999.

[26] Dessouky, M., Hall, R., Zhang, L., and Singh, A., "Real-time control of buses for schedule coordination at a terminal," Transportation Research Part A: Policy and Practice, vol. 37, no. 2, pp. 145–164, 2003.

[27] Ding, Y. and Chien, S., "Improving transit service quality and headway regularity with real-time control," Transportation Research Record: Journal of the Transportation Research Board, no. 1760, pp. 161–170, 2001.

[28] Eberlein, X. J., Wilson, N. H., and Bernstein, D., "Modeling real-time control strategies in public transit operations," in Computer-aided transit scheduling, pp. 325–346, Springer, 1999.

[29] Eberlein, X. J., Wilson, N. H., and Bernstein, D., "The holding problem with real–time information available," Transportation science, vol. 35, no. 1, pp. 1–18, 2001.

[30] Fu, L., Liu, Q., and Calamai, P., "Real-time optimization model for dynamic scheduling of transit operations," Transportation Research Record: Journal of the Transportation Research Board, no. 1857, pp. 48–55, 2003.

[31] Fu, L. and Yang, X., "Design and implementation of bus-holding control strategies with real-time information," Transportation Research Record: Journal of the Transportation Research Board, no. 1791, pp. 6–12, 2002.

[32] Group, K. and others, "Transit capacity and quality of service manual," 2013.

[33] He, S.-X., "An anti-bunching strategy to improve bus schedule and headway reliability by making use of the available accurate information," Computers & Industrial Engineering, vol. 85, pp. 17–32, 2015.

[34] Hernández-Landa, L. G., Morales-Marroquín, M. L., Nigenda, R. S., and Ríos-Solís, Y. Á., "Linear bus holding model for real-time traffic network control," in Applied Simulation and Optimization, pp. 303–319, Springer, 2015.

[35] Hickman, M. D., "An analytic stochastic model for the transit vehicle holding problem," Transportation Science, vol. 35, no. 3, pp. 215–237, 2001.

[36] Horn, R. A. and Johnson, C. R., Matrix analysis. Cambridge university press, 2012.

[37] Jolliffe, J. and Hutchinson, T., "A behavioural explanation of the association between bus and passenger arrivals at a bus stop," Transportation Science, vol. 9, no. 3, pp. 248–282, 1975.

[38] Koffman, D., A simulation study of alternative real-time bus headway control strategies. No. 663, 1978.

[39] Lesley, L., "The role of the timetable in maintaining bus service reliability," in Proceedings of operating public transport symposium, 1975.

[40] Lin, G.-S., Liang, P., Schonfeld, P., and Larson, R., "Adaptive control of transit operations," 1995.

[41] Liu, G. and Wirasinghe, S., "A simulation model of reliable schedule design for a fixed transit route," Journal of Advanced Transportation, vol. 35, no. 2, pp. 145–174, 2001.

[42] Marguier, P. H. and Ceder, A., "Passenger waiting strategies for overlapping bus routes," Transportation Science, vol. 18, no. 3, pp. 207–230, 1984.

[43] Mazloumi, E., Mesbah, M., Ceder, A., Moridpour, S., and Currie, G., "Efficient transit schedule design of timing points: a comparison of ant colony and genetic algorithms," Transportation Research Part B: Methodological, vol. 46, no. 1, pp. 217–234, 2012.

[44] Milkovits, M. N., Simulating service reliability of a high frequency bus route using automatically collected data. PhD thesis, Massachusetts Institute of Technology, 2008.

[45] Minser, J., "Manager, market research," Chicago Transit Authority. Personal communication, 2009.

[46] NAKANISHI, Y., "Bus performance indicators: On-time performance and service regularity," Transportation research record, no. 1571, pp. 3–13, 1997.

[47] Newell, G., "Unstable brownian motion of a bus trip," in Statistical Mechanics and Statistical Methods in Theory and Application, pp. 645–667, Springer, 1977.

[48] Newell, G. F., "Control of pairing of vehicles on a public transportation route, two vehicles, one control point," Transportation Science, vol. 8, no. 3, pp. 248–264, 1974.

[49] Newell, G. F. and Potts, R. B., "Maintaining a bus schedule," in Australian Road Research Board (ARRB) Conference, 2nd, 1964, Melbourne, vol. 2, 1964.

[50] ODell, S. W. and Wilson, N. H., "Optimal real-time control strategies for rail transit operations during disruptions," in Computer-aided transit scheduling, pp. 299–323, Springer, 1999.

[51] Okrent, M. M., Effects of transit service characteristics on passenger waiting time. PhD thesis, Northwestern University, 1974.

[52] Osuna, E. and Newell, G., "Control strategies for an idealized public transportation system," Transportation Science, vol. 6, no. 1, pp. 52–72, 1972.

[53] Platt, B., "The pulse: Bus bunching the problem for North York riders." http://www.thestar.com/news/gta/2014/07/29/the_pulse_bus_bunching_the_problem_for_north_york_riders.html, 2014.

[54] Pritchard, G., Vines, J., Briggs, P., Thomas, L., and Olivier, P., "Digitally driven: how location based services impact the work practices of london bus drivers," in Proceedings of the 32nd annual ACM conference on Human factors in computing systems, pp. 3617–3626, ACM, 2014.

[55] Puong, A. and Wilson, N. H., "A train holding model for urban rail transit systems," in Computer-aided systems in public transport, pp. 319–337, Springer, 2008.

[56] Schmitt, F. and Rothlauf, F., "On the importance of the second largest eigenvalue on the convergence rate of genetic algorithms," in Proceedings of the 14th Symposium on Reliable Distributed Systems, Citeseer, 2001.

[57] Senevirante, P. N., "Analysis of on-time performance of bus services using simulation," Journal of Transportation Engineering, vol. 116, no. 4, pp. 517–531, 1990.

[58] Shen, S. and Wilson, N. H., "An optimal integrated real-time disruption control model for rail transit systems," in Computer-aided scheduling of public transport, pp. 335–363, Springer, 2001.

[59] Sontag, E. D., Mathematical control theory: deterministic finite dimensional systems, vol. 6. Springer Science & Business Media, 1998.

[60] Sontag, E. D. and Boyd, S., "Mathematical control theory: Deterministic finite-dimensional systems," IEEE Transactions on Automatic Control, vol. 40, no. 3, pp. 563–563, 1995.

[61] Strathman, J. G. and Hopper, J. R., "Empirical analysis of bus transit on-time performance," Transportation Research Part A: Policy and Practice, vol. 27, no. 2, pp. 93–100, 1993.

[62] Strathman, J. G., Kimpel, T. J., Dueker, K. J., and Northwest, T., "Bus transit operations control: review and an experiment involving tri-met's automated bus dispatching system," tech. rep., Transportation Northwest, Department of Civil Engineering, University of Washington, 2000.

[63] Suh, W., Chon, K.-S., and Rhee, S.-M., "Effect of skip-stop policy on a korean subway system," Transportation Research Record: Journal of the Transportation Research Board, no. 1793, pp. 33–39, 2002.

[64] Sun, A. and Hickman, M., "The real–time stop–skipping problem," Journal of Intelligent Transportation Systems, vol. 9, no. 2, pp. 91–109, 2005.

[65] Sun, A. and Hickman, M., "The holding problem at multiple holding stations," in Computer-aided systems in public transport, pp. 339–359, Springer, 2008.

[66] Turnquist, M. A., "Strategies for improving reliability of bus transit service," Transportation Research Record, no. 818, 1981.

[67] Turnquist, M. A. and Blume, S. W., Evaluating potential effectiveness of headway control strategies for transit systems. No. 746, 1980.

[68] Turnquist, M. A. and Bowman, L. A., "Control of service reliability in transit networks," tech. rep., 1979.

[69] Van Der Aa, N., Ter Morsche, H., and Mattheij, R., "Computation of eigenvalue and eigenvector derivatives for a general complex-valued eigensystem," Electronic Journal of Linear Algebra, vol. 16, no. 1, pp. 300–314, 2007.

[70] Van Oort, N., Wilson, N., and Van Nes, R., "Reliability improvement in short headway transit services: Schedule-and headway-based holding strategies," Transportation Research Record: Journal of the Transportation Research Board, no. 2143, pp. 67–76, 2010.

[71] Vandebona, U. and Richardson, A. J., "Effect of checkpoint control strategies in a simulated transit operation," Transportation Research Part A: General, vol. 20, no. 6, pp. 429–436, 1986.

[72] Wirasinghe, S., "Cost based approach to scheduling travel time on a public transportation route.," Transportation and traffic theory, 1993.

[73] Wirasinghe, S. C. and Liu, G., "Determination of the number and locations of time points in transit schedule designcase of a single run," Annals of Operations Research, vol. 60, no. 1, pp. 161–191, 1995.

[74] Xuan, Y., Argote, J., and Daganzo, C. F., "Dynamic bus holding strategies for schedule reliability: Optimal linear control and performance analysis," Transportation Research Part B: Methodological, vol. 45, no. 10, pp. 1831–1845, 2011.

[75] Yan, Y., Meng, Q., Wang, S., and Guo, X., "Robust optimization model of schedule design for a fixed bus route," Transportation Research Part C: Emerging Technologies, vol. 25, pp. 113–121, 2012.

[76] Zhao, J., Bukkapatnam, S., and Dessouky, M. M., "Distributed architecture for real-time coordination of bus holding in transit networks," Intelligent Transportation Systems, IEEE Transactions on, vol. 4, no. 1, pp. 43–51, 2003.

[77] Zhao, J., Dessouky, M., and Bukkapatnam, S., "Optimal slack time for schedule-based transit operations," Transportation Science, vol. 40, no. 4, pp. 529–539, 2006.

[78] Zolfaghari, S., Azizi, N., and Jaber, M. Y., "A model for holding strategy in public transit systems with real-time information," International Journal of Transport Management, vol. 2, no. 2, pp. 99–110, 2004.