Iterative Tuning Strategy for Setting Phase Splits with Anticipation of Traffic Demand in Urban Traffic Network

Yu Wang, Danwei Wang, Shangtai Jin, Nan Xiao, Yitong Li and Emilio Frazzoli

Abstract

Facing large-scale urban traffic network, countless effort has been made toward intelligent and efficient urban traffic control system to better use existing traffic infrastructures. Recently, a novelty pre-timed traffic signal control strategy known as Iterative Tuning (IT) has been developed by exploiting repetitive characteristic of junction's vehicle throughput on working days, which is sufficiently efficient in under-saturated traffic conditions. This paper further improves IT strategy in saturated traffic conditions with consideration of traffic demand including vehicle throughput and residual queued vehicles. Unlike conventional pre-timed strategies, preparation of signal schedules is not required in IT strategy and fine-tuning process executes iteratively and automatically. This paper proposes a generalized traffic model to describe urban network dynamics and explicit split tuning algorithm. Without specific control trajectories, rigorous analysis generates sufficient condition for guaranteeing the convergence of IT strategy globally over repetitions. The robustness of IT controller to variations of traffic flow patterns and errors of initial conditions is also analysed in details. Commonwealth Avenue, an area with heavy traffic in Singapore, is demonstrated in simulations and simulation results indicate the effectiveness and robustness of IT strategy.

Index Terms

Iterative Learning Control, Urban Traffic Signal Control, Traffic Flow Patterns.

I. Introduction

With ever-increasing traffic mobility, urban traffic signal control plays an important role to better utilize existing road infrastructures. Roughly speaking, traffic signal control is classified as pre-timed signal control and traffic-responsive signal control [1]. Recently most research focuses on traffic-responsive signal control, such as Store-and-Forward approaches [2], BLX [3], etc. While large numbers of signal controllers in use currently are still pre-timed, and pre-timed signal schedules can be coupled with traffic-responsive signal control to fine-tune signal control parameters to have better control performance [4]. Therefore, this paper aims to further develop pre-timed signal strategy in urban traffic signal control.

For pre-timed signal control, the prerequisite is the repetitive characteristic of traffic demand. Festin [5] studied the general daily traffic profiles in the United States and five areas for the period 1970-1995 and showed that daily traffic profiles were repetitive on a weekly basis. Roess *et al.* [6] showed that typical variations of daily traffic patterns were around 13% and 16% on weekdays and weekends, respectively. Wang *et al.* [7] evaluated the phase-based traffic flow patterns in Singapore and indicated the coefficients of variations were below 20% and 35% during peak hours on working days and weekends, respectively.

Since 1950s, traffic signal settings were proposed to optimize phase splits and cycle length for isolated junction [8]. As increasing traffic demand varied largely in morning and evening peak hours, Time-of-day (TOD) approach was proposed to segment a day into several time intervals during which different signal plans are applied. Hauser *et al.* [9] considered the application of data mining tools to help design switching points in traffic signal control. Smith *et al.* [10] proposed a computational approach to automate the identification of intervals within allowable cluster size using a high-resolution definition of system states and optimized the signal timing plans based on the set of archived data. Brian Park *et al.* [11] optimized the breakpoints through minimization of within-cluster distance and maximization of between-cluster distances. Wong *et al.* [12] used clustering algorithms to re-estimate the TOD intervals and optimize the signal timing plans for each TOD interval. Guo and Zhang [13] proposed the time of traffic occurring as one dimension in clustering to consider the coordination of a corridor. In each grouped cluster, suitable signal plans were calculated by commercial simulation tools, such as TRANSYT series, Synchro series, etc.

For pre-timed strategies, preparation and fine-tuning process of signal schedules for new junctions are quite time-consuming and labour-intensive. Meanwhile, traffic was estimated to experience an additional 3-5% delay time per year as a consequence

Yu Wang and Danwei Wang are with Centre for System Intelligence and Efficiency (EXQUISITUS), School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, wang0676@e.ntu.edu.sg; edwwang@ntu.edu.sg

Shangtai Jin is with the Advanced Control Systems Laboratory, School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, PRC. shtjin@bjtu.edu.cn

Nan Xiao and Yitong Li are with the Singapore-MIT Alliance for Research and Technology Centre, Singapore 138602, Singapore, xiaonan@smart.mit.edu; yitong@smart.mit.edu

Emilio Frazzoli is with the Massachusetts Institute of Technology, Cambridge MA 02139, USA, frazzoli@mit.edu

of gradual changing traffic conditions [14]. Signal re-timing process is required after several years when performance of the signal control system gets relatively worse.

To enable the fine-tuning capability of pre-timed signal control, iterative learning approach was proposed to automatic fine-tune signal schedules. Traditionally, Iterative Tuning Learning (ILC) methodology [15] is usually applied to fine-tune control input to systems that operate repeatedly and to track a specific reference over repetitions. The application of ILC includes many industrial systems [16] in manufacturing, robotics [17] and chemical processing. This approach had been first applied for density control of freeway traffic flows with repeatable patterns and achieved robust performance [18]. Huang *et al.* [19] proposed an iterative learning approach by using traffic assignment model to find the global optimal signal control and flow patterns, then traffic signal controller was designed to drive traffic patterns approaching the desired flow patterns. Wang *et al.* [20] proposed Iterative Tuning (IT) strategy by considering repetitive vehicle throughput and obtained considerable performance. IT was the first trial to apply iterative learning approach without specific control reference.

In order to further improve the performance of IT strategy, repetitive traffic demand is used in this paper to tune traffic signals. Traffic demand includes not only the vehicle throughput at a junction, but also residual queued vehicles before the stop line of junctions. In case of saturated condition, traffic demand rather than vehicle throughput reveals the real traffic conditions.

The main contribution of this paper with respect to proposed IT strategy [20] includes: 1) IT offers generalized nonlinear model to describe traffic dynamics and explicit algorithm for split tuning. The preparation of initial signal plans is arbitrary and the fine-tuning process executes iteratively and automatically. 2) As traffic flows fluctuate largely during different time for every junction, it is difficulty to provide desired control trajectories for traditional ILC approaches. Without trajectories, error function with traffic demand is proposed in this paper. Rigorous analysis provides sufficient condition for guaranteeing the convergence of IT strategy over repetitions. 3) Urban traffic network in Singapore with real traffic demand is simulated and validates the convergence, efficiency and robustness of IT strategy.

In this paper, cycle time and offset are assumed to be fixed; only phase split is tuned here. The rest of this paper is organized as follows: Section II states the problem of urban traffic signal control. Section III describes IT strategy in details and shows the convergence of IT strategy with traffic variations. Section IV presents case studies based on real traffic flows. Section V analyses simulation results, validates the convergence and efficiency of IT strategy. Section VI concludes this paper and suggests some topics for further research.

II. Problem Statement

In this section, a generalized traffic model is proposed to describe traffic dynamics.

A. Notation

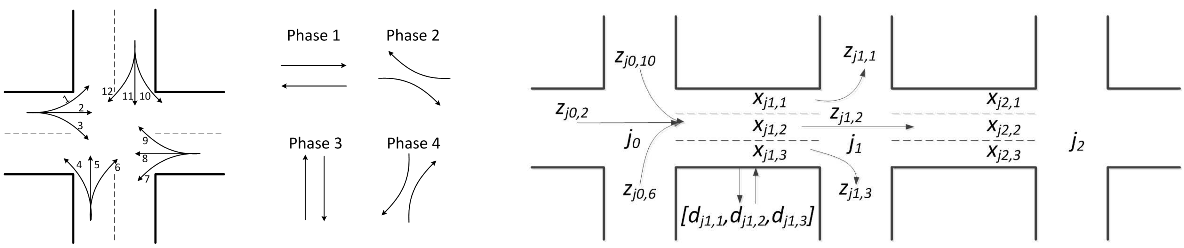
This paper considers traffic dynamics based on lane group $i \in F_j$, where F_j is the set of all lane groups of junction j. Referring to [6], lane group is defined for one or more lanes of a link approaching one junction that have the same direction. Traffic flows are controlled by phase time, which is the green duration of phase $p \in P_j$, where P_j is the set of all phases of junction j. Evolution of traffic dynamics is described cycle by cycle. Main variables defined for junction $j \in J$ and entire network are listed here, where J is the set of junctions.

k	index of cycles, and $k \in [0, K]$, where K is maximum number of cycles.
$x_{j,i}(k)$	number of vehicles on lane group i of junction j when signal turns from red to green at cycle k.
$z_{j,i}(k)$	vehicle throughput on lane group i of junction j during green signal at cycle k.
$q_{j,i}(k)$	number of residual queued vehicles on lane group i of junction j when signal turns from green to red at cycle k.
$y_{j,i}(k)$	traffic demand on lane group i of junction j at cycle k.
$\widetilde{u}_{j,i}(k)$	green time duration for lane group i of junction j at cycle k.
$\widetilde{o}_{j,i}(k)$	degree of saturation (DS) for lane group i of junction j at cycle k.
$d_{j,i}(k)$	disturbances on lane group i of junction j at cycle k.
$u_{j,p}(k)$	green time duration of phase p for junction j at cycle k.
$o_{j,p}(k)$	critical DS for phase p of junction j at cycle k.
$e_{j,p}(k)$	error of critical DS for phase p of junction j at cycle k.
$N_{j,i}$	number of lane groups for junction j.
$\Gamma_{j,i}(k)$	$\{x_{j,i}(k),\ z_{j,i}(k),\ q_{j,i}(k),\ y_{j,i}(k),\ \widetilde{u}_{j,i}(k),\ \widetilde{o}_{j,i}(k),\ d_{j,i}(k)\}$, the set for lane group i of junction j at cycle k.
$\gamma_{j,i}(k)$	any element of $\Gamma_{j,i}(k)$, i.e., $\gamma_{j,i}(k) \in \Gamma_{j,i}(k)$.
$\boldsymbol{\Gamma}_j(k)$	$\{\mathbf{x}_j(k),\ \mathbf{z}_j(k),\ \mathbf{q}_j(k),\ \mathbf{y}_j(k),\ \widetilde{\mathbf{u}}_j(k),\ \widetilde{\mathbf{o}}_j(k),\ \mathbf{d}_j(k)\}$, the set for all lane groups of junction j at cycle k.
$\boldsymbol{\gamma}_j(k)$	any element of $\boldsymbol{\Gamma}_j(k)$, i.e., $\boldsymbol{\gamma}_j(k) \in \boldsymbol{\Gamma}_j(k)$. $\boldsymbol{\gamma}_j(k) = [\gamma_{j,1}(k), \gamma_{j,2}(k), \ldots, \gamma_{j,N_{j,i}}(k)]^\top$.
$N_{j,p}$	number of phases for junction j.
$\Xi_{j,p}(k)$	$\{u_{j,p}(k),\ o_{j,p}(k),\ e_{j,p}(k)\}$, the set for phase p of junction j at cycle k.
$\xi_{j,p}(k)$	any element of $\Xi_{j,p}(k)$, i.e., $\xi_{j,p}(k) \in \Xi_{j,p}(k)$.
$\boldsymbol{\Xi}_j(k)$	$\{\mathbf{u}_j(k),\ \mathbf{o}_j(k),\ \mathbf{e}_j(k)\}$, the set for all phases of junction j at cycle k.
$\boldsymbol{\xi}_j(k)$	any element of $\boldsymbol{\Xi}_j(k)$, i.e., $\boldsymbol{\xi}_j(k) \in \boldsymbol{\Xi}_j(k)$. $\boldsymbol{\xi}_j(k) = [\xi_{j,1}(k), \xi_{j,2}(k), \ldots, \xi_{j,N_{j,p}}(k)]^\top$.

N number of junctions in the network.
$\mathbf{\Gamma}(k)$ $\{\mathbf{x}(k)$, $\mathbf{z}(k)$, $\mathbf{q}(k)$, $\mathbf{y}(k)$, $\widetilde{\mathbf{u}}(k)$, $\widetilde{\mathbf{o}}(k)$ $\mathbf{d}(k)\}$, the set for all lane groups of entire network at cycle k.
$\boldsymbol{\gamma}(k)$ any element of $\mathbf{\Gamma}(k)$, i.e., $\boldsymbol{\gamma}(k) \in \mathbf{\Gamma}(k)$. $\boldsymbol{\gamma}(k) = [\boldsymbol{\gamma}_1^\top(k), \boldsymbol{\gamma}_2^\top(k), \ldots, \boldsymbol{\gamma}_N^\top(k)]^\top$.
$\Xi(k)$ $\{\mathbf{u}(k)$, $\mathbf{o}(k)$, $\mathbf{e}(k)\}$, the set for all phases of entire network at cycle k.
$\boldsymbol{\xi}(k)$ any element of $\Xi(k)$, i.e., $\boldsymbol{\xi}(k) \in \Xi(k)$. $\boldsymbol{\xi}(k) = [\boldsymbol{\xi}_1^\top(k), \boldsymbol{\xi}_2^\top(k), \ldots, \boldsymbol{\xi}_N^\top(k)]^\top$.
n index of iterations.
$\mathbf{\Gamma}(k,n)$ $\{\mathbf{x}(k,n)$, $\mathbf{z}(k,n)$, $\mathbf{q}(k,n)$, $\mathbf{y}(k,n)$, $\widetilde{\mathbf{u}}(k,n)$, $\widetilde{\mathbf{o}}(k,n)$, $\mathbf{d}(k,n)\}$, the set for all lane groups of entire network at cycle k in iteration n.
$\boldsymbol{\gamma}(k,n)$ $\boldsymbol{\gamma}(k,n)$ in iteration n. $\boldsymbol{\gamma}(k,n)$ represents any element of $\mathbf{\Gamma}(k,n)$, i.e., $\boldsymbol{\gamma}(k,n) \in \mathbf{\Gamma}(k,n)$.
$\Xi(k,n)$ $\{\mathbf{u}(k,n)$, $\mathbf{o}(k,n)$, $\mathbf{e}(k,n)\}$, the set for all phases of entire network at cycle k in iteration n.
$\boldsymbol{\xi}(k,n)$ $\boldsymbol{\xi}(k)$ in iteration n. $\boldsymbol{\xi}(k,n)$ represents any element of $\Xi(k,n)$, i.e., $\boldsymbol{\xi}(k,n) \in \Xi(k,n)$.

B. *Traffic Model*

For an urban network, junction's diagram is shown in Fig. 1. Fig. 1(a) shows the indexes of lane groups for single junction with phase specification. In Fig. 1(b), three adjacent junctions j_0, j_1 and j_2 are illustrated with lane groups, where the indexes of lane groups follow the similar way in Fig. 1(a).



(a) Single Junction (b) Three Junctions' Diagram

Fig. 1. Single junction with indexed lane groups and three junctions' diagram. In Fig. 1(b), three adjacent junctions j_0, j_1 and j_2 are illustrated with lane groups. The indexes of lane groups follow the similar way in Fig. 1(a).

As shown in Fig. 1(b), the dynamics of $x_{j_1,1}(k)$, $x_{j_1,2}(k)$, $x_{j_1,3}(k)$ are updated by vehicle throughput $z_{j_0,2}(k)$, $z_{j_0,6}(k)$, $z_{j_0,10}(k)$ and $z_{j_1,1}(k)$, $z_{j_1,2}(k)$, $z_{j_1,3}(k)$ at connectors at junctions j_0 and j_1, respectively with disturbances $d_{j_1,1}(k)$, $d_{j_1,2}(k)$, $d_{j_1,3}(k)$, i.e.,

$$\begin{aligned} x_{j_1,1}(k+1) =& x_{j_1,1}(k) - z_{j_1,1}(k) + r_{j_1,1}(k)(z_{j_0,2}(k) + z_{j_0,6}(k) + z_{j_0,10}(k)) + d_{j_1,1}(k) \\ x_{j_1,2}(k+1) =& x_{j_1,2}(k) - z_{j_1,2}(k) + r_{j_1,2}(k)(z_{j_0,2}(k) + z_{j_0,6}(k) + z_{j_0,10}(k)) + d_{j_1,2}(k) \\ x_{j_1,3}(k+1) =& x_{j_1,3}(k) - z_{j_1,3}(k) + r_{j_1,3}(k)(z_{j_0,2}(k) + z_{j_0,6}(k) + z_{j_0,10}(k)) + d_{j_1,3}(k) \end{aligned} \tag{1}$$

where $r_{j_1,1}(k)$, $r_{j_1,2}(k)$, $r_{j_1,3}(k)$ are turning ratios of vehicles driving into $x_{j_1,1}(k)$, $x_{j_1,2}(k)$, $x_{j_1,3}(k)$ of junction j_1, respectively, and satisfy $r_{j_1,1}(k) + r_{j_1,2}(k) + r_{j_1,3}(k) = 1$.

For junction j, traffic states $x_{j,i}(k), \forall i \in F_j$ is rewritten as state space model based on (1), i.e.,

$$\mathbf{x}_j(k+1) = \mathbf{x}_j(k) + B_j(k)\mathbf{z}_j(k) + \sum_{j^* \in J, j^* \neq j} B_{j^*}^j(k)\mathbf{z}_{j^*}(k) + C_j(k)\mathbf{d}_j(k) \tag{2}$$

where $\mathbf{x}_j(k), \mathbf{z}_j(k), \mathbf{d}_j(k) \in \mathbf{\Gamma}_j(k)$; matrices $B_j(k) = -I$, where I is the identity matrix; $B_{j^*}^j(k)$ and $C_j(k)$ contain network information, such as turning ratios and road connections; $j^* \in J, j^* \neq j$ indicates all junctions except junction j.

For the entire network with N junctions, traffic states $\mathbf{x}_j(k), \forall j \in J$ is rewritten based on (2),

$$\mathbf{x}(k+1) = \mathbf{x}(k) + B(k)\mathbf{z}(k) + C(k)\mathbf{d}(k) \tag{3}$$

where $\mathbf{x}(k), \mathbf{z}(k), \mathbf{d}(k) \in \mathbf{\Gamma}(k)$; $B(k)$ represents all the control effects of $B_j(k)$ and $B_{j^*}^j(k)$; $C(k)$ = diag$\{C_1(k)$, $C_2(k)$, …, $C_N(k)\}$.

Referring to *Traffic Engineering* [6], traffic demand $\mathbf{y}(k)$ indicates *arrival flows*, which is the number of vehicles that desire to travel pass a point during a specified period. In under-saturated conditions, phase durations are sufficient to clear all vehicles and there is enough space in the downstream links, no vehicles are queued. Otherwise in saturated conditions, large numbers of vehicles are queued until the next cycle. Therefore, demand $\mathbf{y}(k)$ is related to number of vehicles $\mathbf{x}(k)$ and disturbances $\mathbf{d}(k)$, which is formulated as nonlinear function, i.e.,

$$\mathbf{y}(k) = g(\mathbf{x}(k), \mathbf{d}(k)) \tag{4}$$

where $\mathbf{y}(k) \in \mathbf{\Gamma}(k)$.

Vehicle throughput $\mathbf{z}(k)$ is the minimum value amongst traffic demand $\mathbf{y}(k)$, number of vehicles driving through within allocated green time $\mathbf{u}(k)$, and available road space in the downstream links. Since maximum volumes in the next links are fixed, available road space is determined by vehicles $\mathbf{x}(k)$ on those links. Therefore, a nonlinear function is illustrated as,

$$\mathbf{z}(k) = f(\mathbf{x}(k), \mathbf{y}(k), \mathbf{u}(k), \mathbf{d}(k)) \tag{5}$$

where $\mathbf{u}(k) \in \mathbf{\Xi}(k)$.

Substituting (4) into (5), there exists,

$$\mathbf{z}(k) = f\left(\mathbf{x}(k), g(\mathbf{x}(k), \mathbf{d}(k)), \mathbf{u}(k), \mathbf{d}(k)\right) = f(\mathbf{x}(k), \mathbf{u}(k), \mathbf{d}(k)) \tag{6}$$

For the entire network, substituting (6) into (3), and traffic demand (4), state space model (7)-(8) represents the traffic dynamics,

$$\mathbf{x}(k+1) = \mathbf{x}(k) + B(k) f(\mathbf{x}(k), \mathbf{u}(k), \mathbf{d}(k)) + C(k)\mathbf{d}(k) \tag{7}$$

$$\mathbf{y}(k) = g(\mathbf{x}(k), \mathbf{d}(k)) \tag{8}$$

The above state space model is a generalized form to characterize the dynamics with consideration of upstream and downstream vehicles. In S model [3], vehicle throughput and demand per sampling time are expressed as leaving average flow rate and arriving average flow rate, respectively. Store-and-forward models [2] also try to formulate the function of vehicle throughput mathematically. IT strategy is a data-driven strategy and two nonlinear functions are not required in details. Vehicle throughput and demand are anticipated based on historical repetitive traffic pattern. Traffic model (7)-(8) are only used for convergence analysis.

C. Control Objective

In this paper, the objective of IT strategy is to balance critical degree of saturation (DS) for all junctions during the entire period iteratively. From Chapter 21 of *Traffic Engineering* [6], the *v/c* ratio, or DS, is the ratio of demand flow rate over the capacity. From Webster algorithm [8], it is suggested that the least delay is obtained when green periods of phases are in proportional to the corresponding ratios of flow to saturation flows. The same to Webster algorithm, the balanced critical DS for all junctions in urban network will obtain the nearly lowest delay time.

The same to Equation (21-3) in [6], DS $\widetilde{o}_{j,i}(k)$ for lane group i of junction j can be rewritten as,

$$\widetilde{o}_{j,i}(k) = \frac{y_{j,i}(k)}{s_{j,i}\widetilde{u}_{j,i}(k)}, \forall j \in J, \forall i \in F_j \tag{9}$$

where $s_{j,i}$ is saturation flow for lane group i of junction j.

Critical DS $o_{j,p}(k)$ is the largest DS of lane groups belonging to the same phase. For a junction shown in Fig. 1(a), phase specification is predefined and fixed in general, there exists,

$$o_{j,1}(k) = \max\{\widetilde{o}_{j,2}(k), \widetilde{o}_{j,8}(k)\}$$
$$o_{j,2}(k) = \max\{\widetilde{o}_{j,3}(k), \widetilde{o}_{j,9}(k)\}$$
$$o_{j,3}(k) = \max\{\widetilde{o}_{j,5}(k), \widetilde{o}_{j,11}(k)\}$$
$$o_{j,4}(k) = \max\{\widetilde{o}_{j,6}(k), \widetilde{o}_{j,12}(k)\}$$

The above relationship is formulated as follows,

$$\mathbf{o}_j(k) = \mathcal{P}(\widetilde{\mathbf{o}}_j(k)) \tag{10}$$

where $\widetilde{\mathbf{o}}_j(k) \in \mathbf{\Gamma}_j(k)$; $\mathbf{o}_j(k) \in \mathbf{\Xi}_j(k)$; function $\mathcal{P}(\cdot)$ formulates phase specification.

As phase time also has the relationship similar to DS, there exists,

$$\mathbf{u}_j(k) = \mathcal{P}(\widetilde{\mathbf{u}}_j(k)) \tag{11}$$

where $\widetilde{\mathbf{u}}_j(k) \in \mathbf{\Gamma}_j(k)$; $\mathbf{u}_j(k) \in \mathbf{\Xi}_j(k)$.

Substituting (9) into (10), i.e.,

$$\mathbf{o}_j(k) = \mathcal{P}\left(\text{diag}\{\mathbf{s}_j\}^{-1}\text{diag}\{\widetilde{\mathbf{u}}_j(k)\}^{-1}\mathbf{y}_j(k)\right) = \text{diag}\{\mathbf{u}_j(k)\}^{-1}\mathcal{P}\left(\text{diag}\{\mathbf{s}_j\}^{-1}\mathbf{y}_j(k)\right) \tag{12}$$

where $\mathbf{s}_j = [s_{j,1}, s_{j,2}, \ldots, s_{j,N_{j,i}}]^\top$.

Then errors of critical DS $\mathbf{e}_j(k), \forall j \in J$ can be obtained as critical DS minus mean value of critical DS for junction j, i.e.,

$$\mathbf{e}_j(k) = \mathbf{o}_j(k) - \frac{1}{N_{j,p}}\mathcal{J}_{N_{j,p},N_{j,p}}\mathbf{o}_j(k) = A_j\mathbf{o}_j(k) \tag{13}$$

where $\mathbf{e}_j(k) \in \mathbf{\Xi}_j(k)$; $A_j = I_{N_{j,p}} - \dfrac{1}{N_{j,p}}\mathcal{J}_{N_{j,p},N_{j,p}}$; $I_{N_{j,p}}$ is the identity matrix with dimensions of $N_{j,p}$; $\mathcal{J}_{N_{j,p},N_{j,p}}$ is all-ones matrix with dimensions of $N_{j,p} \times N_{j,p}$.

Substituting (12) into (13), i.e.,

$$\mathbf{e}_j(k) = A_j \text{diag}\{\mathbf{u}_j(k)\}^{-1}\mathcal{P}\left(\text{diag}\{\mathbf{s}_j\}^{-1}\mathbf{y}_j(k)\right) \tag{14}$$

For overall network, (14) is rewritten as,

$$\mathbf{e}(k) = A\text{diag}\{\mathbf{u}(k)\}^{-1}\mathcal{P}\left(\text{diag}\{\mathbf{s}\}^{-1}\mathbf{y}(k)\right) \tag{15}$$

where $\mathbf{y}(k) \in \mathbf{\Gamma}(k)$; $\mathbf{e}(k), \mathbf{u}(k) \in \mathbf{\Xi}(k)$; $A = \text{diag}\{A_1, A_2, \ldots, A_N\}$; $\mathbf{s} = [s_1^\top, s_2^\top, \ldots, s_N^\top]^\top$.

The objective of IT strategy is to minimize all elements of $\mathbf{e}(k), \forall k \in [0, K]$ to be 0 over repetitions, i.e.,

$$\lim_{n\to\infty} \mathbf{e}(k, n) \longrightarrow 0, \quad \forall k \in [0, K] \tag{16}$$

where $\mathbf{e}(k, n) \in \mathbf{\Xi}(k, n)$ is error of critical DS $\mathbf{e}(k)$ on iteration n.

Traditional iterative learning approaches [16] are designed to track specific trajectories. As traffic demand varies largely at different time of a day for every junction, iterative approach applied to urban traffic control [19] uses traffic assignment model to predefine the specific trajectories. This predefined process is labour-intensive and the control performance will be determined by the accuracy of traffic assignment model. In contrast, error $\mathbf{e}(k, n)$ in IT strategy is calculated in (15) by demand $\mathbf{y}(k, n)$ and phase $\mathbf{u}(k, n)$ without desired trajectories.

With respect to IT strategy proposed in [20], IT strategy is extended to be suitable for under-saturated and saturated conditions with traffic information of traffic demand $\mathbf{y}(k)$ rather than vehicle throughput $\mathbf{z}(k)$. IT considers not only vehicle throughput, but also residual queued vehicles especially in saturated conditions. For certain phase in under-saturated and saturated conditions, degree of saturation will be $o_{j,p}(k) < 1$ and $o_{j,p}(k) \geq 1$, respectively. In this paper, the saturation flow is assumed to be constant in different traffic conditions.

For the single junction with repetitive traffic flows, desired phase durations can be calculated directly by assigning $\mathbf{e}_j(k) = 0$. For the network with multiple connected junctions, balancing critical DS of certain junction will affect other junctions in the network to be unbalanced. IT strategy considers the neighbouring junctions' coupling effects iteratively and approaches the objectives of all junctions during the entire period simultaneously.

III. Iterative Tuning Strategy and Convergence Analysis

Throughout this paper, $\|\cdot\|$ denotes Euclidean norm. In mathematics, for square matrix M, the spectral radius $\rho(M)$ is defined as follows,

$$\rho(M) = \max_{\pi}\{|r_\pi|\}$$

where r_π is the π^{th} eigenvalue of M.

Lemma 1. As indicated of Corollary 10.29, Chapter 10 in *Applied Linear Algebra* [21], if matrix M is symmetric, its Euclidean matrix norm is equal to its spectral radius, i.e.,

$$\|M\| = \rho(M) \tag{17}$$

To facilitate the tuning convergence, λ norm of a vector $\mathbf{e}(k)$ is defined, i.e.,

$$\|\mathbf{e}(k)\|_\lambda = \sup_{k\in[0,K]} a^{-\lambda k}\|\mathbf{e}(k)\| \tag{18}$$

where $\lambda > 0$ and $a > 1$.

A. Assumptions

Assumption 1. Referring to traffic model (7)-(8) and objective function (15), functions $f(\cdot,\cdot,\cdot)$, $g(\cdot,\cdot)$ and $\mathcal{P}(\cdot)$ are uniformly globally *Lipschitz* on a compact set with respect to their arguments for $k \in [0, K]$, i.e.,

$$\begin{aligned}\|f\left(\mathbf{x}(k,n+1),\mathbf{u}(k,n+1),\mathbf{d}(k,n+1)\right) - f(\mathbf{x}(k,n),\mathbf{u}(k,n),\mathbf{d}(k,n))\| \\ \leq k_{f_x}\|\mathbf{x}(k,n+1)-\mathbf{x}(k,n)\| + k_u\|\mathbf{u}(k,n+1)-\mathbf{u}(k,n)\| + k_{f_d}\|\mathbf{d}(k,n+1)-\mathbf{d}(k,n)\|\end{aligned} \tag{19}$$

$$\|g(\mathbf{x}(k,n+1),\mathbf{d}(k,n+1)) - g(\mathbf{x}(k,n),\mathbf{d}(k,n))\| \leq k_{g_x}\|\mathbf{x}(k,n+1)-\mathbf{x}(k,n)\| + k_{g_d}\|\mathbf{d}(k,n+1)-\mathbf{d}(k,n)\| \tag{20}$$

$$\|\mathcal{P}(\text{diag}\{\mathbf{s}\}^{-1}\mathbf{y}(k,n+1)) - \mathcal{P}(\text{diag}\{\mathbf{s}\}^{-1}\mathbf{y}(k,n))\| \leq k_y\|\mathbf{y}(k,n+1)-\mathbf{y}(k,n)\| \tag{21}$$

where $\mathbf{x}(k,n), \mathbf{y}(k,n), \mathbf{d}(k,n) \in \mathbf{\Gamma}(k,n)$, $\mathbf{u}(k,n) \in \mathbf{\Xi}(k,n)$; k_{f_x}, k_u, k_{f_d}, k_{g_x}, k_{g_d}, k_y are *Lipschitz* constants.

Assumption 2. The initial traffic states and error of initial traffic states are bounded by $\overline{b}_x$ and b_x throughout repeated iterations, respectively, i.e.,

$$\begin{aligned} \|\mathbf{x}(0,n)\| &\leq \overline{b}_x, \quad \forall n \\ \|\mathbf{x}(0,n+1)-\mathbf{x}(0,n)\| &\leq b_x, \quad \forall n \end{aligned} \tag{22}$$

where $\mathbf{x}(0,n+1)$ and $\mathbf{x}(0,n)$ are initial traffic states $\mathbf{x}(0)$ in iteration $n+1$ and n, respectively.

Assumption 3. The disturbances $\mathbf{d}(k)$ form traffic flows in the urban network. Disturbances and error of disturbances are bounded by $\overline{b}_d$ and b_d throughout repeated iterations, respectively, i.e.,

$$\begin{aligned} \sup_{k\in[0,K]} \|\mathbf{d}(k,n)\| &\leq \overline{b}_d, \quad \forall n \\ \sup_{k\in[0,K]} \|\mathbf{d}(k,n+1)-\mathbf{d}(k,n)\| &\leq b_d, \quad \forall n \end{aligned} \tag{23}$$

where $\mathbf{d}(k,n+1)$ and $\mathbf{d}(k,n)$ are disturbances $\mathbf{d}(k)$ in iteration $n+1$ and n, respectively.

Assumption 4. There exists control input $\mathbf{u}(k,n) \in [u_{min}, u_{max}]$ when $n \to \infty$ that can exactly drive errors of critical DS $\mathbf{e}(k,n)$ to be zero over the finite cycle $[0,K]$, where u_{min} and u_{max} are predefined minimum and maximum phase duration, respectively.

Assumption 1 requires traffic model to be globally *Lipschitz* continuous. For traffic dynamics in reality, number of vehicles $\mathbf{x}(k,n)$, demand $\mathbf{y}(k,n)$ and vehicle throughput $\mathbf{z}(k,n)$ of lane groups are always bounded by the road maximum volumes and lane groups' capabilities, which can ensure the satisfaction of the global *Lipschitz* condition.

Assumption 2 is reasonable as number of vehicles is bounded by total volumes of certain lane group. In reality, traffic states around 6 a.m. are very small, which can be regarded as initial traffic states. From data analysis of urban network [7], traffic states on every day are similar and approach 0.

Assumption 3 is reasonable since vehicles driving in and out of network are bounded by the capabilities of parking lots. As shown in [7], variations of disturbances is below 20% on working days.

Assumption 4 is acceptable since control assignment should be feasible.

B. *Iterative Tuning Controller*

For entire network, IT strategy is constructed as follows, i.e.,

$$\mathbf{u}(k,n+1) = \mathbf{u}(k,n) + \beta\mathbf{e}(k,n) \tag{24}$$

where β is a diagonal matrix, where diagonal entries are positive real number.

Define β_j as diagonal matrix for junction j, there exists,

$$\beta = \text{diag}\{\beta_1, \beta_2, \ldots, \beta_N\}$$

In urban traffic network, phase constraint should be satisfied as follows,

$$\sum \mathbf{u}_j(k,n) + t_L = C, \forall k \in [0,K], \forall n \tag{25}$$

where $\mathbf{u}_j(k,n)$ is phase duration $\mathbf{u}_j(k)$ in iteration n; t_L is total lost time in one cycle; C is the cycle time.

Lemma 2. For junction $j \in J$, assuming phase durations $\mathbf{u}_j(k,1)$ in iteration 1 satisfy the constraint in (25), i.e.,

$$\sum \mathbf{u}_j(k,1) + t_L = C, \forall k \in [0,K] \tag{26}$$

To satisfy the constraint in (25) over repetitions, all diagonal entries of β_j are set to be equal, i.e., $\beta_j = b_{\beta_j} I$, where b_{β_j} is a scalar, I is identity matrix.

Proof. See Appendix A. □

Lemma 3. For junction j, as defined in (13), there exists matrix $A_j = I_m - \frac{1}{m}\mathcal{J}_{m\times m}$, where m is integer, then A_j is idempotent matrix, i.e.,

$$A_j = A_j A_j \tag{27}$$

Let $\Psi_j = (I_m - \Phi_j A_j)A_j$, and $\Phi_j = \text{diag}\{\phi_1, \phi_2, \ldots, \phi_m\}$, where $\phi_1, \phi_2, \ldots, \phi_m$ are positive scalar. Define $b_{\Phi_j} = \|\Phi_j\|$, if $0 < b_{\Phi_j} < 2$, there exists,

$$\|\Psi_j\| < 1 \tag{28}$$

For entire network, as defined in (15), $A = \text{diag}\{A_1, A_2, \ldots, A_N\}$, then A is also idempotent matrix, i.e.,

$$A = AA \tag{29}$$

Let $\Psi = (I - \Phi A)A$, and $\Phi = \text{diag}\{\Phi_1, \Phi_2, \ldots, \Phi_N\}$. Define $b_\Phi = \|\Phi\|$, if $0 < b_\Phi < 2$, there exists,

$$\|\Psi\| < 1 \tag{30}$$

Proof. See Appendix B. □

Theorem 1. Consider traffic dynamics described by nonlinear model (7) (8) with objective function (15) and IT controller (24) under Assumptions 1-4, choosing an appropriate tuning matrix $\beta = b_\beta I$ such that $0 < b_\beta < \frac{2u_{min}}{b_o}$, the error of critical DS $\mathbf{e}(k,n)$ and phase difference $\delta\mathbf{u}(k,n)$ are bounded over repetitions, i.e.,

$$\lim_{n\to\infty} \|\mathbf{e}(k,n)\|_\lambda \leq c_x b_x + c_d b_d \tag{31}$$

$$\lim_{n\to\infty} \|\delta\mathbf{u}(k,n)\|_\lambda \leq b_\beta c_x b_x + b_\beta c_d b_d \tag{32}$$

where $\delta\mathbf{u}(k,n) = \mathbf{u}(k,n+1) - \mathbf{u}(k,n)$; c_x, c_d, b_β and b_o are defined in Appendix C.

Proof. See Appendix C. □

Remark 1. As stated in (31) and (32), the bounds of $\lim_{n\to\infty} \|\mathbf{e}(k,n)\|_\lambda$ and $\lim_{n\to\infty} \|\delta\mathbf{u}(k,n)\|_\lambda$ depend on the bounds of the initial conditions b_x, variations of disturbances b_d and tuning parameter β. Note that if b_x and b_d tend to 0, errors of critical DS tend to 0, and phase durations approach steady states.

Remark 2. For the entire network, every junction has its own controller. For local controller of junction j, only the local information, such as traffic demand is required. With consideration of traffic states $\mathbf{x}(k)$ in (7), these controllers work cooperatively toward balancing critical DS for all junctions during the entire period. Therefore, IT strategy is decentralized methodology and capable for large-scale network.

Remark 3. The tuning convergence only depends on minimum green time and maximum DS b_o, where b_o is bounded since the storage capacities of lane groups are fixed. The unavailable disturbances do not affect sufficient condition for guaranteeing convergence. Therefore IT strategy is suitable for complex urban traffic control when uncertainties of traffic dynamics exist.

Remark 4. Traffic model (7)-(8) is a generalized form, which can be simplified to be store-and-forward model [2] and BLX model [3]. Theorem 1 validates the convergence of IT controller under store-and-forward model and BLX model.

IV. Case Studies

In this section, Commonwealth Avenue in Singapore with heavy traffic demand is simulated in Vissim (Fig. 2).

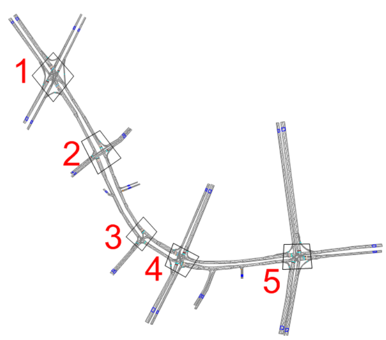


Fig. 2. Commonwealth Avenue in Singapore

A. Case Description

Referring to [22], Origin-Destination (OD) pairs are mainly designed to construct traffic flows close to the real situations on working days. In Vissim, traffic conditions are completely identical for all algorithms. Simulation period is 15 hours from $07:00\ am$ to $10:00\ pm$ with time interval $\tau = 15\ mins$, which includes most of heavy traffic period. Real traffic conditions are simulated with 94.62% accuracy rate.

Two cases are considered here,

1) Case I: OD matrices without any variation
2) Case II: matrices with certain variations

At first, a set of OD matrices $\mathcal{OD}$ are obtained and simulated in Case I, which shows the convergence of IT strategy without traffic variations. Then numbers of sets of OD matrices $\mathcal{OD}_v$ are calculated in (33) to generate traffic variations in Case II. The first term in (33) is designed to generate traffic variations and the second term aims to change traffic flow patterns.

$$\mathcal{OD}_v(\tau) = \varphi_1 \mathcal{R}_1 \cdot \mathcal{OD}(\tau) + \varphi_2 \mathcal{R}_2 \cdot \mathcal{OD}(\tau) + \mathcal{OD}(\tau), \forall \tau \tag{33}$$

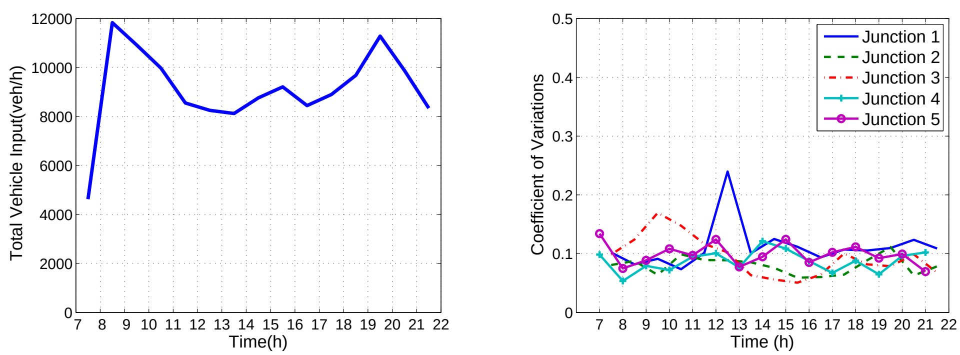
where φ_1 and φ_2 are scalar amplifier factors; $\mathcal{R}_1$ is matrix of random numbers with proper dimensions, and each element in $\mathcal{R}_1 \in [-1, 1]$; for $\mathcal{R}_2$, its $(s,t)^{th}$ entry is calculated as $(-1)^{(s+t)}$.

In case II, 100 sets of different OD matrices are generated as (33), which are categorized into three parts: part i (1-10 runs): $\varphi_1 = 0.15$; $\varphi_2 = 0$. This part is analyzed to calculate conventional pre-timed signal schedules and fine-tuning process for IT strategy. Part ii (11-50 runs): $\varphi_1 = 0.15$; $\varphi_2 = 0$. This part aims to compare the network performance of different strategies with bounded traffic variations. Part iii (51-100 runs): $\varphi_1 = 0.15$; $\varphi_2 = 0.2$. The part aims to show the feasibility of IT strategy with respect to changed traffic flow patterns.

Total vehicle inputs of the network are shown in Fig. 3(a). At the beginning vehicle input is quite small, which initializes traffic dynamics in the network. The overall simulations include morning peak period, evening peak period and afternoon non-peak period. In Case II, there are 100 scenarios with different time-sliced OD matrices designed based on (33). Under single pre-timed strategy, coefficient of variations in part i for five junctions are shown in Fig. 3(b). The average value is 9.45%, which is quite similar to the real conditions during heavy traffic period of working days.

Regarding to the assumptions made in Section III.A, simulation platform has predefined road volumes and junctions' capabilities based on the reality for Assumption 1. The bound in Assumption 2 and 3 are determined by sets of OD matrices generated in (33). For Assumption 4, there exists phase $u_{j,p}(k, n+1)$ beyond the constraints in certain cycle k, which indicates that there is no feasible phase duration to balance critical DS. As function $sat[\cdot]$ in (34), let $u_{j,p}(k, n+1) = u_{min}$ or u_{max}, phases of junction j are tuned to balance critical DS without this phase p.

$$sat[u_{j,p}(k,n+1)] = \begin{cases} u_{j,p}(k,n+1), & \text{if } u_{min} < u_{j,p}(k,n+1) < u_{max}; \\ u_{max}, & \text{if } u_{j,p}(k,n+1) \geq u_{max}; \\ u_{min}, & \text{if } u_{j,p}(k,n+1) \leq u_{min}. \end{cases} \tag{34}$$



(a) Vehicle Inputs (b) Coefficient of Variations

Fig. 3. Traffic Conditions of Commonwealth Avenue in Singapore

B. Signal Control Strategies

Several signal control strategies are simulated for comparisons. Pre-timed signals are obtained from Webster algorithm [8].

Single Pre-timed: one pre-timed signal plan for each junction during the entire period.

TOD Pre-timed: three pre-timed signal plans for morning peak, afternoon non-peak and evening peak hours, respectively.

IT with Init: IT strategy with initialized signal schedules, which is obtained from single pre-timed signal plan.

IT without Init: IT strategy with arbitrarily initialized signal schedules. For junction j, the initial green time are allocated equally, i.e., $u_{j,p}(k, 1) = (C - t_L)/N_{j,p}, \forall k \in [0, K]$.

IT of VT: IT strategy with traffic information of Vehicle Throughput (VT) as stated in [20].

In simulation, rounding operation to nearest integer is taken for phase durations as simulation software can only identify integer. IT controller in (24) is used as follows, i.e.,

$$\mathbf{u}(k, n+1) = \Re\left(\mathbf{u}(k, n) + \beta \mathbf{e}(k, n)\right) \tag{35}$$

where $\Re$ is the function of rounding to nearest integer. In certain iteration, if cycle time accumulated by $\mathbf{u}_j(k)$ is not equal to cycle time C because of rounding operation, the largest phase $u_{j,p}(k), p \in P_j$ is adjusted to be equal to C.

C. Criteria of Comparison

Total Number of Vehicle (TNV): total number of vehicles which have left the network.

Average Delay Time (ADT): delay time per vehicle, where delay time is the subtraction between total travel time of all vehicles and total free-flow travel time.

Average Number of Stops (ANS): stop per vehicle, where stop is counted if the speed of the vehicles was greater than zero at previous time step and is zero currently.

Reductions of I_{ADT} and I_{ANS} are calculated as follows,

$$I_{ADT} = \frac{ADT_{SP} - ADT}{ADT_{SP}} \times 100\%; \quad I_{ANS} = \frac{ANS_{SP} - ANS}{ANS_{SP}} \times 100\%; \tag{36}$$

where ADT_{SP}, ANS_{SP} are ADT and ANS under single pre-timed strategy.

V. Simulation Results

In this section, simulation results in two cases are summarized. Average Phase Difference $\theta(n)$ accumulated indicates the convergence of phase time for iteration n, i.e.,

$$\theta(n) = \frac{1}{K(\sum_{j\in J} N_{j,p})} \sum_{k\in[0,K]} \sum_{j\in J} \sum_{p\in P_j} |u_{j,p}(k, n+1) - u_{j,p}(k, n)| \tag{37}$$

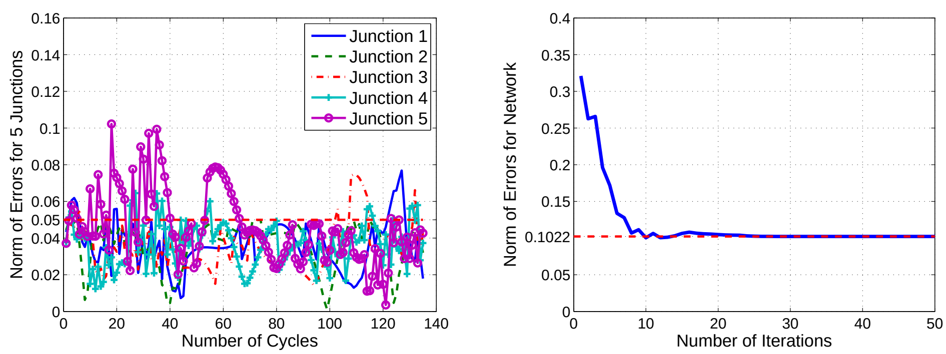
where $u_{j,p}(k, n+1)$ and $u_{j,p}(k, n)$ are the phase $u_{j,p}(k)$ in iteration $n+1$ and n, respectively.

A. Case I: Without Variations

In Case I, one set of OD matrices are simulated over repetitions. In the first iteration, phase durations under single pre-timed strategy obtained from Wester algorithm are applied. Then 50 iterations are run to tune phase splits by IT strategy.

Errors of critical DS shown in Fig. 4 indicates the convergence of tuning process. In the last iteration, norm of errors $\|\text{diag}\{\mathbf{e}_j(k,n)\}\|, \forall k$ for 5 junctions are illustrated respectively in Fig. 4(a). According to (35), and $b_{\beta_j} = 10, \forall j \in J$, phase durations will not be tuned once $\|\text{diag}\{\mathbf{e}_j(k,n)\}\| < 0.05$, since $b_{\beta_j}\|\text{diag}\{\mathbf{e}_j(k,n)\}\| < 0.5$ and rounding operation takes it as 0. From Fig. 4(a), most of errors are below 0.05. There are several points exceeding 0.05 because of rounding operation to nearest integer.

Norm of errors $\|\text{diag}\{\mathbf{E}(n)\}\|$ for the network is illustrated in Fig. 4(b), where $\mathbf{E}(n) = [\mathbf{e}^\top(1,n), \mathbf{e}^\top(2,n), \ldots, \mathbf{e}^\top(K,n)]^\top$. It represents the maximum errors of 5 junctions during the entire period along with number of iterations. The errors $\|\text{diag}\{\mathbf{E}(n)\}\|$ converge very fast within first ten iterations and approaches steady states of 0.1022 after 29^{th} iteration.



(a) Five Junctions (b) Overall Network

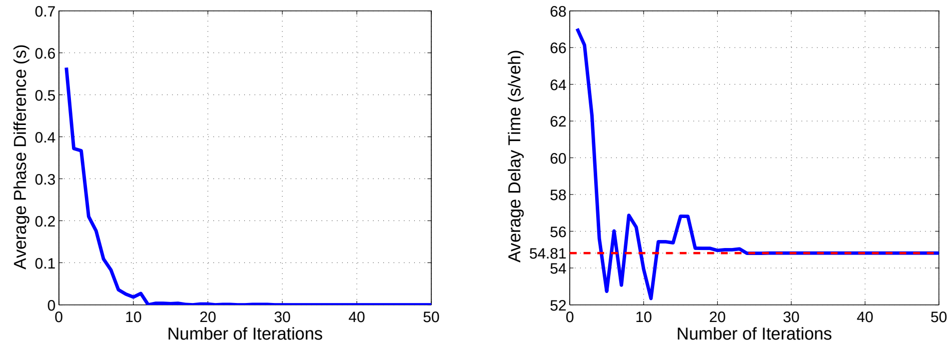
Fig. 4. Norm of Errors of Critical DS for Overall Network and Five Junctions

Meanwhile, average phase difference is calculated to indicate the convergence of phase durations $\mathbf{u}(k,n), \forall k$. As shown in Fig. 5(a), average phase difference is 0 and phase durations reach steady states after 29^{th} iteration. The simulations without variations validate the convergence of IT strategy proved in Theorem 1 and stated in Remark 1.

Moreover, along with phase durations tuned iteratively, average delay time for overall network is deceased and approaches steady state of 54.81 s/veh (Fig. 5(b)). In the first iteration, average delay time under single pre-timed strategy is 67.03 s/veh. IT outperforms pre-timed strategy by 18.23%. As stated by control objective in Section II, balancing degree of saturation can have nearly lowest delay time, but not the best performance. The best performance shown in Fig. 5(b) is 52.54 s/veh, which is quite close to steady state of 54.81 s/veh.

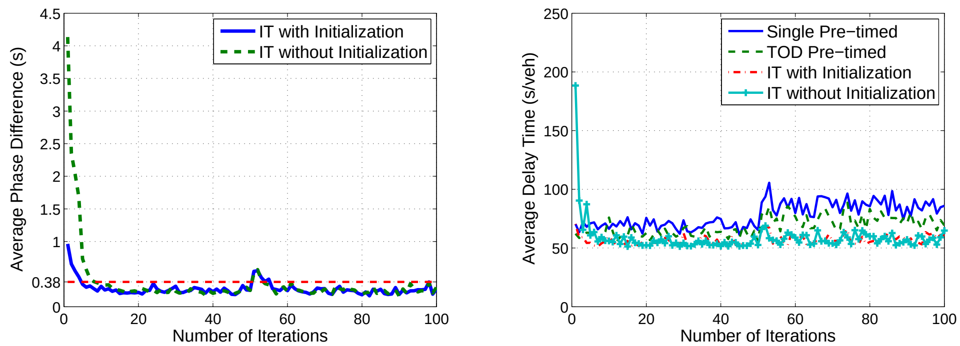
B. Case II: With Variations

As shown in Fig. 6(a), average phase difference is approaching steady state of 0.24 and below 0.38 second per phase, rather than 0 in Fig. 5(a). Without phase initialization, signal schedules are tuned substantially at the first few iterations. During part ii (11-50 runs), average phase difference are quite similar with and without initializations. Since traffic flow patterns are changed



(a) Average Phase Difference (b) Average Delay Time

Fig. 5. Average Phase Difference (s/phase) and Average Delay Time (s/veh) without Variations



(a) Average Phase Difference (b) Average Delay Time

Fig. 6. Average Phase Difference (s/phase) and Average Delay Time (s/veh) with Variations

from 51^{th} iteration, phase difference increases suddenly and return to normal states after 4 iterations. The simulations validate Theorem 1 that total phase difference are bounded by 0.38 second per phase with respect to bounded traffic variations.

The average delay time for four signal schedules are calculated and depicted in Fig. 6(b). Roughly speaking, during part ii (11-50 runs) and part iii (51-100 runs), IT strategy is better than single pre-timed and TOD pre-timed signal schedules. IT strategy with and without initialization looks similar, which indicates labour-intensive preparation of signal schedules for a new junction is not required in IT strategy. During part iii (51-100 runs), performance of single pre-timed and TOD pre-timed signal schedules gets worse due to the changed traffic flow patterns. In contrast, IT strategy learns traffic flow patterns iteratively and automatically. Delay time returns to normal states after 4 iterations. The performance in part iii (51-100 runs) shows that IT strategy adapts to gradual changing traffic flow patterns.

The average performance during 11-50 runs are summarized in Table I. Reduction of ADT and ANS for all strategies presented in bracket are calculated by (36). At the end of simulation period, traffic demand is quite small and there is no congestion everywhere, then TNV is very similar. Average delay time during 11-50 runs is 54.4790 s/veh. Compared with delay time of 54.81 s/veh in Case I, variations of traffic demand may not affect the network performance when traffic demand is highly repetitive as working days. It indicates the feasibility of IT strategy on working days without unforeseen traffic incidents.

TABLE I
AVERAGE PERFORMANCE

Algorithm	TNV (veh)	ADT (s/veh)	ANS
Single Pre-timed	32904	69.1373	1.5694
TOD Pre-timed	32902	62.2507(9.96%)	1.4215(9.42%)
IT with Init	32909	54.4790(21.20%)	1.2494(20.39%)
IT without Init	32908	54.2354(21.55%)	1.2442(20.72%)
IT of VT	32908	57.4241(16.94%)	1.3261(15.50%)

TOD Pre-timed, IT strategy with and without initialization can reduce delay time by 9.96%, 21.20% and 21.55%, reduce number of stops by 9.42%, 20.39% and 20.72%, respectively. Therefore, IT strategy considerably improves the network performance on working days. Besides, with consideration of residual queued vehicles, the performance of IT with traffic demand outperforms vehicle throughput stated in [20] by 4.36% and 4.89% on ADT and ANS, respectively. It shows that traffic demand is more reliable to reflect real traffic conditions.

VI. Conclusion

Iterative Tuning (IT) strategy is the first trial to apply Iterative Learning Control (ILC) approach to urban traffic signal control without specific control reference. This paper extends IT strategy for phase splits to be capable for under-saturated and saturated traffic conditions on working days by taking traffic demand into consideration. IT controllers of all junctions tune daily traffic signal schedules iteratively and automatically to balance critical degree of saturation (DS) for all phases. Time-consuming and labour-intensive preparation and fine-tuning process of pre-timed signal schedules are not required any more. With generalized traffic model and explicit tuning algorithm, sufficient condition for guaranteeing the convergence of IT strategy is investigated in details by analysing the robustness to traffic variations and errors of initial conditions. Simulation results validate the convergence of IT strategy with and without variations. Comparing with pre-timed strategy, IT strategy considerably reduces delay time and number of stops for overall network. With changed vehicle inputs, IT strategy learns traffic patterns and fine-tunes phase durations within a few iterations.

The future work will focus on IT strategy with reactive compensation, which can respond to non-repetitive disturbances.

Appendix A
Proof Of Lemma 2

Proof. In this paper, lost time t_L and cycle time C are assumed to be fixed. Define $\mathbf{o}_j(k,n)$, $\mathbf{e}_j(k,n)$ as $\mathbf{o}_j(k)$, $\mathbf{e}_j(k)$ in iteration n, with IT controller (24), there exists,

$$\mathbf{u}_j(k,n+1) = \mathbf{u}_j(k,n) + \beta_j \mathbf{e}_j(k,n) \tag{A.1}$$

Summing up all elements of all vectors on both sides separately,

$$\sum \mathbf{u}_j(k,n+1) = \sum \mathbf{u}_j(k,n) + \sum \beta_j \mathbf{e}_j(k,n) \tag{A.2}$$

Vector $\mathbf{e}_j(k,n) = A_j \mathbf{o}_j(k,n)$ is defined in (13). With matrix A_j, $\sum \mathbf{e}_j(k,n) = 0$ for any vector $\mathbf{o}_j(k,n)$.
If all entries of β_j are equal, i.e., $\beta_j = b_{\beta_j} I$, for any vector $\mathbf{o}_j(k,n)$, there exists,

$$\sum \beta_j \mathbf{e}_j(k,n) = 0 \tag{A.3}$$

Substitute (A.3) into (A.2), yields,

$$\sum \mathbf{u}_j(k,n+1) = \sum \mathbf{u}_j(k,n) \tag{A.4}$$

Assuming phase durations $\mathbf{u}_j(k,1)$ satisfy,

$$\sum \mathbf{u}_j(k,1) + t_L = C, \forall k \in [0,K] \tag{A.5}$$

With (A.4), there exists,

$$\sum \mathbf{u}_j(k,n) + t_L = C, \forall k \in [0,K], \forall n \tag{A.6}$$

Lemma 2 is proved. □

Appendix B
Proof of Lemma 3

Proof. For $A_j = I_m - \frac{1}{m}\mathcal{J}_{m\times m}$,

$$\begin{aligned} A_j A_j &= (I_m - \frac{1}{m}\mathcal{J}_{m\times m})(I_m - \frac{1}{m}\mathcal{J}_{m\times m}) = I_m I_m - 2I_m \frac{1}{m}\mathcal{J}_{m\times m} + \frac{1}{m}\mathcal{J}_{m\times m}\frac{1}{m}\mathcal{J}_{m\times m} \\ &= I_m - 2\frac{1}{m}\mathcal{J}_{m\times m} + \frac{1}{m}\mathcal{J}_{m\times m} = I_m - \frac{1}{m}\mathcal{J}_{m\times m} = A_j \end{aligned} \tag{B.1}$$

Therefore, A_j is idempotent matrix.
From [23], for idempotent matrix A_j, there exists,

$$rank(A_j) = tr(A_j) = m-1 \tag{B.2}$$

where tr is trace of matrix.

Let $r_\pi \in \mathbf{r}$ be eigenvalue, and $\mathbf{r}$ is the vector of eigenvalues of A_j. Then $\mathbf{r} = [0,1,1,\ldots,1]$, where there are only 1 eigenvalue equal to 0 and $(m-1)$ eigenvalues equal to 1.

Let $\Psi_j = (I_m - \Phi_j A_j)A_j$, where Φ_j is diagonal matrix, and $\Phi_j = \text{diag}\{\phi_1, \phi_2, \ldots, \phi_m\}$, where $\phi_1, \phi_2, \ldots, \phi_m$ are positive scalar. Based on Eigendecomposition, there exists eigenvector Q such that $A_j = Q\Lambda Q^{-1}$, where Q^{-1} is the inverse of Q; $\Lambda = \text{diag}\{\mathbf{r}\}$. Then

$$\Psi_j = (I_m - \Phi_j A_j)A_j = (QI_mQ^{-1} - \Phi_j Q\Lambda Q^{-1})Q\Lambda Q^{-1} = Q(I_m - \Phi_j\Lambda)Q^{-1}Q\Lambda Q^{-1} = Q(I_m - \Phi_j\Lambda)\Lambda Q^{-1} \tag{B.3}$$

where

$$\begin{aligned}(I_m - \Phi_j\Lambda)\Lambda &= \left(\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} - \begin{bmatrix} \phi_1 & 0 & \cdots & 0 \\ 0 & \phi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \phi_m \end{bmatrix}\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}\right) \cdot \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1-\phi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1-\phi_m \end{bmatrix}\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 1-\phi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1-\phi_m \end{bmatrix}\end{aligned} \tag{B.4}$$

The spectral radius of Ψ_j is formulated as,

$$\rho(\Psi_j) = \max\{1-\phi_1, 1-\phi_2, \ldots, 1-\phi_m\} \tag{B.5}$$

Since A_j is symmetric matrix, matrix Ψ_j is also symmetric. Based on Lemma 1, yields,

$$\|\Psi_j\| = \rho(\Psi_j) \tag{B.6}$$

In order to have $\|\Psi_j\| < 1$, then

$$0 < b_{\Phi_j} < 2 \tag{B.7}$$

where $b_{\Phi_j} = \|\Phi_j\|$.

For entire network, define $A = \text{diag}\{A_1, A_2, \ldots, A_N\}$,

$$AA = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_N \end{bmatrix}\begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_N \end{bmatrix} = \begin{bmatrix} A_1A_1 & 0 & \cdots & 0 \\ 0 & A_2A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_NA_N \end{bmatrix} = A \tag{B.8}$$

Therefore, A is also idempotent matrix.

Let $\Psi = (1 - \Phi A)A$, and $\Phi = \text{diag}\{\Phi_1, \Phi_2, \ldots, \Phi_N\}$, every block diagonal matrix in Ψ follows the similar relationship as shown in (B.3)-(B.5), therefore, the spectral radius of Ψ is formulated as,

$$\rho(\Psi) = \max_{j\in J}\{\rho(\Psi_j)\} \tag{B.9}$$

Based on Lemma 1,

$$\|\Psi\| = \rho(\Psi) \tag{B.10}$$

In order to have $\|\Psi\| < 1$, then

$$0 < b_\Phi < 2 \tag{B.11}$$

where $b_\Phi = \|\Phi\|$.

Lemma 3 is proved. □

Appendix C
Proof Of Theorem 1

Proof. Iterative tuning controller is

$$\mathbf{u}(k, n+1) = \mathbf{u}(k, n) + \beta\mathbf{e}(k, n) \tag{C.1}$$

where $\mathbf{e}(k,n) = A\text{diag}\{\mathbf{u}(k,n)\}^{-1}\mathcal{P}\left(\text{diag}\{\mathbf{s}\}^{-1}\mathbf{y}(k,n)\right)$ based on (15).

Define

$$\begin{aligned}\delta\mathbf{u}(k,n) &= \mathbf{u}(k,n+1) - \mathbf{u}(k,n) \\ \delta\mathbf{x}(k,n) &= \mathbf{x}(k,n+1) - \mathbf{x}(k,n) \\ \delta\mathbf{y}(k,n) &= \mathbf{y}(k,n+1) - \mathbf{y}(k,n) \\ \delta\mathbf{d}(k,n) &= \mathbf{d}(k,n+1) - \mathbf{d}(k,n)\end{aligned}$$

With $\text{diag}\{\mathbf{o}(k,n+1)\} = \text{diag}\{\mathbf{u}(k,n+1)\}^{-1}\mathcal{P}\left(\text{diag}\{\mathbf{s}\}^{-1}\mathbf{y}(k,n+1)\right)$ based on (12), yields,

$$\begin{aligned}
&\mathbf{e}(k,n+1)-\mathbf{e}(k,n) \\
=&A\text{diag}\{\mathbf{u}(k,n+1)\}^{-1}\mathcal{P}\left(\text{diag}\{\mathbf{s}\}^{-1}\mathbf{y}(k,n+1)\right)-A\text{diag}\{\mathbf{u}(k,n)\}^{-1}\mathcal{P}\left(\text{diag}\{\mathbf{s}\}^{-1}\mathbf{y}(k,n)\right) \\
=&A\text{diag}\{\mathbf{u}(k,n+1)\}^{-1}\mathcal{P}\left(\text{diag}\{\mathbf{s}\}^{-1}\mathbf{y}(k,n+1)\right)-A\text{diag}\{\mathbf{u}(k,n)\}^{-1}\mathcal{P}\left(\text{diag}\{\mathbf{s}\}^{-1}\mathbf{y}(k,n+1)\right) \\
&+A\text{diag}\{\mathbf{u}(k,n)\}^{-1}\mathcal{P}\left(\text{diag}\{\mathbf{s}\}^{-1}\mathbf{y}(k,n+1)\right)-A\text{diag}\{\mathbf{u}(k,n)\}^{-1}\mathcal{P}\left(\text{diag}\{\mathbf{s}\}^{-1}\mathbf{y}(k,n)\right) \\
=&-A\text{diag}\{\mathbf{u}(k,n+1)\cdot\mathbf{u}(k,n)\}^{-1}\delta\mathbf{u}(k,n)\mathcal{P}\left(\text{diag}\{\mathbf{s}\}^{-1}\mathbf{y}(k,n+1)\right)+A\text{diag}\{\mathbf{u}(k,n)\}^{-1}\mathcal{P}\left(\text{diag}\{\mathbf{s}\}^{-1}\delta\mathbf{y}(k,n)\right) \\
=&-A\text{diag}\{\mathbf{u}(k,n)\}^{-1}\text{diag}\{\mathbf{o}(k,n+1)\}\delta\mathbf{u}(k,n)+A\text{diag}\{\mathbf{u}(k,n)\}^{-1}\mathcal{P}\left(\text{diag}\{\mathbf{s}\}^{-1}\delta\mathbf{y}(k,n)\right)
\end{aligned} \tag{C.2}$$

Based on (C.1), there exists,

$$\delta\mathbf{u}(k,n)=\beta\mathbf{e}(k,n) \tag{C.3}$$

Substituting (C.3) into (C.2), then

$$\begin{aligned}
\mathbf{e}(k,n+1)&=\mathbf{e}(k,n)-A\text{diag}\{\mathbf{u}(k,n)\}^{-1}\text{diag}\{\mathbf{o}(k,n+1)\}\beta\mathbf{e}(k,n)+A\text{diag}\{\mathbf{u}(k,n)\}^{-1}\mathcal{P}\left(\text{diag}\{\mathbf{s}\}^{-1}\delta\mathbf{y}(k,n)\right) \\
&=(I-\beta A\text{diag}\{\mathbf{u}(k,n)\}^{-1}\text{diag}\{\mathbf{o}(k,n+1)\})\mathbf{e}(k,n)+A\text{diag}\{\mathbf{u}(k,n)\}^{-1}\mathcal{P}\left(\text{diag}\{\mathbf{s}\}^{-1}\delta\mathbf{y}(k,n)\right)
\end{aligned} \tag{C.4}$$

Based on Lemma 3, A is idempotent matrix, therefore,

$$\mathbf{e}(k,n)=A\text{diag}\{\mathbf{u}(k,n)\}^{-1}\mathcal{P}\left(\text{diag}\{\mathbf{s}\}^{-1}\mathbf{y}(k,n)\right)=A^2\text{diag}\{\mathbf{u}(k,n)\}^{-1}\mathcal{P}\left(\text{diag}\{\mathbf{s}\}^{-1}\mathbf{y}(k,n)\right)=A\mathbf{e}(k,n) \tag{C.5}$$

Then

$$\begin{aligned}
\mathbf{e}(k,n+1)&=(I-\beta\text{diag}\{\mathbf{u}(k,n)\}^{-1}\text{diag}\{\mathbf{o}(k,n+1)\}A)A\mathbf{e}(k,n)+A\text{diag}\{\mathbf{u}(k,n)\}^{-1}\mathcal{P}\left(\text{diag}\{\mathbf{s}\}^{-1}\delta\mathbf{y}(k,n)\right) \\
&=\Psi\mathbf{e}(k,n)+A\text{diag}\{\mathbf{u}(k,n)\}^{-1}\mathcal{P}\left(\text{diag}\{\mathbf{s}\}^{-1}\delta\mathbf{y}(k,n)\right)
\end{aligned} \tag{C.6}$$

where $\Psi=(I-\beta\text{diag}\{\mathbf{u}(k,n)\}^{-1}\text{diag}\{\mathbf{o}(k,n+1)\}A)A$.

Based on Assumption 1, taking norm operation for (C.6), there exists,

$$\begin{aligned}
\|\mathbf{e}(k,n+1)\|&\le\|\Psi\|\cdot\|\mathbf{e}(k,n)\|+\|A\text{diag}\{\mathbf{u}(k,n)\}^{-1}\mathcal{P}\left(\text{diag}\{\mathbf{s}\}^{-1}\delta\mathbf{y}(k,n)\right)\| \\
&\le\|\Psi\|\cdot\|\mathbf{e}(k,n)\|+k_y\|A\text{diag}\{\mathbf{u}(k,n)\}^{-1}\|\cdot\|\delta\mathbf{y}(k,n)\| \\
&\le\|\Psi\|\cdot\|\mathbf{e}(k,n)\|+\varepsilon_1\|\delta\mathbf{y}(k,n)\|
\end{aligned} \tag{C.7}$$

where $\varepsilon_1=\sup_{k\in[0,K]}k_y\|A\text{diag}\{\mathbf{u}(k,n)\}^{-1}\|$.

Taking norm operation on (8) based on Assumption 1 and 3, yields,

$$\|\delta\mathbf{y}(k,n)\|\le\|\delta g(\mathbf{x}(k,n),\mathbf{d}(k,n))\|\le k_{g_x}\|\delta\mathbf{x}(k,n)\|+k_{g_d}\|\delta\mathbf{d}(k,n)\|\le k_{g_x}\|\delta\mathbf{x}(k,n)\|+k_{g_d}b_d \tag{C.8}$$

Taking norm operation of (7) based on Assumption 1-3,

$$\begin{aligned}
\|\delta\mathbf{x}(k,n)\|&\le\|\delta\mathbf{x}(k-1,n)+B(k-1)\delta f(\mathbf{x}(k-1,n),\mathbf{u}(k-1,n),\mathbf{d}(k-1,n))+C(k-1)\delta\mathbf{d}(k-1,n)\| \\
&\le\|\delta\mathbf{x}(k-1,n)\|+b_B(k_{f_x}\|\delta\mathbf{x}(k-1,n)\|+k_{f_u}\|\delta\mathbf{u}(k-1,n)\|+k_{f_d}\|\delta\mathbf{d}(k-1,n)\|)+b_C\|\delta\mathbf{d}(k-1,n)\| \\
&\le(1+b_Bk_{f_x})\|\delta\mathbf{x}(k-1,n)\|+b_Bk_{f_u}\|\delta\mathbf{u}(k-1,n)\|+(b_C+b_Bk_{f_d})\|\delta\mathbf{d}(k-1,n)\| \\
&\le(1+b_Bk_{f_x})\|\delta\mathbf{x}(k-1,n)\|+b_Bk_{f_u}b_\beta\|\mathbf{e}(k-1,n)\|+(b_C+b_Bk_{f_d})\|\delta\mathbf{d}(k-1,n)\| \\
&\le\varepsilon_2\|\delta\mathbf{x}(k-1,n)\|+\varepsilon_3\|\mathbf{e}(k-1,n)\|+\varepsilon_4\|\delta\mathbf{d}(k-1,n)\| \\
&\le\varepsilon_2^k\|\delta\mathbf{x}(0,n)\|+\varepsilon_3\sum_{\kappa=0}^{k-1}\varepsilon_2^{k-\kappa-1}\|\mathbf{e}(\kappa,n)\|+\varepsilon_4\sum_{\kappa=0}^{k-1}\varepsilon_2^{k-\kappa-1}\|\delta\mathbf{d}(\kappa,n)\| \\
&\le\varepsilon_2^Kb_x+\varepsilon_4\sum_{j=0}^{k-1}\varepsilon_2^{k-\kappa-1}b_d+\varepsilon_3\sum_{\kappa=0}^{k-1}\varepsilon_2^{k-\kappa-1}\|\mathbf{e}(j,n)\| \\
&\le\varepsilon_2^Kb_x+\frac{\varepsilon_4(1-\varepsilon_2^K)}{1-\varepsilon_2}b_d+\varepsilon_3\sum_{\kappa=0}^{k-1}\varepsilon_2^{k-\kappa-1}\|\mathbf{e}(j,n)\|
\end{aligned} \tag{C.9}$$

where $b_C=\sup_{k\in[0,K]}\|C(k)\|$, $b_B=\sup_{k\in[0,K]}\|B(k)\|$, $b_\beta=\|\beta\|$, $\varepsilon_2=1+b_Bk_{f_x}$, $\varepsilon_3=b_Bk_{f_u}b_\beta$, $\varepsilon_4=b_C+b_Bk_{f_d}$.

Substituting (C.8) and (C.9) into (C.7),

$$\begin{aligned}\|\mathbf{e}(k,n+1)\| &\leq \|\Psi\|\cdot\|\mathbf{e}(k,n)\| + \varepsilon_1(k_{g_x}\|\delta\mathbf{x}(k,n)\| + k_{g_d}b_d) \\ &\leq \|\Psi\|\cdot\|\mathbf{e}(k,n)\| + \varepsilon_1 k_{g_x}\left(\varepsilon_3\sum_{\kappa=0}^{k-1}\varepsilon_2^{k-\kappa-1}\|\mathbf{e}(\kappa,n)\| + \varepsilon_2^K b_x + \frac{\varepsilon_4(1-\varepsilon_2^K)}{1-\varepsilon_2}b_d\right) + \varepsilon_1 k_{g_d}b_d \\ &= \|\Psi\|\cdot\|\mathbf{e}(k,n)\| + \varepsilon_1\varepsilon_3 k_{g_x}\sum_{\kappa=0}^{k-1}\varepsilon_2^{k-\kappa-1}\|\mathbf{e}(\kappa,n)\| + \varepsilon_1 k_{g_x}\varepsilon_2^K b_x + \left(\frac{\varepsilon_1 k_{g_x}\varepsilon_4(1-\varepsilon_2^K)}{1-\varepsilon_2} + \varepsilon_1 k_{g_d}\right)b_d \\ &= \|\Psi\|\cdot\|\mathbf{e}(k,n)\| + \varepsilon_1\varepsilon_3 k_{g_x}\sum_{\kappa=0}^{k-1}\varepsilon_2^{k-\kappa-1}\|\mathbf{e}(\kappa,n)\| + \varepsilon_5 b_x + \varepsilon_6 b_d\end{aligned} \tag{C.10}$$

where $\varepsilon_5 = \varepsilon_1 k_{g_x}\varepsilon_2^K$, $\varepsilon_6 = \left(\frac{\varepsilon_1 k_{g_x}\varepsilon_4(1-\varepsilon_2^K)}{1-\varepsilon_2} + \varepsilon_1 k_{g_d}\right)$.

Then the convergence of $\mathbf{e}(k,n)$ is proved by using λ norm operation. From (C.10), multiplying both sides by $\varepsilon_2^{-\lambda k}$, and taking supremum over [0,K], there exists,

$$\begin{aligned}&\sup_{k\in[0,K]}\varepsilon_2^{-\lambda k}\|\mathbf{e}(k,n+1)\| \\ &\leq \|\Psi\|\cdot\sup_{k\in[0,K]}\varepsilon_2^{-\lambda k}\|\mathbf{e}(k,n)\| + \varepsilon_1\varepsilon_3 k_{g_x}\sup_{k\in[0,K]}\varepsilon_2^{-\lambda k}\sum_{\kappa=0}^{k-1}\varepsilon_2^{k-\kappa-1}\|\mathbf{e}(\kappa,n)\| + \sup_{k\in[0,K]}\varepsilon_2^{-\lambda k}(\varepsilon_5 b_x + \varepsilon_6 b_d)\end{aligned} \tag{C.11}$$

Since

$$\begin{aligned}\sup_{k\in[0,K]}\varepsilon_2^{-\lambda k}\sum_{\kappa=0}^{k-1}\varepsilon_2^{k-\kappa-1}\|\mathbf{e}(\kappa,n)\| &\leq \varepsilon_2^{-1}\sup_{k\in[0,K]}\left(\sum_{j=0}^{k-1}\varepsilon_2^{-\lambda\kappa}\|\mathbf{e}(\kappa,n)\|\varepsilon_2^{(\lambda-1)(j-k)}\right) \\ &\leq \varepsilon_2^{-1}\sup_{k\in[0,K]}\left(\sum_{\kappa=0}^{k-1}(\sup_{k\in[0,K]}\varepsilon_2^{-\lambda\kappa}\|\mathbf{e}(\kappa,n)\|)\varepsilon_2^{(\lambda-1)(\kappa-k)}\right) \\ &\leq \varepsilon_2^{-1}\|\mathbf{e}(k,n)\|_\lambda\cdot\sup_{k\in[0,K]}\sum_{\kappa=0}^{k-1}\varepsilon_2^{(\lambda-1)(\kappa-k)} \\ &= \|\mathbf{e}(k,n)\|_\lambda\cdot\frac{1-\varepsilon_2^{-(\lambda-1)K}}{\varepsilon_2^\lambda-\varepsilon_2}\end{aligned} \tag{C.12}$$

Deriving from (C.11),

$$\begin{aligned}\|\mathbf{e}(k,n+1)\|_\lambda &\leq \|\Psi\|\cdot\|\mathbf{e}(k,n)\|_\lambda + \varepsilon_1\varepsilon_3 k_{g_x}\frac{1-\varepsilon_2^{-(\lambda-1)K}}{\varepsilon_2^\lambda-\varepsilon_2}\cdot\|\mathbf{e}(k,n)\|_\lambda + \varepsilon_7 b_x + \varepsilon_8 b_d \\ &\leq \left(\|\Psi\| + \varepsilon_1\varepsilon_3 k_{g_x}\frac{1-\varepsilon_2^{-(\lambda-1)K}}{\varepsilon_2^\lambda-\varepsilon_2}\right)\cdot\|\mathbf{e}(k,n)\|_\lambda + \varepsilon_7 b_x + \varepsilon_8 b_d\end{aligned} \tag{C.13}$$

where $\varepsilon_7 = \sup_{k\in[0,K]}\varepsilon_2^{-\lambda k}\varepsilon_5$; $\varepsilon_8 = \sup_{k\in[0,K]}\varepsilon_2^{-\lambda k}\varepsilon_6$.

With condition of $\|\Psi\| < 1$, and a sufficiently large λ, there exists,

$$\left(\|\Psi\| + \varepsilon_1\varepsilon_3 k_{g_x}\frac{1-\varepsilon_2^{-(\lambda-1)K}}{\varepsilon_2^\lambda-\varepsilon_2}\right) \leq \rho < 1 \tag{C.14}$$

From (C.13),

$$\|\mathbf{e}(k,n+1)\|_\lambda \leq \rho\|\mathbf{e}(k,n)\|_\lambda + \varepsilon_7 b_x + \varepsilon_8 b_d \leq \rho^{n+1}\|\mathbf{e}(k,0)\|_\lambda + \frac{\varepsilon_7(1-\rho^{n+1})}{1-\rho}b_x + \frac{\varepsilon_8(1-\rho^{n+1})}{1-\rho}b_d \tag{C.15}$$

Taking limit as $n\to\infty$ with bounded $\|\mathbf{e}(k,0)\|_\lambda$, there exists,

$$\lim_{n\to\infty}\|\mathbf{e}(k,n)\|_\lambda \leq \frac{\varepsilon_7}{1-\rho}b_x + \frac{\varepsilon_8}{1-\rho}b_d = c_x b_x + c_d b_d \tag{C.16}$$

where $c_x = \frac{\varepsilon_7}{1-\rho}$; $c_d = \frac{\varepsilon_8}{1-\rho}$.

From (C.3), $\|\delta\mathbf{u}(k,n)\| \leq b_\beta\|\mathbf{e}(k,n)\|$, then

$$\lim_{n\to\infty}\|\delta\mathbf{u}(k,n)\|_\lambda \leq b_\beta c_x b_x + b_\beta c_d b_d \tag{C.17}$$

From (C.14), the convergence condition is $\|\Psi\| < 1$. Define

$$\Phi = \beta\text{diag}\{\mathbf{u}(k,n)\}^{-1}\text{diag}\{\mathbf{o}(k,n+1)\}$$

where Φ is a matrix with all non-negative entries.

Then

$$\Psi = (I - \Phi A)A \tag{C.18}$$

Based on Lemma 3, taking Euclidean norm and supremum of Φ over [0,K],

$$\sup_{k\in[0,K]}\|\Phi\| = \sup_{k\in[0,K]}(\|\beta\text{diag}\{\mathbf{u}(k,n-1)\}^{-1}\text{diag}\{\mathbf{o}(k,n)\}\|) = \frac{b_\beta b_o}{u_{min}} \tag{C.19}$$

where $b_o = \sup_{k\in[0,K]}(\|\text{diag}\{\mathbf{o}(k,n)\}\|)$, $u_{min} = \inf_{k\in[0,K]}(\|\text{diag}\{\mathbf{u}(k,n-1)\}\|)$.

If $0 < \frac{b_\beta b_o}{u_{min}} < 2$, $\|\Psi\| < 1$, therefore,

$$0 < b_\beta < \frac{2u_{min}}{b_o} \tag{C.20}$$

Theorem 1 is proved. □

Acknowledgment

The work is supported by the Singapore-MIT Alliance for Research and Technology (SMART) Center for Future Urban Mobility (FM). This work is also supported by National Nature Science Foundation under Grant 61573054.

References

[1] M. Papageorgiou, C. Diakaki, V. Dinopoulou, A. Kotsialos, and Y. Wang, "Review of road traffic control strategies," *Proceedings of the IEEE*, vol. 91, no. 12, pp. 2043–2067, December 2003.
[2] K. Aboudolas, M. Papageorgiou, and E. Kosmatopoulos, "Store-and-forward based methods for the signal control problem in large-scale congested urban road networks," *Transportation Research Part C: Emerging Technologies*, vol. 17, no. 2, pp. 163–174, 2009.
[3] S. Lin, B. De Schutter, Y. Xi, and H. Hellendoorn, "Efficient network-wide model-based predictive control for urban traffic networks," *Transportation Research Part C*, vol. 24, pp. 122 –140, 2012.
[4] E. B. Kosmatopoulos, M. Papageorgiou, A. Vakouli, and A. Kouvelas, "Adaptive fine-tuning of nonlinear control systems with application to the urban traffic control strategy tuc," *Control Systems Technology, IEEE Transactions on*, vol. 15, no. 6, pp. 991–1002, 2007.
[5] S. M. Festin, "Summary of national and regional travel trends 1970-1995," Tech. Rep., 1996.
[6] R. P. Roess, E. S. Prassas, and W. R. McShane, *Traffic Engineering*, 3rd ed. Prentice Hall, 2004.
[7] Y. Wang, D. Wang, C. Yang, N. Xiao, Y. Li, and E. Frazzoli, "Phase-based repetitiveness and pattern classification of urban traffic flow data," in *14th ITS Asia Pacific Forum*, 2015.
[8] F. V. Webster, "Traffic signal settings," Department of Scientific and Industrial Research Road Research Laboratory, Tech. Rep., 1958.
[9] T. Hauser and W. T. Scherer, "Data mining tools for real-time traffic signal decision support & maintenance," in *Systems, Man, and Cybernetics, 2001 IEEE International Conference on*, vol. 3, 2001, pp. 1471–1477.
[10] B. L. Smith, W. T. Scherer, T. A. Hauser, and B. B. Park, "Data–driven methodology for signal timing plan development: A computational approach," *Computer-Aided Civil and Infrastructure Engineering*, vol. 17, no. 6, pp. 387–395, 2002.
[11] B. Brian Park, P. Santra, I. Yun, and D.-H. Lee, "Optimization of time of day plan scheduling using a multi-objective evolutionary algorithm," *Transportation Research Record: Journal of the Transportation Research Board*, no. 1867, pp. 217–223, 2004.
[12] Y. Wong and W. Woon, "An iterative approach to enhanced traffic signal optimization," *Expert Systems with Applications*, vol. 34, no. 4, pp. 2885–2890, 2008.
[13] R. Guo and Y. Zhang, "Identifying time-of-day breakpoints based on nonintrusive data collection platforms," *Journal of Intelligent Transportation Systems*, vol. 18, no. 2, pp. 164–174, 2014.
[14] F. Luyanda, D. Gettman, L. Head, S. Shelby, D. Bullock, and P. Mirchandani, "Acs-lite algorithmic architecture: applying adaptive control system technology to closed-loop traffic signal control systems," *Transportation Research Record: Journal of the Transportation Research Board*, no. 1856, pp. 175–184, 2003.
[15] H.-S. Ahn, Y. Chen, and K. L. Moore, "Iterative learning control: brief survey and categorization," *IEEE Transactions on Systems Man and Cybernetics part C Applications and Reviews*, vol. 37, no. 6, pp. 1099–1121, 2007.
[16] D. Bristow, M. Tharayil, and A. G. Alleyne, "A survey of iterative learning control," *Control Systems, IEEE*, vol. 26, no. 3, pp. 96–114, 2006.
[17] D. Wang and C. C. Cheah, "An iterative learning- control scheme for impedance control of robotic manipulators," *The International Journal of Robotics Research*, vol. 17, no. 10, pp. 1091–1104, 1998.
[18] Z. Hou, J. Yan, J.-X. Xu, and Z. Li, "Modified iterative-learning-control-based ramp metering strategies for freeway traffic control with iteration-dependent factors," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 13, no. 2, pp. 606–618, June 2012.
[19] W. Huang, F. Viti, and C. M. Tampere, "An iterative learning approach for signal control in urban traffic networks," in *Intelligent Transportation Systems Conference (ITSC)*, 2013.
[20] Y. Wang, D. Wang, N. Xiao, Y. Li, and E. Frazzoli, "Iterative tuning strategy for setting phase splits in traffic signal control," in *Intelligent Transportation Systems (ITSC), 2014 IEEE 17th International Conference on*, 2014, pp. 2453–2458.
[21] P. J. Olver and C. Shakiban, *Applied Linear Algebra*. Prentice-Hall, Upper Saddle River, NJ, 2006.
[22] E. Bert, "Dynamic urban origin-destination matrix estimation methodology," Tech. Rep., 2009.
[23] D. A. Harville, *Matrix algebra from a statistician's perspective*. Springer, 2008.