Steady-state Control for Signalized Intersections Modeled as Switched Server System*

Zhonghe He¹, Yangzhou Chen¹, Jianjun Shi², Xingguang Han¹, and Xu Wu³

Abstract—This paper proposes two kinds of signal control strategies for a multiple-phase signalized intersection modeled as the switched server system. Firstly, a class of switched server systems is introduced, and then two kinds of scheduling policies of the server, i.e., priority-based scheduling policy (PBSP) and clear-oldest-buffer scheduling policy (COBSP), are considered, under which the switched server system is periodically stable. Furthermore, the signal control for a multiple-phase signalized intersection is modeled as a switched sever system, and then the proposed PBSP and COBSP are applied to the steady-state control for the signalized intersection. At last, we illustrate our methods by the simulation for an intersection.

I. INTRODUCTION

Signalized intersections can be viewed as nodes in urban road networks. And introducing traffic signal light in an intersection is an effective way to avoiding conflicts between incompatible traffic flows and, thus ensures traffic safety for drivers and pedestrians. At the same time, effective traffic signal control can also enhance utilization efficiency of the capacity in an intersection, and then present the foundation for urban road network coordinated control.

Different kinds of methods and models have been proposed for the signal control of signalized intersections. For example, fixed-time control methods [1], [2], [3], [4], optimal control methods, e.g., linear quadratic optimal control [5], [6], and Pontryagin maximum principle [7], [8], hybrid systems methods [9], [10], extended linear complementarity programming [11], steady-state control methods [12], [13], [14], to name just a few.

Traffic signal light within an intersection serves as a server, and the lane groups on incoming links of an intersection are similar to buffers which store accumulated vehicles planning to pass the intersection. Meantime, traffic signal light alternately provides vehicles on lane groups in different phases with right of way. Thus, the switched server system theory can be applied for the signal control of signalized intersections.

Specifically, a switched server system consists of one server and a finite number of buffers alternately served by the server according to the designed scheduling policy. Under the given scheduling policy of the server, the switched server system is a special class of hybrid systems, and more detailed discussions about switched server systems can be referred to [15]. The switched server system can be used to model some real-world systems, e.g., manufacturing systems [16], [17], traffic signal control systems [14]. Some scheduling policies of the server have been proposed, e.g., priority-based scheduling policy (PBSP) [18], clear-oldest-buffer scheduling policy (COBSP) [19]. In [20], [21], the periodic stability of the switched server system with two buffers was considered, the idea in which was that the optimal periodic solution minimizing some criterion was first solved, and then the scheduling policy of the server was designed such that the solution of the system asymptotically converges to the optimal periodic solution. In [12], [13], steady-state control for signalized intersections with two incompatible vehicle flows was considered, the idea in which was similar to that in [20], [21]. However, for the high-dimensional systems, the analytical form of the optimal periodic solution can not be derived in general. In [14], steady-state control for T-shaped signalized intersections modeled as a switched server system, was considered by applying the PBSP, where saturation flow rates of vehicles on the lane groups were assumed to be equal. The assumptions about the type of the intersections and saturation flow rates in [14] limit its further applications to more general cases.

In this paper, two kinds of signal control strategies for a multiple-phase signalized intersection modeled as the switched server system, are presented. Firstly, a class of switched server systems is introduced, and periodic stability of the switched server system under the PBSP and COBSP is proved. Furthermore, the signal control for a multiple-phase signalized intersection is modeled as a switched sever system, and the proposed PBSP and COBSP are respectively applied to the steady-state control for the signalized intersection.

The paper is organized as follows. In section II, periodic stability of the switched server system under the PBSP and COBSP is proved. In section III, the signal control for a signalized intersection is modeled as a switched server system, and the proposed PBSP and COBSP are respectively applied to the steady-state control for the signalized intersection. In section IV, conclusions and topics for future research are
II. STABILITY OF SWITCHED SERVER SYSTEM

A switched server system (Fig.1) consists of one server and \( n \) \((n \geq 2)\) buffers, where the work arrives at the buffer \( i, i = 1, \ldots, n \) at a constant rate \( d_i > 0 \) and the work leaves the buffer \( i \) at a constant rate \( s_i > 0 \) when the buffer \( i \) is served by the server. Both constant rates \( d_i \) and \( s_i \) are called arriving rate and leaving rate of the work in the buffer \( i \), respectively. We denote by \( x_i(t), i = 1, \ldots, n \) the workload in the buffer \( i \) at time instant \( t \). Due to non-negative constraints on the workload in the buffer, the state space of the system is defined as \( X =: \{ [x_1, \ldots, x_n]^T \in \mathbb{R}^n : x_i \geq 0, i = 1, \ldots, n \} \). When the server switches from serving the buffer \( i \) to serving the buffer \( j \), there is a setup time \( l_{ij} > 0, i, j = 1, \ldots, n, i \neq j \), during which the server does not serve any buffer.

From above descriptions for a switched server system, the dynamics of the workload in the buffer can be described as follows:

\[
\begin{align*}
\dot{x}_1(t) &= d_1 \\
\vdots \\
\dot{x}_{i-1}(t) &= d_{i-1} \\
\dot{x}_i(t) &= d_i - s_i, i = 1, \ldots, n \\
\dot{x}_{i+1}(t) &= d_{i+1} \\
\vdots \\
\dot{x}_n(t) &= d_n
\end{align*}
\]  

(1)

Equation (1) corresponds to the case when the server is serving the buffer \( i \) and equation (2) corresponds to the case when the server is switching from serving one buffer to serving another one.

We assume that the following inequality holds in the succeeding sections,

\[
\sum_{j=1}^{n} \left( \frac{d_j}{s_j} \right) < 1. \tag{3}
\]

The inequality (3) is the necessary condition for keeping the buffer bounded for any scheduling policy [16].

A. Stability of Priority-based Scheduling Policy

In this subsection, we consider priority-based scheduling policy (PBSP) of the server, which is described as follows:

1) The server starts with the buffer 1;
2) Whenever the server has emptied the workload in the buffer \( i \), i.e., \( x_i(t) = 0 \), it switches to serving the buffer \( i+1 \), \( i = 1, \ldots, n-1 \); Whenever the server has emptied the workload in the buffer \( n \), i.e., \( x_n(t) = 0 \), it switches to serving the buffer 1 again;
3) When the server switches from serving the buffer \( i \) to serving the buffer \( i+1 \), \( i = 1, \ldots, n-1 \), there is a setup time \( l_{i(i+1)} > 0 \); When the server switches from serving the buffer \( n \) to serving the buffer 1, the setup time is \( l_{n1} \).

From the descriptions of the PBSP, there exist \( n \) sets \( X_i, i = 1, \ldots, n \), defined as

\[
X_i =: \{ [x_1, \ldots, x_n]^T \in \mathbb{R}^n : x_i = 0, x_j > 0, j \neq i \}, \tag{4}
\]

where the switching of the server from serving the buffer \( i \) to serving the buffer \( i+1 \), occurs in the set \( X_i, i = 1, \ldots, n-1 \), and the switching of the server from serving the buffer \( n \) to serving the buffer 1, occurs in the set \( X_n \).

Theorem I: Consider the switched server system described by (1) and (2) under the PBSP. If arriving rates \( d_i \), and leaving rates \( s_i, i = 1, \ldots, n \) are chosen such that the inequality (3) is satisfied, then the following statements hold

1) There exists a unique periodic solution for the switched server system, which is globally asymptotically stable.
2) The period of the periodic solution is computed by

\[
C = \frac{L}{1 - \sum_{j=1}^{n} \left( \frac{d_j}{s_j} \right)}, \tag{5}
\]

where \( L = l_{12} + \cdots + l_{(n-1)n} + l_{n1} \).
3) Let \( g_i, i = 1, \ldots, n \) denote the time provided by the server to serve the work in the buffer \( i \) during the steady period \( C \). Then,

\[
g_i = (d_i/s_i)C = \frac{(d_i/s_i)}{\sum_{j=1}^{n} \left( \frac{d_j}{s_j} \right)}(C - L). \tag{6}
\]

4) The coordinates of the switching points \( x_i^* \in X_i, i = 1, \ldots, n \)
1, . . . , n of the periodic solution are determined by

\[
x^*_i = \begin{bmatrix} x^*_{i1}, \ldots, x^*_{in} \end{bmatrix}^T
\]

\[
= [d_i(\sum_{j=1}^{i-1} l_{j(i+1)} + \sum_{j=2}^{i} g_j), \ldots, \\
d_h(\sum_{j=h}^{i-1} l_{j(i+1)} + \sum_{j=h+1}^{i} g_j), \ldots, \\
d_{i-1}(l_{(i-1)h} + g_i), 0, \\
d_{i+1}(C - g_{i+1} - l_{i(i+1)}), \ldots, \\
d_{i+k}(C - \sum_{j=i+1}^{i+k} g_j - \sum_{j=i+1}^{i+k-1} l_{j(i+1)}), \ldots, \\
d_n(C - \sum_{j=1}^{n} g_j - \sum_{j=i+1}^{n-1} l_{j(i+1)})]^T.
\]

**Proof:** Consider linear transformation: \( y = Sx \), where \( x = [x_1, \ldots, x_n]^T \in X \) and \( S \) is a diagonal matrix \( \text{diag}(1/s_1, \ldots, 1/s_n) \). A new switched server system can be derived by applying above linear transformation to (1) and (2). For the derived switched server system, arriving rate and leaving rate of the work in the buffer \( i \) are \( d_i/s_i \) and 1, respectively. The statements 1) and 2) in Theorem 1 can be derived by applying Theorem 7 in [18].

For the statement 3), let \( x(t), t \geq 0 \) be the periodic solution to the switched server system with the initial condition \([x_1(0), \ldots, x_n(0)]^T\) and \( \{t_k\}_{k=0}^\infty, t_0 = 0 \) be the strictly monotone increasing switching time sequence of the server for the periodic solution \( x(t) \), where \( t_{2k-2} \) and \( t_{2k-1}, k = 1, 2, \ldots \) are respectively the time instant the server starts serving the buffer and the time instant the server finishes service for the buffer. Then, \( t_{2n} \) is the time instant the server finishes a complete periodic switching \((1 \rightarrow 2 \rightarrow \cdots \rightarrow n \rightarrow 1)\) for the first time, i.e., \( C = t_{2n} \), and \( t_{2i-2}, i = 1, \ldots, n \) is the time instant the server starts serving the buffer \( i \) during the period \( C \). From the descriptions of the PBSP, the workload in the buffer \( i \) keeps increasing before the buffer \( i \) is served by the server, thus, \( g_i \) is the solution to the following system of equations,

\[
\begin{align*}
x_i(t) + d_i t_{2i-2} + (d_i - s_i) g_i &= 0, \\
x_i(0) - d_i(C - t_{2i-2} - g_i) &= 0,
\end{align*}
\]

(7)

Solving (7) can obtain \( g_i = (d_i/s_i) C \), and

\[
g_i = (d_i/s_i) C = \frac{d_i/s_i L - L (1 - \sum_{j=1}^{n} (d_j/s_j))}{\sum_{j=1}^{n} (d_j/s_j)} = \frac{d_i/s_i}{\sum_{j=1}^{n} (d_j/s_j)} (C - L).
\]

For the statement 4), let the coordinates of the switching points of the periodic solution \( x(t) \) be \( x^*_i = [x^*_{i1}, \ldots, x^*_{in}]^T \in X_i, i = 1, \ldots, n \). From the definition of the set \( X_i \) in (4), one has \( x^*_{i1} = 0 \). Furthermore, let \( \{T^g_{ij}\}_{j=1}^\infty, i = 1, \ldots, n \) be the time sequence, where \( T^g_{ij} \) is the time instant the server finishes service for the buffer \( i \) within the \( g \)th periodic switching of the server. Then,

\[
T^h_{i} - T^{h-1}_i = l_{h(i+1)} + g_{h+1} + \cdots + l_{i-1i} + g_i
\]

For the periodic solution under any fix ed switching sequence of the server is equal to \( C = nl/(1 - \sum_{j=1}^{n} (d_j/s_j)) \).

**Fig. 2.** Simulation results under the PBSP. Initial condition is chosen as \([x_1(0), x_2(0), x_3(0)]^T = [50, 80, 70]^T\). Switching sequence of the server: 1 \( \rightarrow \) 2 \( \rightarrow \) 3 \( \rightarrow \) 1 \( \cdots \).

**Remark 1:** If the parameters of the switched server system described by (1) and (2) under the PBSP are known, one can obtain all the information about the periodic solution according to the results in Theorem 1. The physical significance of the periodic solution in Theorem 1 is that for any buffer, the number of works arriving at the buffer within \( C \) equals to the number of works served by the server within \( C \). By the period formula (5), the period \( C \) will increase as the setup times \( s_{ij} \) increase, and if the setup times \( s_{ij} \) among buffers are a common constant, e.g., some \( l > 0 \) not depending on the switching sequence of the server, the period of the periodic solution under any fixed switching sequence of the server is equal to \( C = nl/(1 - \sum_{j=1}^{n} (d_j/s_j)) \).

**Example 1:** Consider a switched server system with three buffers, where arriving rates and leaving rates of the work in each buffer are respectively chosen as \( d_1 = 0.8, d_2 = 0.5, d_3 = 0.7, s_1 = 2, s_2 = 2.5, s_3 = 3.5 \) [lots/s], and the setup times are respectively chosen as \( l_{12} = 3, l_{23} = 4, l_{31} = 3 \) [s]. It is easy to verify that \( d_i \) and \( s_i, i = 1, 2, 3 \) satisfy the inequality (3). Simulation results under the PBSP are illustrated in Fig.2.

The fixed switching sequence of the server is assumed when applying the PBSP. In the next subsection, we consider a scheduling policy of the server, which can optimize the switching sequence of the server.

846
B. Stability of Clear-oldest-buffer Scheduling Policy

Clear-oldest-buffer scheduling policy (COBSP) is described as follows:

Consider the switched server system described by (1) and (2). Assume that the $i$th buffer is currently served by the server, and the switching of the server occurs at time instant $t = t_k$, $t_0 = 0$, then the next served buffer index is determined by (8),

$$i^*(t_k) = \min_i \left\{ i : \frac{x_i(t_k)}{d_i} = \max \left\{ \frac{x_1(t_k)}{d_1}, \ldots, \frac{x_n(t_k)}{d_n} \right\} \right\}. \tag{8}$$

1) The server starts with the buffer $i$ determined by the initial condition $[x_1(0), \ldots, x_n(0)]^T$, $t_0 = 0$ and (8);
2) Whenever the server has emptied the workload in the buffer $i$ at time instant $t_k$, i.e., $x_i(t_k) = 0$, it switches to serving the buffer $i^*$ determined by (8);
3) When the server switches from serving the buffer $i$ to serving the buffer $i^*$, there is a setup time $t_{i,i^*} > 0$.

Theorem 2: Consider the switched server system described by (1) and (2) under the COBSP. If arriving rates $d_i$, and leaving rates $s_i, i = 1, \ldots, n$ are chosen such that the inequality (3) is satisfied, then there exist $(n-1)!$ periodic solutions to the switched server system and any other solution of the system asymptotically converges to one of periodic solutions.

Proof: Construct linear transformation: $y = Sx$, where $x = [x_1, \ldots, x_n]^T \in X$ and $S$ is a diagonal matrix defined in the proof of Theorem 1. The new switched server system with arriving rate $d_i/s_i$ and leaving rate 1 can be derived by applying above linear transformation to (1) and (2). The statements in Theorem 2 can be derived by applying Theorem 5 in [19].

Remark 2: The significance of the COBSP is that the server gives the priority to the buffer having longest waiting time (refer to (8)). Applying the COBSP, the switching sequence of the server is uniquely determined by the initial condition. For example, consider a switched server system with three buffers, if applying the COBSP and choosing the initial condition $[x_1(0), x_2(0), x_3(0)]^T \in X$ such that the condition $x_2(0)/d_2 > x_3(0)/d_3 > x_1(0)/d_1$ holds, then the switching sequence of the server is $2 \rightarrow 3 \rightarrow 1 \rightarrow 2 \rightarrow \ldots$.

III. STEADY-STATE CONTROL FOR SIGNALIZED INTERSECTIONS

In this section, we firstly set up the relation between the switched server system described by (1) and (2) and the signal control for a signalized intersection, and furthermore, apply the PBSP and COBSP proposed in section II to steady-state control for a signalized intersection.

Consider a signalized intersection with $n$ ($n \geq 2$) phases determined in advance. The lane group on the incoming link of a signalized intersection is the basic unit for the analysis of a signalized intersection, which may consist of multiple lanes depending on the traffic conditions. The traffic signal light within a signalized intersection serves as the server and the lane group is similar to a buffer storing the accumulated vehicles planning to pass the intersection. Thus, we can apply the switched server system theory to the signal control for a signalized intersection.

Assume that there are $m_i$ lane groups contained in phase $i$, and arriving rate and saturation flow rate of vehicles on the $p$th lane group in phase $i$, respectively are $d_{i,p}$ and $s_{i,p}$, $i = 1, \ldots, n$, $p = 1, \ldots, m_i$, $m_i \geq 1$. Furthermore, assume that there are no overlapping phases, i.e., there are no vehicle flows continuing right of way among multiple phases. The flow ratio on a lane group is defined as the ratio of arriving rate to saturation flow rate of vehicles on this lane group. In this paper, the queue length of vehicles is defined as the number of vehicles behind the stopline. Under above assumptions, the switched server system model of a signalized intersection can be established by the following procedures:

1) The lane group having the maximum flow ratio among all lane groups in phase $i$, is chosen, which is called the critical lane group and indexed by $(i, i_k)$, $i = 1, \ldots, n$, $i_k \in \{1, \ldots, m_i\}$. The vehicle flow on the critical lane group is called the critical vehicle flow;
2) The dynamics of the queue length of vehicles on the critical lane group in each phase can be described by (1) and (2). The lost time per phase is represented by the setup time of the server switching from serving one
Remark 3: For the switched server system model of a signalized intersection derived from above procedures, arriving rates and saturation flow rates of vehicles are assumed to be constant, which can be computed by the average of corresponding quantities of vehicles in some time interval.

Assume that arriving rates \(d_{(i,i_k)}\) and saturation flow rates \(s_{(i,i_k)}\) of the critical vehicle flows in each phase \(i\) satisfy the inequality (3). If the phase sequence is determined as \(1\rightarrow 2\rightarrow \cdots \rightarrow n\), and the PBSP is applied to the signal control for the signalized intersection, the fixed-time signal timing formula can be derived,

\[
C = \frac{L}{1 - \sum_{j=1}^{n} \left(\frac{d_{(j,j_k)}}{s_{(j,j_k)}}\right)},
\]

\[
g_i = \frac{d_{(i,i_k)}}{s_{(i,i_k)}}C = \frac{d_{(i,i_k)}}{s_{(i,i_k)}}\left(C - L\right), i = 1, \ldots, n
\]

where \(C\) is the cycle length, and \(g_i\) is the green time duration of phase \(i\).

Due to the statements in Theorem 1, \(C\) in (9) is the period of the periodic solution for the switched server system model of a signalized intersection, which coincides with the minimum cycle length avoiding the oversaturation of the signalized intersection (refer to [3]).

If arriving rates \(d_{(i,i_k)}\) and saturation flow rates \(s_{(i,i_k)}\) of the critical vehicle flows in each phase \(i\) satisfy the inequality (3), and the COBSP is applied to the signal control for the signalized intersection, then, due to the results in Theorem 2 the queue length of vehicles on the critical lane group asymptotically converges to the periodic queue length, and the period of the periodic solution will depend on the initial conditions, e.g., refer to the discussions about the COBSP in Remark 2.

Remark 4: In [12, 13], the steady-state control for a signalized intersection with two incompatible vehicle flows (i.e., in the case of \(n = 2\)) was considered. The basic idea is that an optimal periodic solution minimizing some criterion is first solved, and furthermore, the optimal switching time sequences are solved, such that the solution of system under given initial condition converges to the optimal periodic solution. The periodic solution considered in this paper may be not optimal under the criterion in [12, 13], which depends on the choice of the criterion and parameters of the system. More detailed discussions can also be referred to [21]. However, in the case of \(n > 2\), the analytical form of an optimal periodic solution minimizing some criterion can not be derived in general. From the results in Theorem 1, all the information about the periodic solution can be obtained if the parameters of the system are given. Thus, we consider the periodic solution in Theorem 1 as criterion, design the PBSP for the signal control, and optimize the phase sequence by the COBSP.

In the following parts, we consider a cross-shaped intersection (Fig.4) for illustrating above signal timing procedures.

**Fig. 4. Cross-shaped intersection and its phase partition.**

### TABLE I

<table>
<thead>
<tr>
<th>Phase</th>
<th>Traffic parameters for cross-shaped intersection.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(d_1) veh/s</td>
</tr>
<tr>
<td>1</td>
<td>0.08</td>
</tr>
<tr>
<td>2</td>
<td>0.06</td>
</tr>
<tr>
<td>3</td>
<td>0.05</td>
</tr>
<tr>
<td>4</td>
<td>0.07</td>
</tr>
</tbody>
</table>

Without loss of generality, we assume that all the right turning vehicle flows do not be controlled by the traffic signal light. Let \(x_i, i = 1, \ldots, 8\) denote the queue length of vehicles on the \(i\)th lane group with arriving rate \(d_i\) [veh/s] and saturation flow rate \(s_i\) [veh/s] (Fig.4). Necessary traffic parameters are listed in table I.

There are two controlled lane groups in each phase (Fig.4). Thus, we first choose the critical lane group in each phase, which has larger arriving rate of vehicles in that phase. From table I, the lane groups 2, 4, 5, 7 are respectively chosen for the signal control of the intersection. Moreover, \(d_i\) and \(s_i, i = 2, 4, 5, 7\) satisfy the inequality (3).

When the PBSP is applied and the lost time per phase is assumed to be 3s, simulation results are given in Fig.5. The cycle length and green time duration of each phase for the periodic solution are respectively determined by (5) and (6). In this example, \(C = 120\)s, \(g_1 = 30\)s, \(g_2 = 18\)s, \(g_3 = 45\)s, \(g_4 = 15\)s.

When the COBSP is applied, simulation results are given in Fig.6. The cycle length and green time duration of each phase for each of 3! periodic solutions are the same as those.
Comparing with the PBSP, the initial condition in Fig.6 satisfies the condition \( x_4(0)/d_4 > x_7(0)/d_7 > x_2(0)/d_2 > x_5(0)/d_5 \), thus, the phase sequence is \( 2 \rightarrow 4 \rightarrow 1 \rightarrow 3 \rightarrow 2 \rightarrow \cdots \). (refer to the discussions in Remark 2).

IV. CONCLUSIONS

In this paper, the relation between the scheduling policy for the switched server system and the signal control for a signalized intersection is established. Furthermore, two kinds of signal control strategies, i.e., PBSP and COBSP, are proposed for the steady-state control for signalized intersections. The COBSP can be applied to optimize the phase sequence of signalized intersections. For the design of the PBSP and COBSP, we do not consider the capacity constraints (i.e., \( x_i(t) \leq x_{i}^{\text{max}} \)) on the lane groups on incoming links of the signalized intersection. The analytical form of the periodic solution considered in Theorem 1 can be derived if the parameters of the system are given, the future research will focus on the design of signal control strategies by using the information about the periodic solution, such that the solution to the system asymptotically converges to this periodic solution while satisfying the capacity constraints on the lane groups. Furthermore, we will consider the execution of the PBSP and COBSP in stochastic case.

REFERENCES