Transportation Research Part B 45 (2011) 667–681

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Optimal allocation of turns to lanes at an isolated signal-controlled junction

C.K. Wong [a,*], B.G. Heydecker [b]

[a] Department of Building and Construction, City University of Hong Kong, Tat Chee Avenue, Kowloon Tong, Hong Kong SAR, Hong Kong
[b] Centre for Transport Studies, Department of Civil, Environmental, and Geomatic Engineering, University College London, Chadwick Building, Gower Street, London WC1E 6BT, United Kingdom

ARTICLE INFO

Article history:
Received 2 February 2010
Received in revised form 26 November 2010
Accepted 4 December 2010

Keywords:
Traffic signal control
Signal timing optimisation
Junction geometric design
Lane-based optimisation
Mixed-Integer-Programming

ABSTRACT

Conventional design methods require the lane marking patterns, which are painted on ground showing road users the permissible turning directions on different approach lanes, as exogenous inputs to define the traffic stream grouping for analysis. This predefined grouping of traffic movements may restrict the design of signal timings in the optimisation procedures. More recently, a lane-based design method has been developed to relax the lane markings as binary-type control variables in a mathematical programming approach. The lane marking patterns and the signal timings can then be optimised simultaneously in a unified framework. This paper presents an extension work to further relax the numbers of approach lane in traffic arms as new integer variables which can then be optimised to give optimal lane arrangement in various arms of a junction to manage the given traffic demands more efficiently. All well-defined signal timings variables in the phase-based approach as well as the lane marking and lane flow variables in the lane-based approach together with their governing constraints are all preserved in the new formulation for the reserve capacity optimisation of isolated signal-controlled junctions.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

The earliest designs for traffic signal settings taking on uniform traffic arrival patterns as inputs were fixed-time signals in which the signal timings for different traffic groups were changed periodically and repeated in cycles (Webster, 1958; Webster and Cobbe, 1966) Allsop (1972) enhanced the signal setting calculation to determine the traffic capacity for isolated signal-controlled junctions. A common flow multiplier was first introduced into the design framework as the objective for optimisation indicating the overall junction performance. The problem of determining maximum junction capacity has been formulated as a linear programming (LP) model that can be effectively solved by the simplex method. The procedures were implemented in a computer program, called SIGCAP, for assessing junction capacities at isolated signal-controlled junctions (Allsop, 1975, 1976). Stage structures and sequences in a signal cycle are key inputs in the design framework. To relax this requirement and to ensure the inclusion of an optimal stage plan while designing signal settings, Tully (1976) developed an algorithm to generate candidate stages and all admissible stage sequences based on the concept of graph theory. Based on Tully's procedure, Gallivan and Heydecker (1988) designed a new algorithm that does not restrict clique formations and there might be multiple appearances for an identical clique (stage) occurring in one clique sequence (stage sequence). Extra inter-stage loss time induced for those traffic streams involved may seriously degrade the overall junction performance. Heydecker (1992) further reduced this difficulty by grouping all the possible cycle structures into a much smaller number of equivalence classes. A successor function was then introduced in the formulation for the representation of each of these

* Corresponding author.
E-mail addresses: wongck@cityu.edu.hk (C.K. Wong), ben@transport.ucl.ac.uk (B.G. Heydecker).

0191-2615/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.trb.2010.12.001

equivalence classes. This reduction in the number of cases for evaluation greatly enhanced the overall computational attractiveness. As the successor function was operated down to individual traffic streams, it was well suited for all phase-based formulations. Improta and Cantarella (1984) formulated a group-based approach for the problem of signal settings optimisation as a Binary-Mixed-Integer-Linear-Program (BMILP). Traffic groups that cannot move together while at the same time fulfilling safety requirements are regarded as incompatible groups and sufficient intergreen or clearance times have to be provided to separate them well within a signal cycle. A set of binary variables and relevant constraints were introduced to control the relative order of receiving right of way for traffic groups. The algorithm assigned signal timings to each traffic group individually while optimising the overall junction capacity and total delay for signalised junctions without the necessity to maintain an explicit stage structure and sequence. A standard branch-and-bound technique was then employed to solve the approximated BMILP problems. Allsop (1992) summarised the mathematical formulations of both the stage- and phase- (or group-) based approaches and the optimisation methodologies for the design of signal-controlled junctions. Various constraints on signal timings and optimisation criteria were discussed.

Recently, the lane-based optimisation method combining the design of lane markings and signal settings for isolated signal-controlled junctions was developed and its formulation was a direct extension of the phase-based method in which all phase-based variables and relevant governing constraints are preserved and revised accordingly to amalgamate with other new lane-based variables and constraints. The design approach was termed the lane-based method as all key design variables were given on a lane-basis. The lane markings were defined as discrete binary variables and integrated into the design optimisation framework. The set of binary variables also included the successor functions governing the order of signal displays and the continuous set of variables comprised the assigned lane flows, common flow multiplier, cycle length, as well as starts and durations of green for traffic movements and on approach lanes. The lane-based method was formulated as a mathematical program in which a set of linear constraints is set up to ensure the feasibility and safety of the resultant optimised lane markings and signal settings. The junction capacity maximisation and cycle length minimisation problems were formulated as Binary-Mixed-Integer-Linear-Programs (BMILP) which are solvable by any standard branch-and-bound routine (Wong and Wong, 2003a,b; Wong et al., 2006).

All of the methods reviewed so far take on the junction geometric layout such as the numbers of approach and exit lanes in different arms as fixed input parameters although the original lane-based method has developed to optimise the individual lane usage. It is believed that the lane arrangement within each arm in a junction is crucial to the overall junction performance. If one more approach lane is given to serve a traffic stream, the saturation flow representing the rate of discharge of that traffic stream are increased substantially and lesser green times are required to let through the same amount of traffic. To add an approach lane, however, implies that one exit lane is lost within the same arm to serve the leaving traffic from all other arms while retaining the same junction size. Besides, sufficient numbers of exit lanes of other arms should be arranged to serve the traffic due to the additional approach lane. It becomes a very complex interaction even without considering the tradeoffs in signal timings. The present paper aims to formulate the junction design problem into a mathematical program to include the junction geometric layout, individual lane usages, and signal timings as design variables which can be optimised simultaneously to achieve better junction capacity. Section 2 summarises the necessary data inputs in the design framework. Section 3 introduces the control variables used in the present formulation. Section 4 presents the necessary governing constraints confining the feasible solution region for all of the control variables. The optimisation problem for maximising the junction reserve capacity is formulated in Section 5. A solution algorithm is given in Section 6. Numerical examples are provided in Section 7 for demonstration of the proposed algorithms for practical junction design problems. Finally, there is a brief conclusion in Section 8.

2. Input data

To implement the present method for the design analysis of isolated signal-controlled junctions, various input data, specifying the demand pattern, junction geometries, allowable capacity limits, and other signal timing restrictions, are required. The required input data set is given as follows. This section lists all the required input data and detailed discussions will be given in Section 4.

Number of traffic arms	N_T
Demand flows	$Q_{i,j}$, $\forall i = 1, \ldots, N_T$; $j = 1, \ldots, N_T - 1$
Total number of lanes in traffic arms	L_i, $\forall i = 1, \ldots, N_T$
Minimum (Maximum) cycle length	$c_{\min}(c_{\max})$
Minimum green durations	$g_{i,j}$, $\forall i = 1, \ldots, N_T$; $j = 1, \ldots, N_T - 1$
Minimum clearance times	$\omega_{u,v}$, $\forall (u, v) \in \Psi^{\#}$
Lane saturation flows for straight-ahead movement	$s_{i,k}$, $\forall i = 1, \ldots, N_T$; $k = 1, \ldots, L_i$
Maximum acceptable degree of saturation	$p_{i,k}$, $\forall i = 1, \ldots, N_T$; $k = 1, \ldots, L_i$
Conversion factor of turning flow to through car unit	$\tau_{i,j}$, $\forall i = 1, \ldots, N_T$; $j = 1, \ldots, N_T - 1$
Time difference between actual and effective greens	e

[#] Ψ refers to a set containing all traffic movements entering a signal-controlled junction.

3. Control variables

In the present formulation, the control variables can be defined as follows. Let $\Lambda = (\Lambda_I, \Lambda_B, \Lambda_C)$ be the set of control variables, where Λ_I is a subset of the integer variables, Λ_B is a subset of the binary variables, and Λ_C is a subset of continuous variables. Again, detailed diagrammatic explanations of the variables will be given in Section 4.

The subset, Λ_I, consists of the following integer variables

Number of approach lanes	$\alpha_i, \forall i = 1, \ldots, N_T$
Number of exit lanes	$\varepsilon_i, \forall i = 1, \ldots, N_T$

The subset, Λ_B, consists of the following binary variables

Permitted movements	$\Delta_{i,j,k}, \forall i = 1, \ldots, N_T; j = 1, \ldots, N_T - 1; k = 1, \ldots, \alpha_i$
Successor functions	$\Omega_{i,j,l,m}, \forall ((i,j), (l,m)) \in \Psi_s^*$

The subset, Λ_C, consists of the following continuous variables

Assigned lane flows	$q_{i,j,k}, \forall i = 1, \ldots, N_T; j = 1, \ldots, N_T - 1; k = 1, \ldots, \alpha_i$
Common flow multiplier	μ
Cycle length	ζ
Starts of green for turning movements	$\theta_{i,j}, \forall i = 1, \ldots, N_T; j = 1, \ldots, N_T - 1$
Durations of green for turning movements	$\phi_{i,j}, \forall i = 1, \ldots, N_T; j = 1, \ldots, N_T - 1$
Starts of green on approach traffic lanes	$\Theta_{i,k}, \forall i = 1, \ldots, N_T; k = 1, \ldots, \alpha_i$
Durations of green for approach traffic lanes	$\Phi_{i,k}, \forall i = 1, \ldots, N_T; k = 1, \ldots, \alpha_i$

* Ψ_s is a subset of Ψ referring to incompatible movements only.

4. Governing constraint sets

4.1. Total number of approach and exit lanes within arm

For a traffic arm i connecting to a junction, the total number of lanes L_i is equal to the sum of the number of approach lane α_i and the number of exit lane ε_i as in Eq. (1). In conventional designs, numbers of approach and exit lanes are pre-specified fixed parameters. With this set of constraints, the lane arrangements in different arms can be varied and optimised taking the actual demand patterns into consideration. For a heavy turning movement, a long green time of the associated signal phase is usually required to attain an acceptable degree of saturation. However, if a lane is reallocated to serve the movement of approaching traffic, the saturation flow (discharge capacity) will be increased so that only shorter green time is needed. The saved green time can then be re-assigned to serve other critical traffic streams. The design concept aims to improve the overall junction performance by optimising the lane configurations as well as the signal settings in a unified framework.

$$L_i = \alpha_i + \varepsilon_i, \qquad \forall i = 1, \ldots, N_T \tag{1}$$

Fig. 1 shows an example arm i with seven traffic lanes ($L_i = 7$). The configuration of the arm that is shown provides four approach and three exit lanes.

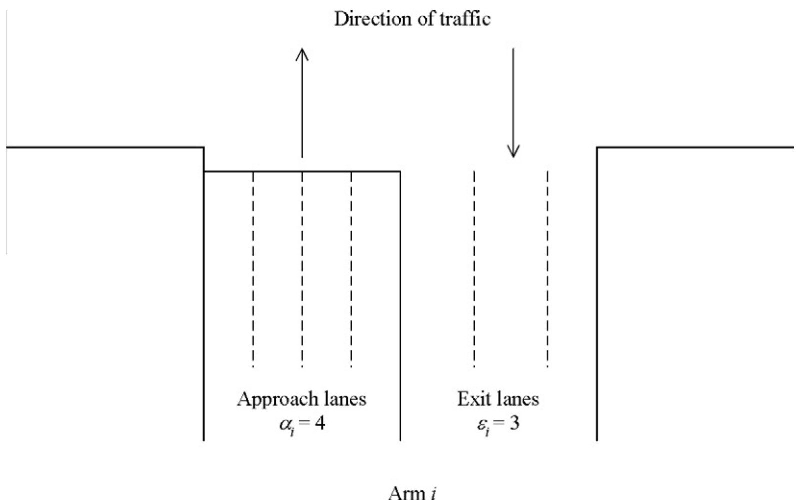

Fig. 1. Example of an arm configuration.

4.2. Conservation of assigned flows on approach lanes

In the present formulation, it is assumed that the demand turning flows $Q_{i,j}$ are known and given as inputs in the optimisation framework. Lane turning flows $q_{i,j,k}$ to be assigned on different approach lanes have been represented as variables in the present formulation. When a common flow multiplier μ is also introduced to scale the given demand flows, it is possible to maximise it and obtain the highest possible flow level in the optimisation framework to represent the junction capacity. The assigned turning flows on different approach lanes should therefore be governed by a set of flow conservation constraints as follows:

$$\mu Q_{i,j} = \sum_{k=1}^{\alpha_i} q_{i,j,k}, \qquad \forall i = 1, \ldots, N_T; \quad j = 1, \ldots, N_{T-1}, \tag{2}$$

where $Q_{i,j}$ is the given demand for turn j from arm i and $q_{i,j,k}$ is the traffic flow to be assigned on lane k for the traffic movement of turn j from arm i.

Based on the demand flows listed in Table 1, one possible flow pattern assigned on the two approach lanes is given in Fig. 2 if the common flow multiplier μ is 1.2. Because there is only one left-turn lane marking in the nearside lane, all the scaled left-turn flow has to be assigned on that lane. As there are straight-ahead lane markings on both of the approach lanes, the scaled straight-ahead demand flow can be assigned on both of the lanes as long as their sum is equal to 1200 (units). In the example arm, there is a shared lane marking on the nearside lane for both the left-turn and the straight-ahead movements. Thus, the two movements form one traffic stream. The total lane flows on the two approach lanes should be well balanced which are now 850 and 950 (units) to give the same flow factor. Further requirements under this circumstance will be discussed later in this section.

4.3. Minimum permitted movement on approach lanes

For each traffic arm connecting to a junction, there are approach and exit lanes as already shown in Fig. 1. It is expected that each approach lane should provide at least one lane marking to permit one traffic movement and ensure that every approach lane is in use. The following set of constraints is thus required.

$$\sum_{j=1}^{N_T-1} \Delta_{i,j,k} \geqslant 1, \qquad \forall i = 1, \ldots, N_T; \quad k = 1, \ldots, \alpha_i. \tag{3}$$

4.4. Minimum exit lanes in arms

For a traffic movement of turn j from arm i, the number of exit lanes of the corresponding exit arm as denoted by the function $\Gamma(i,j)$ should always be at least as great as the total number of lane markings assigned to permit such a turning

Table 1
Numerical example of demand flows.

Arm, i	Turn, j	Demand flow, $Q_{i,j}$	Scaled demand flow, $\mu Q_{i,j}$	Remarks
1	1	500	600	Left-turn
1	2	1000	1200	Straight-ahead

Note: Assume μ = 1.2.

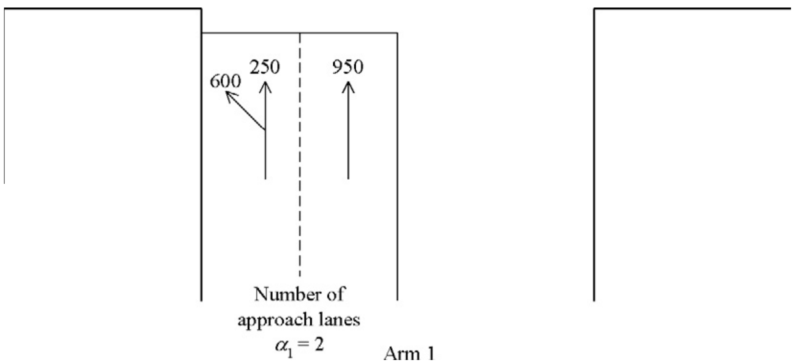

Fig. 2. Example of assigned flow conservation.

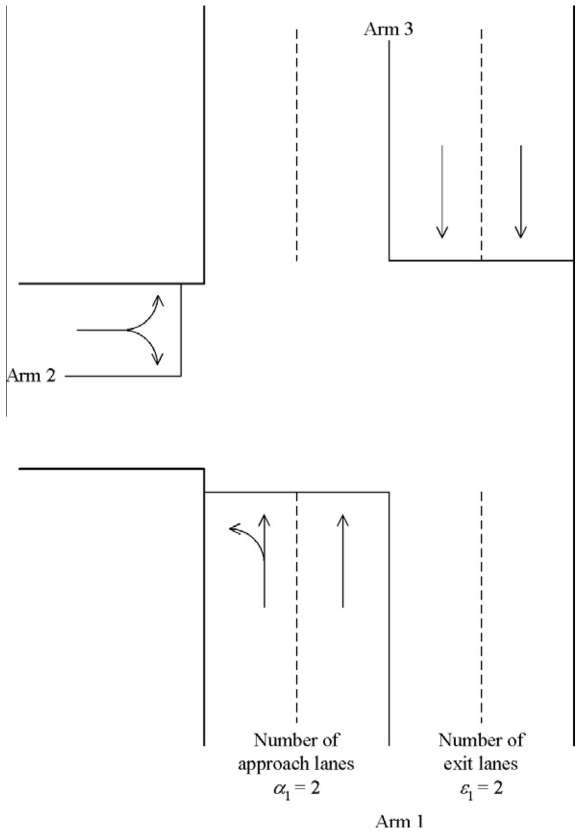

Fig. 3. Example junction layout.

Table 2
Model parameters of the example junction.

Arm, i	Turn, j	Lane, k	Lane marking, $\Delta_{i,j,k}$	Number of approach lane, α_i	Exit arm, $\Gamma(i,j)$	Number of exit lane, $\varepsilon_{\Gamma(i,j)}$
1	1	1	1	2	2	$\underline{\geqslant 1}$
1	2	1	1		3	$\underline{\geqslant 2}$
1	2	2	1		3	
2	1	1	1	1	3	$\geqslant 1$
2	2	1	1		1	$\geqslant 1$
3	1	1	1	2	1	$\underline{\geqslant 2}$
3	1	2	1		1	

Note: Underlined "$\underline{\geqslant 1}$" or "$\underline{\geqslant 2}$" are binding constraints for the corresponding exit arms.

movement from arm i. Due to safety and operational considerations, the provision of adequate number of exit lanes at downstream for all turning movements from all traffic arms is essential to prevent undesirable traffic merging activities. The following constraint set is required to make sure that the leaving traffic can depart smoothly from the junction.

$$\varepsilon_{\Gamma(i,j)} \geqslant \sum_{k=1}^{\alpha_i} \Delta_{i,j,k}, \qquad \forall i = 1,\ldots,N_T; \quad j = 1,\ldots,N_T - 1. \tag{4}$$

Fig. 3 shows an example junction layout and corresponding model parameters are tabulated in Table 2. Since there are two traffic lanes from arm 2 ($L_2 = 2$), one exit lane ($\varepsilon_2 = 1$) has to be given to serve the left-turn traffic from arm 1 and the remaining lane from arm 2 becomes the only approach lane ($\alpha_2 = 1$) for both the left- and right-turn traffic simultaneously through adopting a shared lane marking. The respective exit arms for the left- and right-turn traffic from arm 2 are arms 3 and 1. The required number of exit lane is equal to one for these two arms. However, there are two straight-ahead lane markings on each of arms 1 and 3 demanding at least two exit lanes respectively in arms 3 and 1. Thus, the effective minimum number of exit lanes on each of arms 1 and 3 must be two.

4.5. Permitted movements across adjacent approach lanes

For any two adjacent traffic lanes, k (left-hand) and $k+1$ (right-hand) lanes from arm i, if the traffic movement of turn j is permitted on lane $k+1$, then traffic movements of all other turns, $j+1, \ldots, N_T-1$, should be prohibited on lane k to eliminate potential internal-cross conflicts within an arm. This can be specified by the following constraint set.

$$1 - \Delta_{i,j,k+1} \geqslant \Delta_{i,m,k}, \qquad \forall i = 1,\ldots,N_T; \quad j = 1,\ldots,N_T-2; \quad m = j+1,\ldots,N_T-1; \quad k = 1,\ldots,\alpha_i - 1. \tag{5}$$

Given the binary nature of the lane marking variables, if $\Delta_{i,j,k+1} = 1$, then $\Delta_{i,m,k} = 0$ for all $m = j+1, \ldots, N_T-1$, meaning that the traffic movement (i, m, k) is prohibited. These movements have to be served by other available approach lanes (such as lanes $k+1, \ldots, \alpha_i$). However, if $\Delta_{i,j,k+1} = 0$, then $\Delta_{i,m,k}$ can take on either a value of 0 or 1.

In Fig. 4, an example arm is given for illustrations and corresponding model parameters are tabulated in Table 3. There are three approach lanes from arm 1 and it is assumed that the junction consists of four arms with $N_T = 4$. The nearside lane from arm 1 permits left-turn and straight-ahead movements. The middle lane allows straight-ahead and right-turn movements. The third (right-most) approach lane only serves the right-turn traffic. If the binary lane marking variable $\Delta_{i,j,k}$ is equal to "1" implying that the lane marking of turn j from arm i on lane k exists and "0" otherwise.

The establishment of a turn marking on an approach lane depends on the lane marking pattern on its immediate right-hand lane. In Table 3, it can be traced that the left-turn lane marking on lane 2 (middle lane) does not exist so that the straight-ahead lane marking on lane 1 (nearside lane) can exist, according to the constraint set. Since there is a straight-ahead lane marking assigned on lane 2, a right-turn on lane 1 is prohibited. And the permitted right-turn on lane 2 causes no further effect on lane 1 for the lane marking establishment. It just bans the right-turn on lane 1 so far. For lane 3, the lane markings of left-turn and straight-ahead movements do not exist so that all kinds of lane markings can be given on lane 2 including left-turn, straight-ahead, and right-turn movements. The given lane marking pattern in the example is a feasible set satisfying the requirements of the present constraint set. Still, it is worth noting that it is possible to have a left-turn lane marking on lane 2 due to the only right-turn lane marking on lane 3. If the left-turn lane marking is assigned on lane 2, then the constraint set must prohibit the straight-ahead movement on lane 1 due to the internal-cross conflict and only a left-turn lane marking can be given on lane 1.

4.6. Cycle length

Let the minimum and maximum cycle times in the junction be $c_{\min}$ and $c_{\max}$. Instead of defining the cycle time directly as the control variable, its reciprocal, $\zeta = 1/c$, is adopted to preserve the linearity in the mathematical formulation. The governing constraint on the cycle length can then be given by

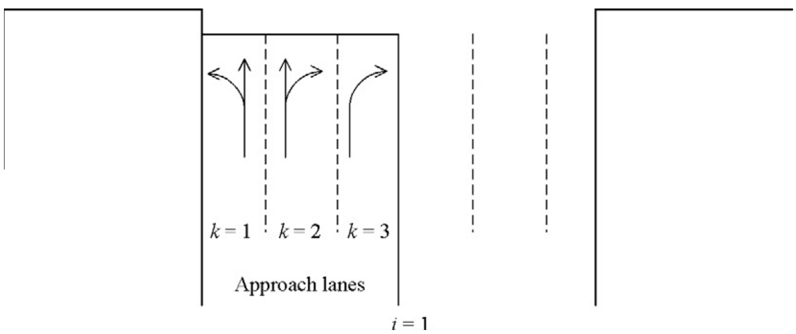

Fig. 4. Example of a junction arm.

Table 3
Model parameter of the example junction arm.

Arm, i	Turn, j	Lane $k+1$	Lane marking, $\Delta_{i,j,k+1}$	Turn, m	Lane marking, $\Delta_{i,m,k}$
1	1	2	0	2 3	0 or 1
1	2	2	1	3	0
1	3	2	1	Not defined	N.A.
1	1	3	0	2 3	0 or 1
1	2	3	0	3	0 or 1
1	3	3	1	Not defined	N.A.

$$\frac{1}{c_{\min}} \geqslant \zeta \geqslant \frac{1}{c_{\max}} \tag{6}$$

so that the cycle length will always be selected between the feasible range of the two extreme boundaries $c_{\min}$ and $c_{\max}$.

4.7. Lane signal settings

In practical signal control strategies, if an approach lane permits two or more turning movements through the provision of a shared lane marking, then the involved movements are generally controlled by an identical signal display to avoid ambiguity. Considering a lane k from arm i, if a turning movement j is permitted on this lane, then the following two constraint sets can be established to fulfil the above condition,

$$M(1-\Delta_{i,j,k}) \geqslant \Theta_{i,k} - \theta_{i,j} \geqslant -M(1-\Delta_{i,j,k}), \qquad \forall i = 1,\ldots,N_T; \quad j = 1,\ldots,N_T-1; \quad k = 1,\ldots,\alpha_i, \tag{7}$$

and

$$M(1-\Delta_{i,j,k}) \geqslant \Phi_{i,k} - \phi_{i,j} \geqslant -M(1-\Delta_{i,j,k}), \qquad \forall i = 1,\ldots,N_T; \quad j = 1,\ldots,N_T-1; \quad k = 1,\ldots,\alpha_i, \tag{8}$$

where M is an arbitrary large positive constant number. If a movement (i,j) is permitted on lane k, then the lane marking $\Delta_{i,j,k} = 1$, and hence the values on both sides of the above two inequalities become zero. This puts forward that $\Theta_{i,k} = \theta_{i,j}$ and $\Phi_{i,k} = \phi_{i,j}$ from the constraints. There are identical signal settings for all j that are permitted on lane k. However, if this movement is not permitted on the lane, then the constraints become ineffective because $(1-\Delta_{i,j,k})$ is equal to unity, and hence the signal settings could be different.

4.8. Start of green

As the signal settings at the junction are cyclical, the starts of green can be picked arbitrarily along the time axis as long as they satisfy other relevant constraints in the formulation. However, for convenience, all of the starts of green variables are confined within a signal cycle of a range between 0 and 1.

$$1 \geqslant \theta_{i,j} \geqslant 0, \qquad \forall i = 1,\ldots,N_T; \quad j = 1,\ldots,N_T-1. \tag{9}$$

4.9. Duration of green

The duration of green of a turning movement is subject to a minimum value for safety reasons so as to avoid sudden stop-and-go motion and provide sufficient times for vehicle drivers to prepare the changes of signal states. These constraints can be set as

$$1 \geqslant \phi_{i,j} \geqslant g_{i,j}\zeta, \qquad \forall i = 1,\ldots,N_T; \quad j = 1,\ldots,N_T-1, \tag{10}$$

where $g_{i,j}$ is a user specified minimum duration of green for a signal phase controlling the turning movement j from arm i.

4.10. Order of signal displays

Any two signal phases are said to be mutually incompatible if for some movement of turn j from arm i and movement of turn l from arm m movements j and l conflict with each other. The set of incompatible signal phases, Ψ_s, can therefore be derived from Ψ, which is the set of pairs of mutually incompatible movements. For any two incompatible signal phases (i,j) and (l,m) in Ψ_s, the order of signal displays is governed by the binary-type successor function (Heydecker, 1992), $\Omega_{i,j,l,m}$ (=0 or 1), where $\Omega_{i,j,l,m} = 0$ if the start of green of signal phase (l,m) follows that of signal phase (i,j), and =1 for the reverse order. Therefore, the following constraints can be set for the successor functions:

$$\Omega_{i,j,l,m} + \Omega_{l,m,i,j} = 1, \quad \forall((i,j),(l,m)) \in \Psi_s. \tag{11}$$

4.11. Clearance time

For any pair of incompatible traffic movements u from one arm and v from another arm, when their lane markings indicate that both movements are permitted, a clearance time (also known as intergreen) must be given in order to separate their rights-of-way in a signal cycle for safety reasons. Such intergreen period can be defined as: "the period between the end of the green signal giving right of way for one signal phase and the beginning of the green signal giving right of way for the next signal phase". It makes up of an amber period (usually 3 s) after one green period and a red-amber period (usually 2 s) before the next green period. The intergreen can still be fixed at a longer period (>5 s) depending on conflict distances according to the actual physical dimensions of the junction and even incorporate an all-red period. A short intergreen period is potentially dangerous although more effective green times can be available for allocating to traffic movements so as to increase the overall junction capacity. For a pair of mutually incompatible traffic movements

$u = (i, j, k)$ and $v = (l, m, n)$, then the following constraints can be set to put forward the clearance time requirements in the present formulation.

$$\theta_{l,m} + \Omega_{i,j,l,m} + M(2 - \Delta_{i,j,k} - \Delta_{l,m,n}) \geqslant \theta_{i,j} + \phi_{i,j} + \omega_{u,v}\zeta, \qquad \forall (u, v) \in \Psi \tag{12}$$

where M is an arbitrary large positive constant number and $\omega_{u,v}$ is a user specified minimum clearance time for the movements u and v. These constraints are effective only when the two mutually incompatible movements are permitted with $\Delta_{i,j,k} = \Delta_{l,m,n} = 1$.

4.12. Prohibited movement

If a traffic movement is prohibited meaning that the corresponding demand flow $Q_{i,j}$ is given to be zero, then the associated turn in form of a lane marking should not be existed on the approach lane i.e. $\Delta_{i,j,k} = 0$ according to Eq. (13). With $\Delta_{i,j,k} = 0$, the assigned lane flow $q_{i,j,k}$ will then be forced to be zero for consistency that can be achieved by the constraint set given in Eq. (14). For one-way streets, no other turning demand flows exist except the one moving to an immediate downstream direction that is likely regarded as a straight-ahead movement. Eq. (13) is able to prevent the establishment of other redundant lane markings (left-turn and/or right-turn) as long as the given demand flow is zero.

$$MQ_{i,j} \geqslant \sum_{k=1}^{\alpha_i} \Delta_{i,j,k}, \qquad \forall i = 1, \ldots, N_T; \quad j = 1, \ldots, N_T - 1, \tag{13}$$

and

$$M\Delta_{i,j,k} \geqslant q_{i,j,k} \geqslant 0, \qquad \forall i = 1, \ldots, N_T; \quad j = 1, \ldots, N_T - 1; \quad k = 1, \ldots, \alpha_i, \tag{14}$$

where M is again an arbitrary large positive constant number. In some practical cases, $\Delta_{i,j,k} = 0$ will be set specifically to ban a movement turn j from arm i to be moved on lane k implying that the movement turn on the lane is prohibited, then the lane flow again must be zero. However, if $\Delta_{i,j,k} = 1$ indicating that the movement is permitted, then the allocated flow can take on any non-negative value, as long as it satisfies other related constraints in the formulation.

4.13. Flow factor

As the way to distribute (demand) turning flows on different approach lanes should follow the queuing theory, it is required that the degrees of saturation must be identical for a pair of adjacent lanes having a common lane marking (permitted movement). Moreover, from the constraints that were set in Eqs. (7) and (8), the signal settings on this pair of adjacent lanes have been forced to be identical. Therefore, to ensure identical degrees of saturation, it suffices to equalise the flow factors (which are defined as the total lane flow divided by the lane saturation flow) of these adjacent lanes. Let $y_{i,k}$ be the flow factor of lane k from arm i, which can be expressed as

$$y_{i,k} = \frac{\sum_{j=1,\ldots,N_T-1} (q_{i,j,k}\tau_{i,j})}{s_{i,k}},$$

where $\tau_{i,j}$ is a user specified numerical factor to convert vehicular flow of turn j from arm i into through car unit (TCU) and $s_{i,k}$ is the lane saturation flow for straight-ahead movements of lane k from arm i. Therefore, the following set of constraints can be used to equalised the lane flow factors,

$$M(2 - \Delta_{i,j,k} - \Delta_{i,j,k+1}) \geqslant \left(\frac{\sum_{j=1}^{N_T-1} (q_{i,j,k}\tau_{i,j})}{s_{i,k}}\right) - \left(\frac{\sum_{j=1}^{N_T-1} (q_{i,j,k+1}\tau_{i,j})}{s_{i,k+1}}\right) \geqslant -M(2 - \Delta_{i,j,k} - \Delta_{i,j,k+1}),$$
$$\forall i = 1, \ldots, N_T; \quad j = 1, \ldots, N_T - 1; \quad k = 1, \ldots, \alpha_i, \tag{15}$$

where M is an arbitrary large positive constant number, and k (left) and $k + 1$ (right) are adjacent lanes from arm i. Again, these constraints are effective only when same lane markings exist on two adjacent lanes with $\Delta_{i,j,k} = \Delta_{i,j,k+1} = 1$.

4.14. Maximum acceptable degree of saturation

Let $p_{i,k}$ be the maximum permitted degree of saturation on lane k from arm i. For a traffic lane k from arm i, the degree of saturation can be expressed as

$$\bar{\rho}_{i,k} = \frac{y_{i,k}}{\Phi_{i,k} + e\zeta} \leqslant p_{i,k}, \qquad \forall i = 1, \ldots, N_T; \quad k = 1, \ldots, \alpha_i,$$

where $\bar{\rho}_{i,k}$ is the degree of saturation on lane k from arm i and e is the extra effective green time that devises from the difference between actual and effective greens (measured in time units, which are usually taken as 1 s). Together with the flow factor equation above, the following constraint set can be generated to ensure that the degree of saturation of every approach lane is no greater than the maximum acceptable limit.

$$\Phi_{i,k} + e\zeta \geqslant \frac{1}{p_{i,k}s_{i,k}} \sum_{j=1}^{N_T-1} (q_{i,j,k}), \qquad \forall i = 1, \ldots, N_T; \quad k = 1, \ldots, \alpha_i. \tag{16}$$

4.15. Optional signal phase constraints

For practical operations, it is sometimes useful to apply relative timing of starts and ends of greens for different signal phases. These constraints can be set as follows. Let $z^s_{i,j,l,m}$ and $z^e_{i,j,l,m}$ be the relative time for respectively the starts and ends of greens that are required to appear between signal phase (i, j) and signal phase (l, m). For the starts of green,

$$\theta_{i,j} + z^s_{i,j,l,m} = \theta_{l,m}, \tag{17}$$

and between the ends of green,

$$\theta_{i,j} + \phi_{i,j} + z^e_{i,j,l,m} = \theta_{l,m} + \phi_{l,m}. \tag{18}$$

These constraints are mainly used to constrain two signal phases to start and/or end simultaneously (i.e. setting $z^s_{i,j,l,m} = 0$ and/or $z^e_{i,j,l,m} = 0$) for the practicability of signal timing at a junction.

5. Optimisation of junction reserve capacity

One important design aspect of traffic signal settings is to maximise the operating capacity of a junction with the considerations of the requirements of geometric layout and signal timings. Based on the useful assumption that the traffic flows for the traffic movements in the junction will increase in proportion to the demand matrix (Allsop, 1972; Gallivan and Heydecker, 1988; Wong, 1996; Wong and Yang, 1997a,b; Wong et al., 2006, 2007), the problem becomes one of determining the largest common multiplier, $\mu_{\max}$, that can be accommodated without violating any of the constraints that are specified in the previous section. A value of $\mu_{\max} < 1$ then indicates that the junction is overloaded by $100(1 - \mu_{\max})$ percent, and a value of $\mu_{\max} > 1$ indicates a reserve capacity of $100(\mu_{\max} - 1)$ percent. The junction capacity maximisation problem is in turn to maximise the common flow multiplier μ and can be effectively formulated as a Binary-Mixed-Integer-Linear-Program (BMILP) as shown below.

$$\underset{\Lambda=(\Lambda_I,\Lambda_B,\Lambda_C)}{\text{Maximise}} \quad \mu$$

subject to the linear constraints in (1)–(18). The above mathematical program is linear in nature and can be solved effectively by standard branch-and-bound routines.

6. Solution algorithms

In this section, solution algorithm is outlined to solve the capacity maximisation in the present extended lane-based formulation for isolated signal-controlled junctions. Capacity maximisation can be effectively formulated as a Binary-Mixed-Integer-Linear-Program (BMILP) and standard branch-and-bound technique can be applied to solve for the global optimum solution. The BMILP problems for the capacity maximisation can be standardised into an integer-programming problem defined as follows:

$$\begin{aligned} &\text{Maximise} \quad \mathbf{d}^T \cdot \mathbf{x} \\ &\text{Subject to} \quad \mathbf{A}\mathbf{x} \leqslant \mathbf{b} \quad \text{and} \quad \mathbf{x} \geqslant 0 \\ &\qquad\qquad\quad \mathbf{x} \text{ contains both integer and continuous components.} \end{aligned}$$

where $\mathbf{x}$ is the variable vector collecting all the continuous and integer components. Binary variable is considered as a special type of integer variable with either "0" or "1" value. The vector of coefficients $\mathbf{d}$ multiplied by the vector of variables forms the objective function. $\mathbf{A}$ and $\mathbf{b}$ are the coefficient matrix and a constant vector for the problem constraints. The branch-and-bound technique starts by solving the program to maximise the objective function that is the common flow multiplier, μ, in the present problem involving mainly binary type variables. The solution algorithm is implemented to avoid explicit searching over the full enumeration tree containing all combinations of the integer variables. The strategies are to maintain the 0–1 restriction at the expense of retaining the feasibility of all linear constraints and also to prune the solution tree branches from a partial solution node containing some integer and binary variables with fixed values and some other free variables awaiting for determination. Initially, all variables are treated as free variables without fixed values and the free variables can be fixed by expanding the branches from a solution node. However, some of the constraints may be violated during this solution process making the solution infeasible to yield the optimal integer solution. Although this property is common in the set of binary variables, the objective function μ can still be maximised to give a bounding function value setting an upper solution bound for the actual optimal value of μ. No more branches are needed to grow further whenever (1) the set of the fixed binary variables violate any one of the governing constraint sets (no descendent node can be feasible no matter how the remaining free variables are set) or (2) the solutions obtained by expanding the descendent nodes cannot attain optimum (always below the best feasible lower bound or the incumbent). However, if the infeasibility is found due to the maximisation of the bounding function value by setting some free variables with values violating the governing constraint sets, then further subdivision branches are required to fix those free variables. By continuously branching the nodes within the enumeration tree, more integer and binary variable values can be determined to satisfy the governing constraint sets and more infeasible descendent branches can be identified and trimmed to reduce the problem size based on the feasibility testing. An optimal solution found in a sub-problem node, that is feasible in the full problem, is not necessarily globally optimal. It can be regarded as a lower solution bound to trim the rest of the solution tree. The best complete feasible solution found during the branching process is known as an incumbent that can also be a choice for the lower solution bound. Once a partial solution node has a bounding function value that is worse than the current incumbent, then that node can be pruned immediately. As the lower solution bound is a magic figure in solving the full problem, it should be replaced if a better one is found as a new incumbent. The search continues until either all necessary branching sub-problems have been solved or no further branching exists (Vanderbei, 2001; Bradley et al., 1977). A computer package called MPL with a CPLEX solver, (Maximal, 2002), is used to implement the branch-and-bound algorithm to solve the present BMILP capacity maximisation problems.

7. Numerical examples

In this section, a set of numerical examples is evaluated to demonstrate the effectiveness of the proposed algorithms. A four-arm junction ($N_T = 4$) as shown in Fig. 5 is studied. Four cases with different number of traffic lanes given in traffic arms are evaluated for maximising the junction capacity. The purpose of studying these cases is to demonstrate that there are sensitivities in the optimisation results, mainly the common flow multiplier, lane marking results and numbers of approach and exit lanes in different arms, with respect to the given junction size for analysis. Relevant input data and assumptions are given as follows.

Table 4 summarises the input assumptions of the four study cases. Each of them has the same number of arms in which a four-arm junction is analysed. Maximum cycle length is set to be 120.00 s in all cases. The maximum acceptable degrees of

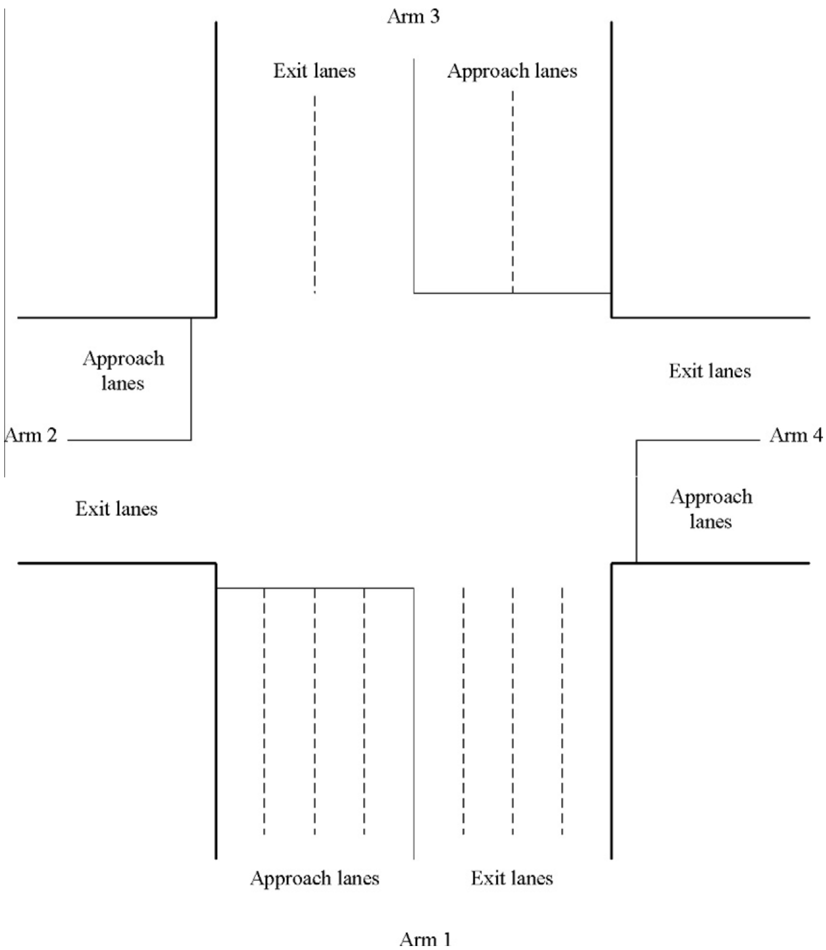

Fig. 5. Layout of the example four-arm junction.

saturation on different traffic lanes are all 90%. The effective green is always 1.0 s longer than the actual green for the design computation in all cases. Only the junction size varies in terms of the numbers of traffic lanes given in different arms. For each traffic lane, the saturation flow for straight-ahead movement (in tcu/h) is estimated by the formula below (Kimber et al., 1986).

$$s_{i,k} = 2080 - 140\delta_{i,k} + 100 \times (w - 3.25) - 42G$$

where w is the width of a lane in metres and G is slope of the arm in percentage (only for uphill reduction and no effects for downhill traffic). In the present example, it is assumed that all lane widths are 3.5 m and 0% in the gradients. The lane saturation flow should be 1965 tcu/h/lane and 2105 tcu/hr/lane for respectively the nearside ($\delta_{i,k} = 1$) and non-nearside ($\delta_{i,k} = 0$) lanes. Relevant traffic demand data $Q_{i,j}$ is tabulated in column 4 of Table 5 in pcu/h. To account for the effects of different turning movements, tcu conversion factors $\tau_{i,j}$ for the traffic demand of different turnings to the through car unit (tcu) is given in column 6 of Table 5. "1.6" and "1.4" are the numerical factors applied to all left- and right-turn traffic, respectively in view of the effect of their turning paths. And 1.0 will be applied for all straight-ahead movements. The minimum green durations $g_{i,j}$ are 5 s for all traffic movements. The minimum clearance time for a pair of conflicting movements (u, v) is listed in Table 6.

Table 7 summarises the general optimisation results including the maximised common flow multipliers, the corresponding optimised reserve capacities, the optimised reciprocal of cycle lengths, the number of integer variables defined in the problems and the required computation times of the four study cases. It can be observed that the four-arm junction is overloaded by about 6% when there are only four traffic lanes in each arm to serve the given demand traffic including the usages of approach and exit lanes. By adding one more lane in each arm, the junction performance improves from overloading to possessing 25.12% reserve capacity. It means that 25.12% more traffic based on the given demand flow matrix could enter the junction and the degrees of saturation of all approach traffic lanes are still under 90%. While extra lanes are adding to enlarge the junction size, the operational performance in terms of the reserve capacity keeps improving. There is 88.21% reserve capacity in Case 4 with seven lanes in each arm of the junction. When the design objective is to maximise the reserve capacity, the optimised cycle lengths are all binding at the maximum allowable limit that is 120.0 s for all the cases. It is well expected that more integer variables such as the lane marking variables are required to model a larger junction involving more traffic arms in the present mathematical formulation. The numbers of integer variables as well as the numbers of governing constraints increase linearly with the junction size. The more the integer variables, the higher the difficulties in solving the programming problems since more branches have to be investigated in the solution algorithms. And therefore longer computing time is usually required. It is still considered computationally affordable to spend around half an hour for a 7-lane 4-arm junction.

Fig. 6 plots the results of the optimal lane configurations for the junction capacity maximisation in the four study cases, (1) four traffic lanes, (2) five traffic lanes, (3) six traffic lanes, and (4) seven traffic lanes in all arms of the 4-arm junction. Case 1 in Fig. 6 shows the lane usages of the four traffic lanes in each arm of the junction. In the numerical example, there are left-turn, straight-ahead, and right-turn flows from each arm of the junction meaning that all three corresponding lane markings

Table 4
Input assumptions of the four test cases.

Case	N_T	$L_i \forall i = 1, \ldots, N_T$	$c_{\max}$ (s)	$p_{i,k} \forall i, k$ (%)	e (s)
1	4	4	120.00	90	1.0
2	4	5	120.00	90	1.0
3	4	6	120.00	90	1.0
4	4	7	120.00	90	1.0

Table 5
Traffic demand and minimum green time for the example four-arm junction.

Arm, i	Turn, j	Exit arm, $\Gamma(i,j)$	$Q_{i,j}$ (pcu/h)	$g_{i,j}$ (s)	$\tau_{i,j}$ (tcu/pcu)	Remarks
1	1	2	500	5.0	1.6	Left-turn
1	2	3	200	5.0	1.0	Straight-ahead
1	3	4	100	5.0	1.4	Right-turn
2	1	3	100	5.0	1.6	Left-turn
2	2	4	500	5.0	1.0	Straight-ahead
2	3	1	100	5.0	1.4	Right-turn
3	1	4	300	5.0	1.6	Left-turn
3	2	1	300	5.0	1.0	Straight-ahead
3	3	2	300	5.0	1.4	Right-turn
4	1	1	100	5.0	1.6	Left-turn
4	2	2	400	5.0	1.0	Straight-ahead
4	3	3	400	5.0	1.4	Right-turn

Table 6
Minimum clearance time matrix of conflicting movements (S).

Arm, i		1	1	1	2	2	2	3	3	3	4	4	4
	Turn, j	1	2	3	1	2	3	1	2	3	1	2	3
1	1									6		6	
1	2				6	6	6			6		6	6
1	3					6	6	6	6			6	6
2	1		6										6
2	2		6	6				6	6	6			6
2	3		6	6					6	6	6	6	
3	1			6		6							
3	2			6		6	6				6	6	6
3	3	6	6			6	6					6	6
4	1						6		6				
4	2	6	6	6			6		6	6			
4	3		6	6	6	6			6	6			

Table 7
Summary of optimisation results.

Case	$\mu_{\max}$	Optimised reserve capacity (%)	Optimised, ζ (s^{-1})	Integer variables	Computation time[a] (min)
1	0.9397	−6.03	1/120	353	1
2	1.2512	25.12	1/120	369	3
3	1.6795	67.95	1/120	385	18
4	1.8821	88.21	1/120	401	35

[a] Times were recorded from a Pentium Duo 2.0 processor.

must exist and the provision of one exit lane from all arms for the leaving traffic is a minimum: this leads to a design in which there is just one movement in each approach lane so that only one exit lane is required for each arm. In Case 1, four traffic lanes are available from all arms and it has been found in Table 7 that the junction is overloaded. The lane marking result shows that the optimal usage of the three remaining approach lanes in each arm is to assign one single lane marking on each of them permitting one single movement only. It can also be explained that adding one more lane marking on a non-critical lane to handle a critical movement is demanding one more exit lane from the corresponding exit arm. This forces a reduction in the number of the approach lane of that exit arm and therefore a shared lane usage has to be arranged. The optimisation results prove that it may not be a favourable design in the overloaded four-arm junction. Referring to Case 2 in Fig. 6, when one more lane is given in each arm, it will be adopted as an exit lane for most of the arms except from arm 1 using it as an approach lane. Simultaneously, more shared lane designs are observed so that the assigned lane flows could be distributed more evenly. Taking arm 4 as an example, the heaviest traffic are the straight-ahead and right-turn while comparing to the left-turn flow. Two additional lane markings for the straight-ahead and right-turn are added. The lightest left-turn is merged to the straight-ahead movement on the nearside lane through a shared lane marking. When the junction size is enlarged by adding new lanes in the arms, it is a decision to be made for their usages either as approach or exit lanes. In Case 3, most of the additional lanes become approach lanes so that some of the shared lane designs are removed. Numbers of the lane markings remain almost the same as in Case 2. In Case 4, a more balanced design is produced in which arms 1 and 4 provide five approach and two exit lanes while arms 2 and 3 offer four approach and three exit lanes. Comparing with Case 3, three more exit lanes and four more lane markings in total are generated from the design for the junction.

For further illustrations, a typical result format as given in Table 8 collecting all the optimised lane flow details and signal timings is shown and provided for evaluations of Case 4. Table 8 contains columns from (a) to (m) and the assigned lane flows are given in columns (c)–(f). If the numerical sum of the assigned lane flows are calculated vertically for each arm, there return the input demand flows for each turning movement in pcu/h. Column (g) gives the total lane flows for each approach lane in tcu/h with accounting for the associated conversion factors of different turning movements. Column (h) collects all the lane saturation flows with only one nearside lane and many non-nearside lanes from each arm based on the road widths and gradients. Flow factors in Column (i) are calculated with dividing the total lane flows in Column (g) by the lane saturation flows in Column (h). Column (j) is the times of the starts of green and Column (l) gives the times of the ends of green in the signal cycle of 120.0 s for the lanes and the associated traffic movements. The effective green times in Column (k) are just adding one second to the durations of green. The last Column (m) is the degrees of saturation of all approach lanes which can be regarded as the flow-to-capacity ratios and 0.9 (90%) is the allowable maximum limit. For those non-critical movements with degree of saturation below 0.9, their durations of green may not be fully maximised which can be obtained by conducting a post optimisation process. Whenever there is a shared lane marking given in Case 4 of Fig. 6, the respective approach lane should contain more than one assigned lane flow values. For those approach lanes involving the shared lane movements, the flow factors of them must be equal and should be considered as one single traffic stream receiving the same

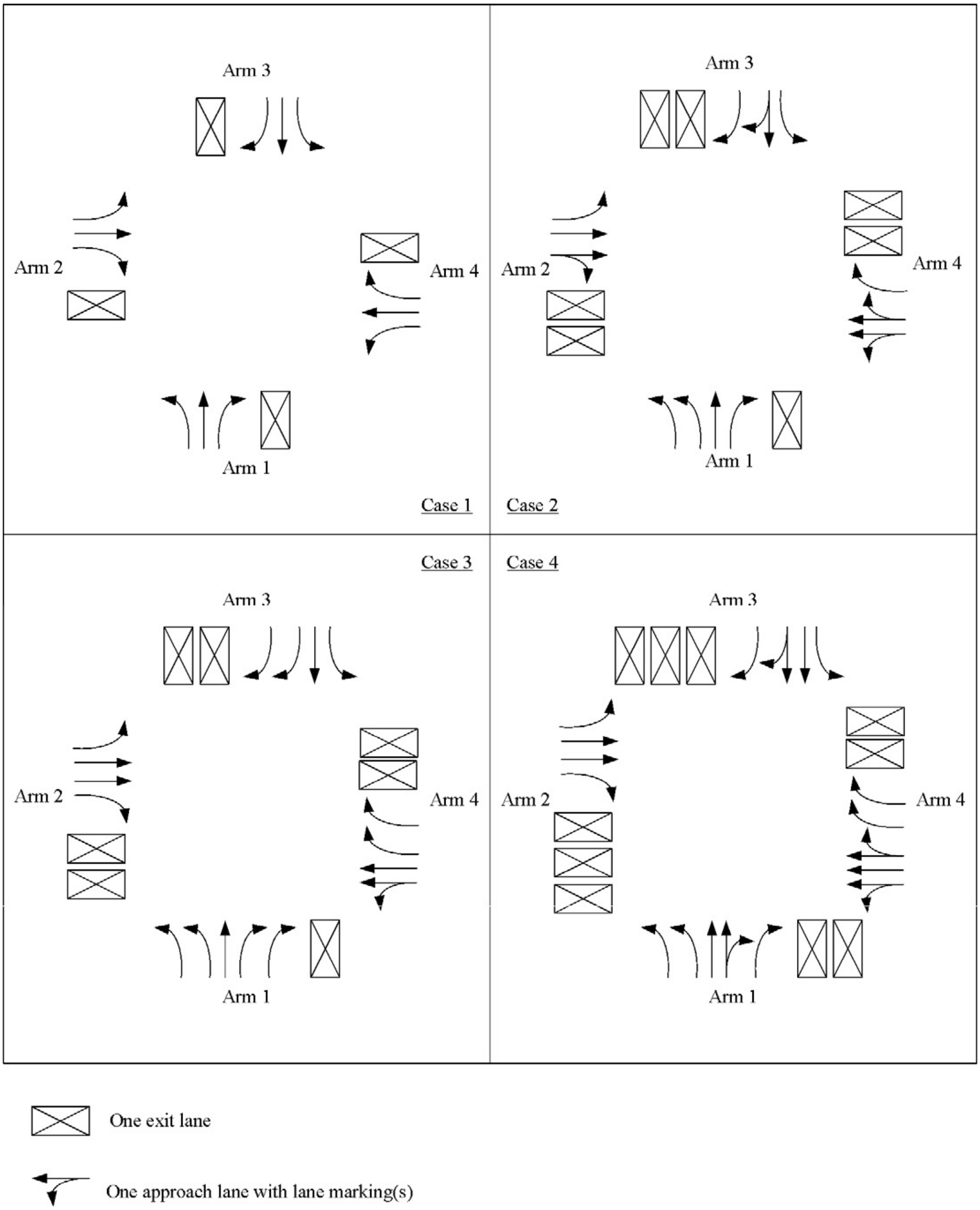

Fig. 6. Optimisation results of lane arrangements.

signal control settings including the starts and ends of green times. All these can be verified in Table 7. Arm 4 is a good example as two shared lane markings have combined the three turning movements to form one single traffic stream and all the flow factors and signal timings become identical.

To further illustrate the potential benefits in applying the proposed methodology for the practical design of a signal-controlled junction, it is essential to make comparisons with the performances and results from the general practice which is the manual design. The main feature of the proposed algorithms is to optimise signal settings as well as lane and turn allocations for the junction geometry in a unified optimisation framework automatically. Using the 4-arm 7-lane case (Case 4 in Section 7) as a reference for further investigations, there are 400 binary variables with around 2^{400} combinations in the solution space (actual size could be smaller due to elimination of some trivial infeasible settings). Manual trial and error solution process to test all feasible junction layouts to obtain optimal settings seems quite impossible. Currently, it refers to a common

Table 8
Optimisation results of lane flows and signal timings. Case 4: Seven traffic lanes for reserve capacity maximisation (μ = 1.8821 and cycle length = 120.0 s).

From Arm (a)	Lane (b)	To arm-assigned lane flow (pcu/h) (c)	(d)	(e)	(f)	Total lane flow (tcu/h) (g)	Saturation flow (tcu/h) (h)	Flow factor (i)	Start of green (s) (j)	Effective green (s) (k)	End of green (s) (l)	Degree of saturation (m)
		1	2	3	4							
1	1		241.40			386.24	1965.00	0.1966	0.00	49.33	48.33	0.9
1	2		258.60			413.76	2105.00	0.1966	0.00	49.33	48.33	0.9
1	3			113.33		113.33	2105.00	0.0538	0.00	13.51	12.51	0.9
1	4			86.67	19.05	113.33	2105.00	0.0538	0.00	13.51	12.51	0.9
1	5				80.95	113.33	2105.00	0.0538	0.00	13.51	12.51	0.9
2	1			100.00		160.00	1965.00	0.0814	18.51	64.43	81.94	0.3
2	2				250.00	250.00	2105.00	0.1188	18.51	30.19	47.70	0.9
2	3				250.00	250.00	2105.00	0.1188	18.51	30.19	47.70	0.9
2	4	100.00				140.00	2105.00	0.0665	18.51	30.82	48.33	0.9
3	1				300.00	480.00	1965.00	0.2443	53.70	61.30	114.00	0.9
3	2	240.00				240.00	2105.00	0.1140	54.33	28.61	81.94	0.9
3	3	60.00	128.57			240.00	2105.00	0.1140	54.33	28.61	81.94	0.9
3	4		171.43			240.00	2105.00	0.1140	54.33	28.61	81.94	0.9
4	1	100.00	51.92			211.92	1965.00	0.1078	87.94	27.06	114.00	0.9
4	2		227.02			227.02	2105.00	0.1078	87.94	27.06	114.00	0.9
4	3		121.06	75.69		227.02	2105.00	0.1078	87.94	27.06	114.00	0.9
4	4			162.16		227.02	2105.00	0.1078	87.94	27.06	114.00	0.9
4	5			162.16		227.02	2105.00	0.1078	87.94	27.06	114.00	0.9

Table 9
Optimised common flow multiplier μ under different lane and turn allocations.

Case	Lane configuration Arm A	Arm B	Arm C	Arm D	Maximised μ	Shared lane markings	% deviated from benchmark
I	5(2)	4(3)	4(3)	5(2)	<u>1.8821</u>	4	0.00
II	5(2)	4(3)	4(3)	5(2)	1.6795	0	−10.76
III	4(3)	4(3)	4(3)	4(3)	1.7386	5	−7.62
IV	4(3)	4(3)	4(3)	4(3)	1.6110	0	−14.40
V	4(3)	5(2)	5(2)	4(3)	1.8149	2	−3.57
VI	4(3)	5(2)	5(2)	4(3)	1.6192	0	−13.97
VII	5(2)	5(2)	5(2)	5(2)	1.8501	1	−1.70
VIII	5(2)	5(2)	5(2)	5(2)	1.8333	0	−2.59

5(2) – Approach lane number = 5 (exit lane number = 2); underlined – benchmark figure for comparisons.

practice that the manual design process will stop once the computed reserve capacity attains certain acceptable level say generally +20–30% and thus true optimum junction design can hardly be evaluated.

Using the Case 4 results in Section 7 as a benchmark (Case I in this section), 7 more relevant cases (II–VIII) in an attempt to imitate human designers' logic are also compiled in Table 9 for comparisons in which they all satisfy the stopping criteria in the manual design with more than 60% reserve capacity. For the same junction size, the optimised junction layout as listed in Case I can then be compared with other possible junction layouts. In Case II, retaining the optimal numbers of approach and exit lanes from various arms but removing all shared lane designs show poorer performances in terms of the maximised common flow multipliers (reserve capacity). Generally, designs with appropriate shared lane markings perform better when Cases I, III, V, and VII are compared with Cases II, IV, VI, and VIII, respectively. It should be noted that all of the shared lane designs in the present comparisons are optimised and cases with inefficient shared lane markings are not considered. Cases III and VII are of symmetrical junction layouts with fixed four and five approach lanes from all arms respectively and consistently worsen the overall junction performances even with optimised shared lane designs while comparing with the benchmark case. Junction geometry designs in Case V can be considered as an inappropriate one but also likely occur in manual designs by inexperienced engineers as it is a skewed mirror image with respect to the layout given in the optimised Case I. Again, overall performances are lowered even with (Case V) or without (Case VI) shared lane designs comparing with the benchmark case. Although there are still numerous feasible junction layout designs that cannot be covered in the present comparisons including all other symmetrical and asymmetrical lane arrangements, a general trend should be observed that manual designs in lane and turn allocations for signal-controlled junctions involving traffic stream groupings and complex interactions among arms are not easy tasks and very tedious to incorporate even small changes for testing purposes and hardly accomplishing the optimal settings which can readily be completed by the proposed algorithms.

8. Conclusions

In this paper, an extension of the lane-based design method for the reserve capacity optimisation of isolated signal-controlled junctions is given. New integer variables are defined to represent the numbers of approach and exit lanes in various arms of a junction and introduced into the optimisation framework which can thus be optimised with the lane markings, lane flows, and signal timings in a unified framework. Reserve capacity of a junction is maximised through maximising the common flow multiplier that is formulated as a Binary-Mixed-Integer-Linear-Programming (BMILP) problem and is solved by a standard branch-and-bound routine. Indeed, the cycle length minimisation problem is a similar kind problem that can be formulated to maximise the reciprocal of the cycle length and solved by the same solution algorithm. Pedestrian phases can also be added in the present formulation and treated as special traffic movements subject to signal timing constraints only without other flow related constraints such as the flow factor constraints. Designs of a four-arm junction with different numbers of traffic lanes in various arms are conducted numerically as a demonstration of the proposed algorithms. Detailed explanations and discussions of the modelling results including the junction geometries, individual lane usages, and signal timings have been given. Computational requirements in solving the design problem increase with the junction sizes as more integer variables are needed implying that more branches and bounds have to be considered in the solution algorithms but are still manageable in practice.

Acknowledgements

The work that is described in this paper has been carried out with the support of The Croucher Foundation Fellowship (2006–2008) and from the Research Grants Council of the Hong Kong Special Administrative Region, China (9041157 and 9041261).

References

Allsop, R.E., 1972. Estimating the traffic capacity of a signalized road junction. Transportation Research 6 (3), 245–255.
Allsop, R.E., 1975. Computer program SIGCAP for assessing the traffic capacity of signal-controlled road junctions – description and manual for users. Transportation Operations Research Group Working Paper, 11, University of Newcastle upon Tyne.
Allsop, R.E., 1976. SIGCAP: a computer program for assessing the traffic capacity of signal-controlled road junctions. Traffic Engineering and Control 17 (8–9), 338–341.
Allsop, R.E., 1992. Evolving application of mathematical optimisation in design and operation of individual signal-controlled road junctions. In: Griffiths, J.D. (Ed.), Mathematics in Transport and Planning and Control. Clarendon Press, Oxford, pp. 1–24.
Bradley, S.P., Hax, A.C., Magnanti, T.L., 1977. Applied Mathematical Programming. Addison-Wesley.
Gallivan, S., Heydecker, B.G., 1988. Optimising the control performance of traffic signals at a single junction. Transportation Research Part B 22 (5), 357–370.
Heydecker, B.G., 1992. Sequencing of traffic signals. In: Griffiths, J.D. (Ed.), Mathematics in Transport and Planning and Control. Clarendon Press, Oxford, pp. 57–67.
Improta, G., Cantarella, G.E., 1984. Control system design for an individual signalized junction. Transportation Research Part B 18 (2), 147–167.
Kimber, R.M., McDonald, M., Hounsell, N., 1986. The prediction of saturation flows for road junctions controlled by traffic signals. TRRL Report, RR 67. Transport and Road Research Laboratory, Crowthorne.
Maximal Software, Inc., 2002. User Manual, MPL Modeling System Release 4.2.
Tully, I.M.S.N.Z., 1976. Synthesis of sequences for traffic signal controllers using techniques of the theory of graphs. PhD thesis, OUEL Report, 1189/77, University of Oxford.
Vanderbei, R.J., 2001. Linear Programming: Foundations and Extensions. Kluwer Academic.
Webster, F.V., 1958. Traffic signal settings. Road Research Technical Paper, No. 39, HMSO, London.
Webster, F.V., Cobbe, B.M., 1966. Traffic signals. Ministry of Transport, Road Research Technical Paper, No. 56.
Wong, C.K., Wong, S.C., 2003a. Lane-based optimization of signal timings for isolated junctions. Transportation Research Part B 37 (1), 63–84.
Wong, C.K., Wong, S.C., 2003b. Lane-based optimization of traffic equilibrium settings for area traffic control. Journal of Advanced Transport 36 (3), 349–386.
Wong, C.K., Wong, S.C., Lo, H.K., 2007. Reserve capacity of a signal-controlled network considering the effect of physical queuing. In: Allsop, R.E., Bell, M.G.H., Heydecker, B.G. (Eds.), Transportation and Traffic Theory 2007. Elsevier, pp. 533–553.
Wong, C.K., Wong, S.C., Tong, C.O., 2006. Lane-based optimization method for multi-period analysis of isolated signal control junctions. Transportmetrica 2 (1), 53–85.
Wong, S.C., 1996. On the reserve capacities of priority junctions and roundabouts. Transportation Research Part B 30 (6), 441–453.
Wong, S.C., Yang, H., 1997a. The estimation of reserve capacity in traffic control. Hong Kong Institution of Engineers Transactions 4 (1), 21–30.
Wong, S.C., Yang, H., 1997b. Reserve capacity for signal-controlled network. Transportation Research Part B 31 (5), 397–402.