Optimizing Traffic Signal Settings for Public Transport Priority[*]

Robert Scheffler[1] and Martin Strehler[2]

1   Brandenburgische Technische Universität, Cottbus, Germany
    robert.scheffler@b-tu.de
2   Brandenburgische Technische Universität, Cottbus, Germany
    martin.strehler@b-tu.de

— Abstract —

In order to promote public transport many municipalities use traffic signal control with a priority for buses or trams. In this paper, we address the problem of finding optimal passive transit signal priority settings. Building on a cyclically time-expanded network model for the combined traffic assignment traffic signal coordination problem, we introduce a suitable queuing model and several modifications to model public transport vehicles appropriately. We evaluate the applicability of this approach by computing and analyzing optimal solutions for several instances of a real-world scenario.

1   Motivation and literature overview

The success of public transport depends on many factors. Certainly, one of them is travel time. Consequently, many city administrations try to accelerate public transport by bus lanes, general right-of-way for trams, or adaption of traffic signal settings. Focusing on the latter, there are mainly two strategies of transit signal priority (TSP). In passive priority, traffic lights along bus lines simply get more green time whether or not a bus is approaching. Active strategies try to detect buses or trams via sensors or the vehicles register themselves via radio, such that signal settings can be adapted on demand.

However, changing signal settings interferes with the optimization of traffic signals for the other road users. Moreover, bus lines also cross each other. Thus, these strategies are not applicable at every intersection or every situation since one cannot increase green times of the crossing lines at the same time. Even worse, one may not exclude the possibility that transit signal priority disturbs a perfectly optimized signal coordination in a city leading to heavy traffic congestion. Eventually, buses and trams may suffer from waiting times in the bumper-to-bumper traffic caused by the system which was originally intended to privilege them.

In this paper, we develop a model and several strategies to optimize traffic signal settings for public transport and the individual traffic simultaneously. The model is based on a cyclically-time expanded network model which was already shown to produce very good

solutions for the simultaneous *traffic assignment traffic signal coordination problem* in urban road networks. This model is extended by a new queuing model to capture the impact of traffic jams on public transport travel times. Afterwards, we discuss both approaches of integrating bus lines in an already optimized coordination and of optimizing signal settings for car and bus commodities simultaneously. This leads to a passive traffic signal priority which not only considers traffic on a single bus line but also takes the network wide impact on traffic congestion into account.

Literature overview

Several cities use priority traffic signal regulations for public transport, also known as *transit signal priority* (TSP). As already mentioned, there are two main concepts: *Passive transit signal priority* simply tries to improve the ease of traffic along the bus lines by a generally increased green time. It is therefore easy and cheap to implement. The Traffic Signal Timing Manual [9] subclassifies *active transit signal priority* in two different concepts depending on the arrival time of the public transport vehicle. Arrival during the green phase is supported by *green extension* to remove the queue in front of the bus and to guarantee passing of the intersection within the same cycle. *Red truncation* is used to end the red phase ahead of schedule when otherwise the bus would arrive at a red signal. Both approaches rely on detecting buses or trams beforehand which requires additional technical installations.

Despite the widespread use of TSP, there are only few studies or optimization approaches attacking the integrated traffic assignment traffic signal coordination problem for the entire network. There are some analytic studies that consider the impact of (active) transit signal priority on a single intersection [1, 12, 17]. Although these studies try to measure the overall impact of the prioritization to the network traffic, they often do not take into consideration that coordinated signals lead to a clustering of traffic, so called platoons of cars. The downstream impact of traffic signal priority, especially with respect to these platoons, is usually not captured by these analytic models. Moreover, the approaches above often implicitly assume that the exceptional traffic situation can be resolved within two cycles and they do not consider a spreading of the disturbance to surrounding intersections due to queues and high traffic volumes also of crossing traffic streams. Furthermore, the route choice of drivers experiencing delay due to transit signal priority is not covered by these models either, since they assume the routes to be fixed.

There is also a number of simulation based studies considering the impact of transit signal priority, e.g., [13, 18]. These approaches are much more realistic and can both capture the downstream effects and the route choices of the drivers. Thus they often give a very realistic picture of the outcome of TSP. Ngan et al. [13] confirmed the assumption that TSP is especially useful with less traffic on crossing roads, whereas traffic on the same route as the bus may significantly benefit from the transit priority. Wahlstedt [18] analyzed PRIBUSS, the standard implementation of TSP in most Swedish signal controllers. In a case study of the city center of Stockholm he showed that TSP could not only have a negative impact on the delay of the crossings streets, but also on the streets that are used by public transport. However, the problem here is that these simulation methods usually are very hard to capture by strict mathematical models. As a consequence, one can empirically measure the effect of a certain priority rule but one can hardly optimize this decision with mathematical optimization tools.

From a practical point of view, there are some approaches that integrate *Automatic Vehicle Location (AVL)* systems into TSP [4]. Using this concept, it is possible to only prioritize transit vehicles that are behind their schedule. Liu et al. [11] presented a dynamic approach for TSP which considers real-time traffic flow conditions.

There is only a comparatively small number of publications that consider passive TSP. Brilon and Laubert [2] highlight that a public transport friendly pretimed signal setting is an important basis for public transport priority. Here, public transport friendly coincides with a signal setting that causes less congestion for parallel traffic. For that reason, they suggest that traffic signal optimization algorithms should consider the stops and transfer times of public transport. Skabardonis [15] provides such an extension for public transport in the signal optimization tool TRANSYT.

A sound overview of transit signal priority control is given by Lin et al. [10]. In their conclusion, the authors formulate (among others) the following most urgent research questions. *Firstly, within coordinated networks, how can one analyze the impact of TSP on the overall coordination? Secondly, TSP only uses slight modifications of the existing base plan. Can bus priority already be considered in the design of these base plans?*

Successful traffic signal coordination very much relies on the idea of constantly repeating a certain set of control patterns for the different traffic intersections. This framework allows for optimization approaches to minimize overall travel time, waiting time or other suitable objectives. Recently, we presented a cyclically time-expanded network flow model to optimize traffic signals and traffic assignment simultaneously [6]. The properties of the model [7] allow applying mixed-integer linear programming techniques. That is, as a main advantage compared to other approaches, the optimization process also yields dual solutions which bound the gap towards the global optimum and which prove optimality of a primal solution. Furthermore, we presented approaches for more realistic queuing and spill-back in [14].

Our contribution

In this paper, we discuss how the special requirements of public transport can be handled when optimizing a pretimed traffic signal control. We give an overview how to implement aspects like different types of bus stops, timetables, queues, and bus priority. Furthermore, we study properties of this new approach and we optimize the traffic assignment simultaneously to examine the impact of passive transit signal priority on the route choice of other road users.
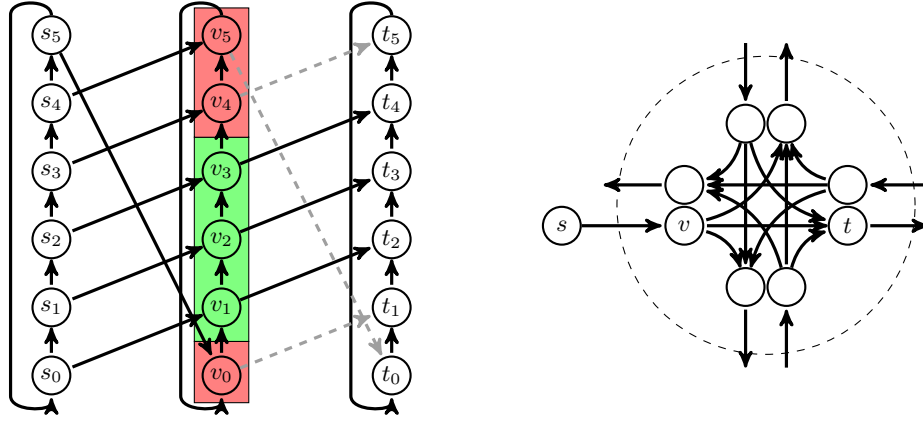
This paper is organized as follows. In Section 2, we present the basic concept of our model for traffic light optimization. In Section 3, we discuss the extensions to integrate public transport in our model. Since these changes have different consequences on the feasibility and optimality of the related mixed integer program, we study the properties of this new model in Section 4. Afterwards, we present the computational results for a real-world scenario in Section 5.

2 Basic model

Flows in networks with traffic signals obviously require a dynamic treatment, since flow values change over time. Already Ford and Fulkerson introduced time-expanded networks in their seminal work about flows more than 60 years ago [3] to cope with such time-dependent behavior. However, since the time horizon determines the size of such networks, this and similar approaches almost directly yield models of pseudo-polynomial size.

Disregarding the individual car and viewing flow in total, inner-city traffic with traffic signals is dominated by a very high periodicity. Therefore, we define a cyclically time-expanded network.

▶ **Definition 1** (Cyclically time-expanded network [6]). Let $G = (V, A, u)$ be a network with capacities $u : A \to \mathbb{N}$ and non-negative transit times t_e for each $e \in A$. For a given cycle

Figure 1 A very simple example of a cyclically time-expanded network G^6 with three consecutive nodes s, v, and t and six time steps. At node v, a traffic signal and a sample timing together with corresponding capacities of the outgoing arcs (implemented with binary variables: dashed $b_i = 0$, standard $b_i = 1$) are shown. On the right-hand side, a standard approach of handling different turning directions with separate arcs in the original network G is shown.

time Γ and a given number k of time steps of length $t = \frac{\Gamma}{k}$, the corresponding *cyclically time-expanded network* $G^k = (V^k, A^k, u^k)$ is constructed as follows.

- For each node $v \in V$, we create k copies $v_0, v_1, \ldots, v_{k-1}$, thus $V^k = \{v_t | v \in V, t \in \{0, \ldots, k-1\}\}$.
- For each link $e = (v, w) \in A$, we create k copies $e_0, e_1, \ldots, e_{k-1}$ in A^k where arc e_t connects node v_t to node $w_{(t + \lfloor \frac{t_e k}{\Gamma} \rfloor) \mod k}$. These arcs are called *transit arcs* and e_t has capacity $u(e_t) := \frac{u(e)}{k}$ and cost t_e.
- Additionally, we add *waiting arcs* from v_t to v_{t+1} $\forall v \in V$ and $\forall t \in \{0, \ldots, k-2\}$ and from v_{k-1} to v_0 $\forall v \in V$ with cost $\frac{\Gamma}{k}$ and infinite capacity to A^k.

In general, we assume all times (transit times, cycle times, signal parameters, ...) to be integral. Γ is chosen as the least common multiple of all cycle times of signals in the road network. Throughout this paper, we choose $k = \Gamma$, i.e., one time step is exactly one second long. An example of such a cyclically time-expanded network is presented in Figure 1.

Commodities are expanded similarly. Let $\Phi \subset V \times V \times \mathbb{R}^+$ be the set of commodities consisting of triples $\varphi = (s, t, d) \in \Phi$ of a source (or origin) s, a sink (or destination) t, and demand d in the original network. Here, demand denotes the amount of flow starting during one cycle, i.e., it is scaled to Γ. For simplicity, traffic demand is uniformly distributed over all copies of the original source. In other words, the net outflow of each s_i of commodity φ is $\frac{d}{k}$. Note that flow may also directly use a waiting arc from s_i to s_{i+1}. In contrast, the net inflow of each sink is not fixed. Flow of commodity φ may leave the network at any time-expanded copy t_i of the sink t. All together, this allows to formulate a standard multi-commodity network flow $f_\varphi : A^k \to \mathbb{R}_{\geq 0}$ $\forall \varphi \in \Phi$ in the cyclically-time-expanded network G^k. As usual, flow has to meet flow conservation at each node of G^k for each commodity. Capacity constraints apply in total over all commodities, i.e., $\sum_{\varphi \in \Phi} f_\varphi(e_t) \leq u(e_t)$ for all $e_t \in A^k$.

Traffic signals can now be implemented with the help of variable capacities. To achieve this, each turning direction at an intersection in G (cf. Figure 1) is modeled by an arc. Let $e \in A$ be an arc of an intersection with capacity c, where flow should be controlled by a traffic light. For the arcs $e_0, \ldots, e_{k-1}$ in the cyclically time-expanded network G^k, we use binary variables $b_0, \ldots, b_{k-1} \in \{0, 1\}$ and set the capacity of e_i to $\frac{u(e)}{k} b_i$. In other words, a binary variable b_i set to 1 corresponds to a green traffic light and the normal capacity at

time step i is available. A binary variable set to 0 corresponds to a red signal and flow may not use this arc at time step i. A new set of binary variables b_i^{on}, $i \in \{0, \ldots, k-1\}$, with constraints $\sum_{i=0}^{k-1} b_i^{\mathrm{on}} = 1$ and $b_{i-1} \geq b_i - b_i^{\mathrm{on}}$ for all $i \in \{0, \ldots, k-1\}$ guarantees that the signal switches from red to green at most once per cycle. For two arcs e_1 and e_2, the linear constraints $b_{1,i} + b_{2,i} \leq 1 \ \forall i \in \{0, \ldots, k-1\}$ prohibit green at the same time. Several other requirements like minimum and maximum green times can be modeled similarly [8].

Now, this model allows the simultaneous optimization of traffic assignment (multi-commodity flow) and traffic signal timings (coordination). We use the total travel time of all road users as objective function, that is, we are looking for a minimum cost flow. In the context of traffic signals, the variables b_i^{on} correspond to the *offset* of the light, i.e., the point in time when the signal switches to green. *Split times* refer to the lengths of the green periods $\sum_{i=0}^{k-1} b_i$. *Phase order* describes the sequence of turning directions which get green and is also determined by the variables b_i^{on} for the respective arcs. Split times and phase order are crucial for traffic safety and have to be chosen carefully. Thus, many applications only permit optimization of offsets. But with all other parameters fixed, choosing the offset of a single light already determines the signal setting of the whole intersection. This canonically defines the *offset of an intersection* in such cases. Consequently, we refer to *offset optimization* if we shift the whole signal plans of intersections in time, but the internal switching patterns at each intersection are unaltered.

Since the model is linear, we can use exact mathematical programming techniques and solvers like CPLEX or GUROBI benefiting from all the advantages, e.g., proof of optimality or dual bounds. Although the model is a linear one, it still provides very realistic travel times and link performance functions, e.g., the raise in the travel times is non-linear in relation to an increasing traffic demand [8].

3   Integrating public transport

In this section, we are going to integrate public transport commodities in our model. Here, we face three major challenges. Firstly, up to now, the cyclically time-expanded model does not provide any first-in first-out property in its queues. Whereas this was of minor importance for routing a huge number of cars and finding a system optimal coordination, we now have to consider the exact delay caused by queuing for public transport. Secondly, due to bus stops, public transport commodities experience other travel times in the road network. Moreover, in the basic model, we used arbitrarily splittable, i.e., non-atomic, commodities. In contrast, a bus or tram should use a single path in the time expanded network, that is, it is an unsplittable or atomic commodity, respectively. Thirdly, we have to define suitable objectives which correspond to our intuitive idea of transit signal priority.

Other aspects of public transport are easier to model. In general, we assume that public transport has no routing options (in space). That is, line planning is completed and the bus is fixed to one route. Still, we have to route over time in the cyclically time-expanded network. Thus, public transport is modeled like a common commodity, but a commodity-specific capacity is set to zero on all non-route arcs. In the following, we will consider only one public transport commodity and we will use f_{bus} to distinguish flow of this commodity φ_{bus} from flow f_φ from the standard commodities $\varphi \in \Phi$. Furthermore, a public transport commodity also has a source s and a sink t, but contrary to standard commodities, we assign the whole demand (usually $d = 1$, i.e., one bus) to one copy s_i as net outflow.

Queues

Flow on waiting arcs can be interpreted as a queue and a capacity on waiting arcs limits the length of this queue. Hence, even effects like spill-back may occur in the cyclically time-expanded model. A detailed study of queues in this model was published in [14]. Unfortunately, these queues do not fulfill the *first-in-first-out* property, that is, waiting flow will be shuffled in general.

Yet, considering only a single bus which may transport several dozen passengers and transit signal priority, it is important to know the exact delay of this bus. Therefore, we introduce a new waiting arc model. The main idea is to split the queue into three parts: in the cars waiting in front of the bus (*queue 1*), the *bus* itself, and the cars waiting behind the bus (*queue 2*). Thus, we use three parallel waiting arcs instead of one.

▶ **Definition 2.** Let v be a node in the original network. In the cyclically-time expanded network, *bus queuing waiting arcs* at v are built by the following construction:

- Instead of copies v_i as in Definition 1, insert four nodes w_i, x_i, y_i, and z_i for each time step $i \in \{0, \ldots, k-1\}$.
- Connect the time-expanded copies of incoming arcs at v to the corresponding copies w_i and copies of outgoing arcs of v to z_i.
- Add arcs (w_i, x_i), (x_i, y_i), and (y_i, z_i) with capacity of the summed up capacities of the upstream arcs ending in v and zero travel time.
- Add arcs $(x_i, w_{i+1 \mod k})$, $(y_i, x_{i+1 \mod k})$, and $(z_i, y_{i+1 \mod k})$ with travel time $\frac{\Gamma}{k}$ and infinite capacity.

In Figure 2, the construction of the new waiting arcs is shown in principle. Of course, this is again realized in a cyclic matter. The zigzag arcs $(x_i, w_{i+1 \mod k})$ and $(z_i, y_{i+1 \mod k})$ are only usable by standard commodities $\varphi \in \Phi$. The dotted arcs $(y_i, x_{i+1 \mod k})$ are exclusive for the public transport commodity φ_{bus}. This can be realized by simply not defining flow variables for commodities on arcs which should not be used by these commodities.

Moreover, we require flow f_{bus} on the dotted arcs $(y_i, x_{i+1 \mod k})$ for public transport to be integral. Since the tail nodes y_i of these dotted waiting arcs are the only chance for the public transport commodity to split, this implies integral flow values for public transport on all arcs of the network. Hence, this constraint yields an atomic flow for public transport.
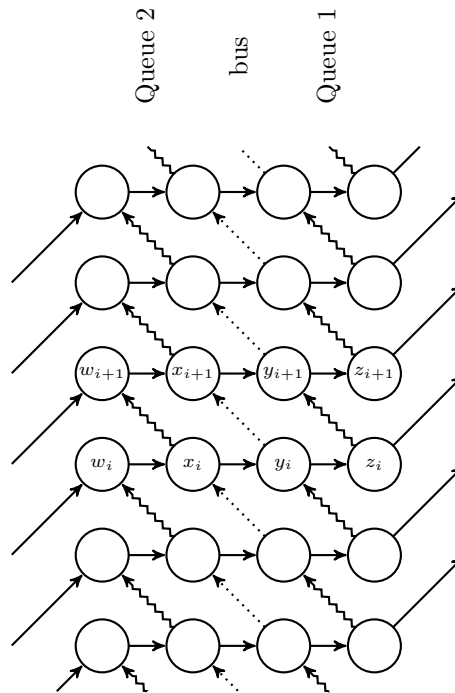
▶ **Lemma 3.** *Integral flow values on the bus queuing waiting arcs as given in Figure 2 for public transport commodities with integral demand imply integral flow values of public transport commodities on every arc in the network.*

Now, it is crucial that public transport flow cannot leave its waiting arcs in the bus queue when there is still flow on the waiting arcs in queue 1. This can be achieved by setting a suitable capacity constraint on all horizontal arcs $e = (w_i, x_i)$, $e = (x_i, y_i)$, and $e = (y_i, z_i)$ in the queue construction, namely

$$\sum_{\varphi \in \Phi} f_\varphi(e) + \frac{c}{k} f_{\text{bus}}(e) \leq \frac{c}{k}$$

where c is the capacity of arc e. As a consequence, flow f_{bus} of the public transport commodity on the horizontal arcs has to be zero as long as there is flow from any other commodity $\varphi \in \Phi$ on these arcs, since flow values of public transport are integral. Especially, flow from a public transport commodity φ_{bus} cannot pass queue 1 if it is not empty.

Additionally, the same argument applies for queue 2 of cars behind the public transport vehicle. Flow f_{bus} cannot pass queue 2 if it is not empty. Thus, flow from queue 2 is forced to

Figure 2 A waiting arc construction that allows FIFO for public transport. Dotted arcs can only be entered by public transport commodities and flow on these arcs has to be integral. Zigzag arcs can only be entered by standard commodities. Horizontal arcs have capacity for either one public transport vehicle or an arbitrary amount of flow of standard commodities.

switch to queue 1, before flow from the public transport commodity can enter its specific bus queue. Vice versa, flow from the standard commodities cannot pass flow φ_{bus} in its queue.

▶ **Theorem 4.** *The bus queuing waiting arc construction in Figure 2 realizes the first-in first-out property for flow of the public transport commodity.*

As described above, this approach is only applicable for a single public transport commodity. If we want to route more than one bus commodity over an arc, we need a bus queue for each of these public transport commodities and additionally, we also need a queue for the standard commodities between each consecutive pair of such bus queues. Please note that this approach only assures the FIFO principle between public transport and individual traffic. Individual traffic may still overtake each other within its queues.

Bus stops

The main difference between public transport and individual traffic are regular stops for boarding passengers, also resulting in longer travel times of the public transport vehicles. To model such stops appropriately, we also have to consider whether the standard traffic is affected by these stops.

In the case of a bus bay, traffic can just bypass the bus. To implement such stops on an arc (u, v) in G, we add new arcs $(u_i, v_{i+\tau_{\text{bus}} \mod k})$ to G^k for all $i = 0, \ldots, k - 1$, such that τ_{bus} corresponds to the sum of pure transit time τ on this edge plus the average time τ_{stop} for boarding passengers at the stop. As a matter of course, these new edges are dedicated to

public transport. Vice versa, a public transport commodity φ_{bus} must not use the original copies of (u, v) in G^k.

If there is no bus bay, we assume the subsequent traffic to wait behind the public transport vehicle. We may use a similar construction as in the previous paragraph. The bus-exclusive arc $(u_i, v_{i+\tau_{\text{bus}}})$ timely overlaps with the arcs $(u_j, v_{j+\tau})$ for $i \leq j \leq i + \tau_{\text{stop}}$ (where all indices apply modulo k, of course). Thus, we use the bundle constraint $\sum_{j=i}^{i+\tau_{\text{stop}}} \sum_{\varphi \in \Phi} f_\varphi((u_j, v_{j+\tau})) + \frac{c}{k} f_{\text{bus}}((u_i, v_{i+\tau_{\text{bus}}})) \leq \frac{c}{k}$, $i = 0, \ldots, k-1$, with c being the capacity of (u, v), to prevent overtaking. Alternatively, we can use the FIFO queue construction if there is one at the downstream signal. Here, we set a minimum waiting time constraint for the public transport commodity to force a stop equivalent to the time τ_{stop} spent at the regular stop.

Traffic Signal Priority

From the point of view of a traffic manager, a minimum total travel time of all road users is a worthwhile objective. On the one hand, buses carry much more passengers than cars do, hence, buses should get higher priority. On the other hand, as mentioned in the introduction, we also have to consider the individual traffic. Traffic signals should be optimized with respect to both kinds of commodities, even before active priority strategies are applied. Traffic parallel to bus routes may even benefit from bus priority, whereas crossing traffic is often conflicting.
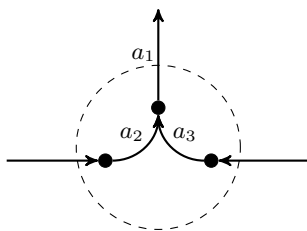
We will distinguish between two approaches. In the first setting, we use a weighted objective where travel time of the bus accounts with a significantly higher factor, usually 20 up to 50. This approach yields a passive transit signal priority where we try to find a good trade-off between the average travel times of all traffic participants.

In the second setting, we will add an additional constraint for the maximum delay of the public transport commodity. That is, we bound the maximum time that a bus can spend on waiting arcs. Here, we may limit the waiting time at every intersection separately, or we can limit the waiting time in total. Obviously, such a bound can be implemented with capacity-like constraints limiting the total flow on all waiting arcs of the bus commodity at each or all intersections. Setting the maximum delay to zero yields a signal setting which is even more aggressive than an active transit signal priority, since it also considers the queues. That is, not only the signal is green when a bus arrives, but also the queue has to be removed in time. In each of these cases, total travel time of all commodities is minimized under these constraints.

Furthermore, it is interesting whether it is sufficient to just (re-)optimize signals on the bus route when adding a new public transport commodity or whether we should optimize all signals in the whole network. This yields many possible combinations and we will study the impact of some of them in Section 5.

4   Feasibility and optimality

In the cyclically time-expanded model, various integer variables control a multi-commodity flow by affecting the capacities. Thus, given demands, it is not even a priori clear whether there exists a feasible solution at all. In fact, there are instances where the problem is infeasible, but the relaxation, i.e., dropping the binary constraints $x \in \{0, 1\}$ and replacing them by $0 \leq x \leq 1$, is solvable. In other words, traffic signals not only reduce capacity of the network due to red phases, also the inner switching logic may cause an additional loss of capacity. An example is given in Figure 3. Please note that the inner logic of the signal in this example is completely fixed and just the offset of the signal can be chosen.

Checking the feasibility of an integer program is one of the classical 21 $\mathcal{NP}$-complete problems of Karp [5]. Yet, in some cases, it is possible to check feasibility of our model in polynomial time, although finding the optimal solution of the combinatorial problem is still $\mathcal{NP}$-hard [6]. In the case of offset optimization and unlimited queue length, that is, green split and phase order are fixed as in the example in Figure 3, we introduce *extended bundled capacities* in the mixed integer programming formulation of our model. For this purpose, we compute the total net green time $T_{\text{net}}^{\mathcal{A}}(e)$ of arcs that end at the same arc $e = (v, w) \in A$, i.e., we compute the length of the time intervals where it is possible to enter a specific arc from arcs in $\mathcal{A} \subseteq \delta^-(v)$. Since flow can only enter the arc $e \in A$ at these times, the net capacity of e is $\frac{T_{\text{net}}(e)}{\Gamma} u(e)$.

Now, the extended bundled capacities are

$$\sum_{t=0}^{\Gamma-1} \sum_{a \in \mathcal{A} \subseteq \delta^-(v)} \sum_{\varphi \in \Phi} f_\varphi(a_t) \leq T_{\text{net}}^{\mathcal{A}}(e) \, u(e)$$

for each arc $e = (v, w) \in A$ and each subset $\mathcal{A} \subseteq \delta^-(v)$. By construction, it is clear that every solution of the mixed integer program also fulfills these constraints. Note that every subset $\mathcal{A} \subseteq \delta^-(v)$ has to be considered. These constraints are also sufficient to provide the following result.

▶ **Theorem 5.** *In the case of pure offset optimization and infinite capacity on waiting arcs, the relaxation of the mixed integer programming formulation with extended bundled capacities of the cyclically time-expanded model is feasible if and only if the mixed integer programming formulation itself is feasible.*

Proof. It remains to show that feasibility of the relaxation implies feasibility of the MIP. Let f be the optimal flow of the relaxation and choose arbitrary offsets for each intersection. Now, we construct a feasible flow f^* in the expanded network with this choice of offsets by using the same underlying paths for each flow particle as in f. If a binary variable is zero for an arc, we use the preceding waiting arc. However, since f fulfills the extended bundled capacities, there is at least one point in time where the subsequent arc has unused capacity left such that every flow particle can be routed over time. ◀

Since we choose an arbitrary coordination of the traffic lights in the proof, we have also shown that each choice of offsets is feasible, if there is at least one feasible choice. However, the cyclically time-expanded network has size $\mathcal{O}(\Gamma n)$ where n is the size of the original

network. Hence, solving the relaxation with a polynomial-time linear programming algorithm is still only a pseudo-polynomial time approach. But we can use the previous result to check feasibility without actually applying the time-expansion.

▶ **Theorem 6.** *If the node degree in G is bounded by a constant, then the feasibility of an instance of the combined offset optimization and traffic assignment problem with unlimited queues can be checked in polynomial time.*

Proof. The main idea is to adjust the capacities of the turning direction arcs in the network G, whose copies in G^k are equipped with binary variables.

Let $T_{green}(a)$ be the length of the green phase of such a turning direction a. Now, we set the new capacity of a in G to $\frac{T_{green}(a)}{\Gamma} u(a)$. Furthermore, we formulate the corresponding extended bundled capacities in G as $\sum_{a \in \mathcal{A} \subseteq \delta^-(v)} \sum_{\varphi \in \Phi} f_\varphi(a) \leq \frac{T_{net}^{\mathcal{A}}(e)\ u(e)}{\Gamma}$ for each $e = (v, w) \in A$ and each subset $\mathcal{A} \subseteq \delta^-(v)$. Since the node degree is bounded, we can bound the number of subsets $\mathcal{A}$ of $\delta^-(v)$ by a constant. Hence, the multi-commodity flow problem in G with these additional constraints is still linear in the input size.

When we have a feasible solution f of this multi-commodity flow problem on the modified graph G, we can copy this flow to G^k by assigning $\frac{f(e)}{k}$ to each copy $e_i \in A^k$ of an arc $e \in A$. Furthermore, we set each binary variable $b_i = \frac{T_{green}(a)}{\Gamma}$, where a is the corresponding arc of b_i. In consequence, each such arc a has a capacity of $\frac{T_{green}(a)}{\Gamma} \frac{u(a)}{k}$, which is also the maximum amount of flow that we have assigned to it. The capacity constraints also hold for arcs without binary variables in the expanded network. Additionally, the extended bundled capacities of the network G assure the extended bundled capacities of G^k. So we know that there is a feasible solution of the relaxation in G^k, when there is a feasible solution in G. Theorem 5 yields a feasible solution of the MIP in this case.

Vice versa, there is also a feasible solution of the flow problem in the original network G, when there is a feasible solution of the relaxation in G^k. Firstly, we re-arrange such a feasible flow of the relaxation as follows: We sum up the flow over each path of the original network. Afterwards, we uniformly distribute this flow over time, i.e., each copy in the cyclically time-expanded network gets the same amount of flow. Furthermore, we set each $b_i = \frac{T_{green}(a)}{\Gamma}$. Obviously, this solution is also feasible, since each capacity constraint is fulfilled. Otherwise, the sum of the flow of one path in the original network would be greater than the total capacity of that path, which would be a contradiction. The obtained solution has the same structure as the solution which was created in the other direction of the proof. Hence, we know that there must be a corresponding feasible solution in the original network. ◀

In the previous section, we introduced constraints limiting the maximum waiting time of a bus to establish transit signal priority. In general, it will be $\mathcal{NP}$-hard to decide, whether the offset optimization problem with such constraints is still feasible. However, in a very restricted case it is decidable in polynomial time.

▶ **Theorem 7.** *Let V_i be the set of intersections that are visited by the public transport commodity i, no intersection is visited twice by a commodity, and no bus queuing waiting arcs are used. Assume that $V_i \cap V_j = \varnothing$ for each pair (i, j) of public transport commodities. Then the problem with constraints that limit the maximum waiting times of public transport is feasible iff it is feasible without these constraints.*

Proof. Since public transport commodities do not interfere at intersections in this case, we can build a progressive signal setting ("green wave") for each public transport commodity.

Since there is no FIFO queuing, public transport can overtake any standard traffic. So we can find a signal setting where none of the public transport flow particles has to use waiting arcs.   ◀

Please note that this result does not hold anymore, if bus queuing waiting arcs are used. In this case, it is possible that a large amount of standard traffic reaches an intersection just before the public transport vehicle. If the outgoing arc has less capacity, then queuing is unavoidable and prevents the public transport vehicle from directly passing the intersection.

5   Computational results

In this section, we will present a computational case study for the public transport extension of the cyclically time-expanded network model. For this purpose, we use a real-world scenario, namely the inner-city of Denver, the capital of the U.S.-state of Colorado. We consider a 6 × 6-grid between *15th Street* and *20th Street* and between *Stout Street* and *Larimer Street*. Nearly every street is a one-way street, indicated by arcs in Figure 4. Furthermore, the *16th Street* between vertex 7 and 12 is a pedestrian and transit mall, so individual traffic is not permitted there. The underlying network was provided by Wünsch [19] and we already studied common signal optimization in this network in [16]. Here, our scenario contains eight different commodities, which have in total a demand of 285 cars per minute. Moreover, we consider two bus lines. Their routes and stops are inspired by bus line 9 (red line) and 36L (green line) of the *Regional Transportation District*, the transit authority of Denver. We slightly change the route of line 36L to fit it into our grid. On each stop, we consider a waiting time of 20 seconds. Since there are bus bays at the two lines in Denver, we allow overtaking of buses while stopping there. Moreover, the extended bundled capacities introduced in Section 4 do not significantly strengthen the MIP formulation in this scenario, since there are only three intersections where two or more incoming streets can have green signals at the same time.
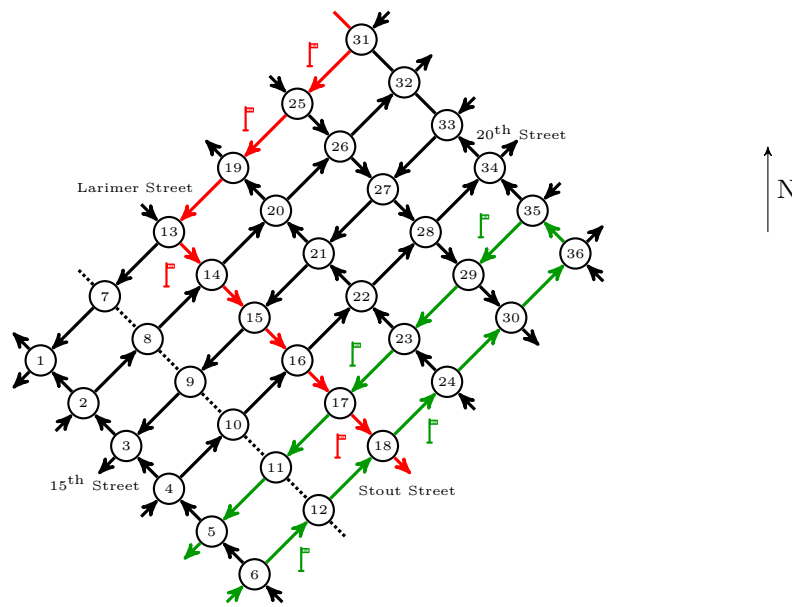
Optimizing coordinations for public transport

First of all, we optimized the offset coordination of the 36 intersections for the standard traffic commodities $\varphi \in \Phi$. Here and in all upcoming computations, we spent one hour of computation time. But as already described in previous results [8], this time is in most cases not sufficient to close the duality gap due to difficulty of computing good lower bounds. Note that the usual MIPs for this network consist of about 36.000 variables including 800 binaries and 19.000 constraints.

The optimization with respect to standard traffic yields a traffic signal coordination which is quite disadvantageous for buses. The two bus lines have to wait 28 and 70 seconds in total at red traffic lights, respectively, although we did not use the FIFO queues in this first attempt.

Thus, we applied the two approaches discussed in Section 3. Firstly, we limited the total waiting time for buses. The results are shown in Table 1. There, we also compare the two variants of optimizing signals only on bus routes versus optimizing all signals.

Secondly, we used weights to account for the high number of passengers in buses. The results for various choices of the weight are presented in Table 2. Several but not all solutions of the first approach could be achieved by an appropriate weight.

Figure 4 Network of downtown Denver, Colorado. The red arcs show the route of line 9, the green arcs the route of the modified line 36L. The flag pictograms mark the bus stops.

Table 1 Total waiting times of the individual traffic in seconds for different maximal waiting times of public transport. On the left hand side, only signals at intersections of public transport routes are re-optimized. On the right hand side, all crossings are optimized. Please note that no new solutions occur for maximum waiting times for 40 to 60 seconds.

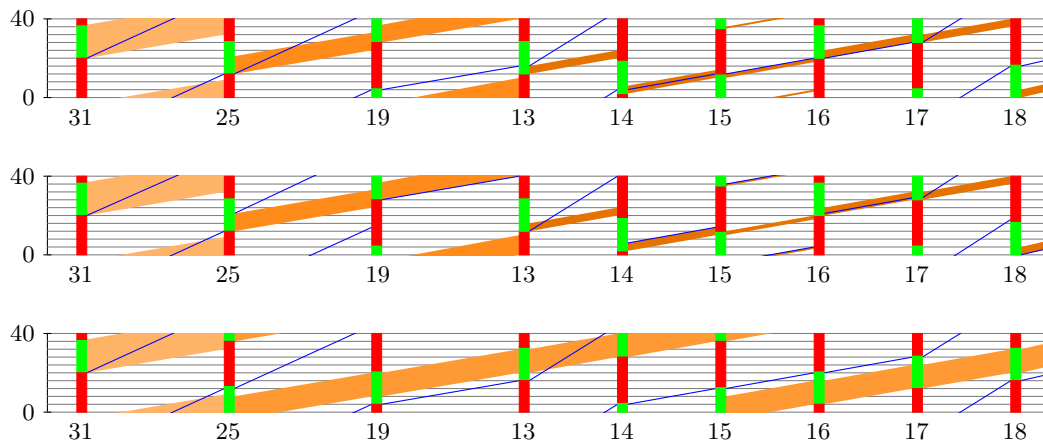| maximal waiting times	only PT-crossings optimized waiting times of			all crossings optimized waiting times of		
	individual	red line	green line	individual	red line	green line
0	3,181	0	0	2,824	0	0
10	2,810	8	6	2,660	6	4
20	2,766	14	6	2,569	20	4
$30 - 60$	2,556	22	30	2,528	24	30
70	2,470	28	70	2,470	28	70

Analyzing the FIFO queues

In the following, we will show that our FIFO queues developed in Section 3 are also applicable in practice. Firstly, we consider the red bus line and a parallel standard commodity on the same route. Figure 5 shows a space-time-plot of both commodities. In the coordination that is given in the upper diagram in Figure 5, the bus (blue line) can pass all signals without waiting, but it violates the FIFO principle. In detail, there is flow (various shades of orange) at intersections 25, 14 , and 16 which arrives during the red phase before the bus and therefore, this flow queues. However, the bus leaves first, when the signal switches to green implying it has overtaken the queue. Please note that traffic density of the other traffic varies, since there is a change from two to four lanes at intersection 13. Moreover, traffic can only become denser, if some flow units are delayed by a red light.

If we use FIFO queues in this example with the same signal setting, the bus falls out of its schedule, since it is overtaken at its four bus stops. The queues formed by these passing vehicles cannot be overtaken anymore and they result in additional delay for the bus which

| weight | only PT-crossings | | | all crossings | | |
| factor | waiting times of | | | waiting times of | | |
	individual	red line	green line	individual	red line	green line
1	2,470	28	70	2,470	28	70
5	2,556	22	30	2,569	20	4
10 − 15	2,810	8	6	2,678	6	2
20 − 30	2,908	2	6	2,824	0	0
≥ 35	3,181	0	0	2,824	0	0

■ **Figure 5** The space time diagrams show traffic of a standard commodity (orange) and a bus of
line 9 (blue) on their route through the network. Time is shown in vertical direction and has to be
interpreted in a cyclic manner as in the cyclically time-expanded network. Horizontal distances are
chosen with respect to real distances. Traffic density is shown by different shades where darker color
means denser traffic. There are four bus stops on arcs 31–25, 25–19, 13–14, and 17–18. In the upper
diagram, no FIFO queues are used and the bus overtakes the queue in front of it at intersections 25,
14, and 16. In the diagram in the middle, the same traffic signal setting is evaluated with FIFO
queues. The bus is significantly delayed by the queues and misses five green phases. In the bottom
diagram, an optimized solution with FIFO queuing is shown. Now, no waiting time for the bus can
be realized even with respect to queues.

culminates in five missed green phases at downstream signals and a total delay of 104 seconds
for the bus. This solution is shown in the diagram in the middle in Figure 5.

Thus, we need to re-optimize signals with respect to the queues. This yields a solution
where the bus does not need to stop at red signals. As one can see in the bottom diagram in
Figure 5, there is always enough green time to empty the queues before the bus arrives at
the signal.

Although the example seems to be designed badly on purpose, the same effect occurs in
the whole network. As an example, we consider the solutions shown in Table 1 where we
optimized signals on the bus route with a maximum waiting time for buses of 0 seconds.
There, we computed an overall waiting time of all commodities of 3,181 seconds (see Table 3

Table 3 Waiting times for two different coordinations with and without FIFO queues.

scenario	waiting times without FIFO			waiting times with FIFO		
	individual	red line	green line	individual	red line	green line
scenario 1	3,181	0	0	3,419	26	0
scenario 2	3,188	0	0	3,307	0	0

scenario 1). Yet, using the same signal settings with FIFO queues, the waiting time for the bus of the red line rises to 26 seconds. Also the waiting time of other commodities significantly increases by about 240 seconds due to the interaction of the queues.

Re-optimizing the same scenario with FIFO queues, we obtain a coordination with objective values shown in row *scenario 2* of Table 3. Thus, already accounting for queues in the optimization yields an additional delay of only 126 seconds in total for the standard commodities, but we now have a solution respecting the FIFO principle with zero waiting time for buses of both lines. Since this coordination is slightly worse without FIFO queues than the first solution, the new solution cannot be found by an optimization tool without considering queuing effects. This example also shows the great importance of considering other road users when designing transit signal priority. An excellent solution for public transport vehicles in an empty network can turn into a rather bad solution when queuing occurs.

6   Discussion

In this paper, we have introduced an extension of the cyclically-time expanded network model for traffic signal optimization supporting public transport. Together with the queuing model, this approach yields competitive solutions for passive transit signal priority.

While parameters like delays at bus stops had to be estimated, one can assume that these parameters will vary in practice. Moreover, in passive TSP, crossing transit lines conflict in general, whereas in most cases in practice buses arrive at different times at such intersections. Thus, many open research questions remain. For example, one may use a cyclically time-expanded network which is expanded with respect to the cycle time of the time table of public transport. Hence, one may use a different signal setting when the bus actually arrives at an intersection. Is it possible to develop an active transit signal priority strategy based on this approach? Can one use solutions of our model to identify intersections suitable for active TSP? In other words, if the bus is already waiting in the passive approach at a certain intersection, the signal setting of this intersection is most likely crucial for travel times of many commodities and one should maybe not use TSP here. Vice versa, how can one find intersections where small changes like active TSP do not destabilize the coordination in the whole network? If one uses active TSP, it is also an open question how to optimally synchronize the traffic signal after the public transport vehicle has passed. In principle, one may also use a time-expanded network expanded over two or more cycles the compute an optimal repair strategy for the network-wide coordination. Can one model passenger flows similar to the traffic flow of the standard commodities? A varying number of passengers in the bus may lead to different solutions in the case of weighted waiting times. Is it even possible to integrate a mode choice model, i.e., some commodities may choose between traversing the network as an individual traffic commodity or by using the public transit network. Here, flow units of public transport commodities would need to act as containers to which one can assign flow particles of other commodities.

────── **References** ──────

1   Z. R. Abdy and B. R. Hellinga. Analytical method for estimating the impact of transit signal priority on vehicle delay. *Journal of Transportation Engineering*, 137(8):589–600, 2011.

2   W. Brilon and W. Laubert. Priority for public transit in Germany. *Journal of Advanced Transportation*, 28(3):313–340, 1994. `doi:10.1002/atr.5670280309`.

3   L. R. Ford and D. R. Fulkerson. *Flow in Networks.* Princeton University Press, Princeton, 1962.

4   N. B. Hounsell and B. P. Shrestha. AVL based bus priority at traffic signals: A review of architectures and case study. *European Journal of Transportation and Infrastructure Research*, 5(1):13–29, 2005.

5   R. M. Karp. Reducibility Among Combinatorial Problems. In R. E. Miller and J. W. Thatcher, editors, *Complexity of Computer Computations*, pages 85–103. Plenum Press, New York, 1972.

6   E. Köhler and M. Strehler. Traffic Signal Optimization Using Cyclically Expanded Networks. In *10th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS'10)*, volume 14 of *OpenAccess Series in Informatics (OASIcs)*, pages 114–129, Dagstuhl, Germany, 2010. `doi:10.4230/OASIcs.ATMOS.2010.114`.

7   E. Köhler and M. Strehler. Combining static and dynamic models for traffic signal optimization – inherent load-dependent travel times in a cyclically time-expanded network model. *Procedia - Social and Behavioral Sciences*, 54(0):1125–1134, 2012.

8   E. Köhler and M. Strehler. Traffic signal optimization using cyclically expanded networks. *Networks*, 65(3):244–261, 2015. `doi:10.1002/net.21601`.

9   P. Koonce, L. Rodegerdts, K. Lee, S. Quayle, S. Beaird, C. Braud, J. Bonneson, P. Tarnoff, and T. Urbanik. *Traffic Signal Timing Manual.* Report Number FHWA-HOP-08-024. Federal Highway Administration, 2008.

10 Y. Lin, X. Yang, N. Zou, and M. Franz. Transit signal priority control at signalized intersections: a comprehensive review. *Transportation Letters*, 7(3):168–180, 2015.

11 H. Liu, A. Skabardonis, and W. Zhang. A dynamic model for adaptive bus signal priority. In *82nd Transportation Research Board Annual Meeting*, 2003.

12 H. Liu, J. Zhang, and D. Cheng. Analytical approach to evaluating transit signal priority. *J. Transp. Syst. Eng. Inf. Technol.*, 8(2):48–57, 2008.

13 V. Ngan, T. Sayed, and A. Abdelfatah. Impacts of various parameters on transit signal priority effectiveness. *J. Public Transp.*, 7(3):71–93, 2004.

14 R. Scheffler and M. Strehler. Optimizing Traffic Signal Timings for Mega Events. In M. Goerigk and R. Werneck, editors, *16th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2016)*, volume 54 of *OpenAccess Series in Informatics (OASIcs)*, pages 8:1–8:16, Dagstuhl, Germany, 2016.

15 A. Skabardonis. Control strategies for transit priority. research report, University of California: California Partners for Advanced Transportation Technology, Berkely, 1998.

16 M. Strehler. *Signalized Flows – optimizing traffic signals and guideposts and related network flow problems.* Phd thesis, Brandenburgische Technische Universität, Cottbus, Germany, 2012.

17 S. R. Sunkari, P. S. Beasley, T. Urbanik, and D. B. Fambro. Model to evaluate the impacts of bus priority on signalized intersections. In *Transportation Research Record 1494*, pages 117–123, 1995.

18 J. Wahlstedt. Impact of bus priority in coordinated traffic signals. In *Procedia Social and Behavioral Sciences*, volume 16, pages 578–587. Elsevier, 2011.

19 G. Wünsch. *Coordination of Traffic Signals in Networks.* Phd thesis, Technische Universität Berlin, 2008.