Traffic Signal Optimization Using Cyclically Expanded Networks*

Ekkehard Köhler[1] and Martin Strehler[2]

1   Mathematisches Institut, BTU Cottbus
     Postfach 10 13 44, 03013 Cottbus, Germany
     ekoehler@math.tu-cottbus.de
2   strehler@math.tu-cottbus.de

── Abstract ──────────────────────────

Traditionally, the coordination of multiple traffic signals and the traffic assignment problem in an urban street network are considered as two separate optimization problems. However, it is easy to see that the traffic assignment has an influence on the optimal signal coordination and, vice versa, a change in the signal coordination changes the optimal traffic assignment. In this paper we present a cyclically time-expanded network and a corresponding mixed integer linear programming formulation for simultaneously optimizing both the coordination of traffic signals and the traffic assignment in an urban street network. Although the new cyclically time-expanded network provides a model of both traffic and signals close to reality, it still has the advantage of a linear objective function. Using this model we compute optimized signal coordinations and traffic assignment on real-world street networks. To evaluate the practical relevance of the computed solutions we conduct extensive simulation experiments using two established traffic simulation tools that reveal the advantages of our model.

1   Introduction

Urban traffic congestion is increasing day by day. Growing cities and increasing population doubled the traffic volume in the last two decades in Europe and North America and an even higher rise has to be expected for the urban regions in Asia or South America in the next years. Traffic congestion causes delays which add up to huge costs for society and business. The urban mobility report 2009 [17] states a total loss of 4.2 billion hours and 87.2 billion dollars for the 439 urban areas in the United States in only one year. Wasted fuel of 2.8 billion gallons, noise, and pollution accumulate.

A reduction of congestion by simply expanding the infrastructure is often impossible due to space limitations. Besides strengthening public transport an increase of network performance seems to be a suitable way to cope with the boosting traffic. While the roads itself do not allow much control of the traffic flow, signalized intersections provide a lot of traffic signal parameters ready to be optimized. Especially the coordination of numerous traffic signals at different intersections in a network seems worthwhile. The main concept of a so-called 'green wave', i.e., the coordination of traffic signals along an arterial road,

was already introduced by Adolph [1] in 1925. But it was not before 1964, when Morgan and Little [15] started a broad analysis and presented a graphical solution for calculating maximal time slots and bandwidths for a single road. By now there exist various approaches to transform the concept of a 'green wave' to a whole network of roads.

Today's research in traffic signal optimization can be considered to be split into two main streams. On the one hand, big efforts have been made on developing highly adaptive traffic signals which are able to react on changes in traffic volume immediately (see for example [12]). On the other hand, assuming fairly stable demands within certain divisions of time, the usage of pre-timed traffic signals appears reasonable [9, 13, 21]. Considering rush hour situations with high demand at every road, adaptive coordinations will fall back to fixed time coordinations. Furthermore, due to the high costs of sensor technologies and a lacking acceptance of car-to-car or car-to-traffic-signal communication for data privacy reasons, we believe fixed traffic signal coordination to be the choice for many cities.

Another distinction can be made between heuristic optimization methods and purely mathematical programming approaches and their underlying models. The first class is based on genetic algorithms, fuzzy logic, or neural network approaches. They often provide good and fast solutions but do not give bounds or guarantees on the quality of the solution compared to an optimal solution. The second class usually can not cover all features and parameters of a real traffic network, but it provides provably good solutions that are of interest of their own or can supply a good bound or starting solution for models of the first class.

The *aging* of fixed-time traffic signal plans [4] is important but often unappreciated in the optimization of traffic signals: It can be observed that some traffic signal coordinations, once having been very efficient, become worse over time. Every change of fixed time signal plans may influence the road choice of the users. Increasing traffic in general and road users changing to optimized arterial roads lead to higher waiting times and may disturb the fine-tuned coordination. Therefore traffic signal coordination and the traffic assignment should be considered as one single optimization problem. Hereby, traffic signal coordination means an optimal choice of the parameters of the traffic signals such that the road users reach their destinations fast and congestion is avoided as much as possible. Traffic assignment deals with route choices of the users and describes the distribution of traffic in the road network. Traffic assignment can be seen from an administrative point of view, where a central authority tries to reduce the overall travel time of the whole system. Or it can be seen from a game theoretical point of view, where each road user decides locally to optimize her/his experienced travel time. A precise definition will be given after the introduction of the necessary parameters. Obviously there is a feedback between the problem to coordinate the traffic signals and assigning traffic units to optimal paths in the network. For the combined problem Allsop and Charlesworth [3] studied a first partially-unified approach and presented an iterative method while Chiou [6] used a gradient projecting method for calculating local optima within his traffic model. Recent results were made by Bell and Ceylan [5] as well as Teklu et al. [20] using genetic programming.

In this paper we present a new approach for optimizing traffic signal coordination and traffic assignment simultaneously. Our objective function here is the minimum of the overall travel time of all road users. We denote this as *traffic signal coordination and traffic assignment problem* (TSCTAP). The model is based on a time-expanded network, that uses the periodicity of traffic signals to limit the time horizon. More precisely we use a cyclic time-expansion and provide a realistic implementation of traffic signals. Our model captures flow-dependent transit times on edges but omits nonlinear link-performance functions

together with their complex analysis. Instead, the traffic assignment is based on separate functions for travel and waiting times. Therefore, the traffic assignment problem can be solved efficiently for fixed coordinations. The TSCTAP itself can be formulated as a mixed integer program. This provides guarantees on the quality of the solution and a rather easy concept of handling user equilibria.

Furthermore, with our new model we can capture some special properties of innercity traffic, e.g. platoons of cars. Platoons naturally develop mainly due to the red and green phases of the traffic signals. Our model can capture these changes in traffic density and we can observe effects like splitting or merging of platoons. Also, the existence of several different platoons on one road and realistic departure times at signals (usually at beginning of the green phase, but not before the platoon arrives at this signal) are modeled.

The paper is organized as follows. Initially, basic definitions are presented. Then we shortly discuss different measurements of performance and the complexity of the traffic signal coordination problem in Section 2 and 3. In Section 4 we define the cyclically expanded network model. To evaluate the practicability of the optimized signal coordination and traffic assignment we use two traffic simulation tools. Section 5 gives the results of these simulation experiments and some further computational results. Finally, in Section 6, we summarize the findings and give an outlook on future work.

2 Preliminaries

In this section we sketch some basic facts and definitions on traffic networks that are prerequisites for understanding the rest of the paper. Due to the numerous aspects of inner-city traffic we concentrate on concepts for traffic signals and traffic assignment. Traffic signals at intersections are characterized by various parameters: cycle time, red-green split, order of the phases and the offset between traffic signals at adjacent intersections. Each single traffic light has a characteristic sequence of *red* and *green* that appears periodically with *cycle time* Γ. The proportion between red and green is called *red-green split*. All lights at an intersection are grouped together to signal groups which again are grouped together to a traffic signal. The (interior) offset of their sequences is fixed to avoid collisions.

If all traffic signals in a network have the same cycle time, one can also look at the *offset* of the intersections with respect to a global system time. Let a car need t units of time to bridge the distance between two consecutive intersections. If the green phase of the second traffic signal group starts exactly t time units after the green phase of the first one, the car experiences a *green wave*.

The most important parameters are depicted in the *signal plans* in Figure 1. All of them can be used for traffic signal optimization. In this paper we will concentrate on optimizing the offset between consecutive traffic signals and we assume all other parameters to be fixed. Additionally, we assume a common cycle time at each traffic signal. Note that this is not a hard restriction; if traffic signals in the network have different cycle times then the least common multiple of all cycle times can be used as a unified cycle time instead.

Let the street network be represented as a directed graph $G = (V, A)$ with node set V and arc set A. We will use the notation in Table 1. We also refer to the sequence of *red* and *green* lights as *operating sequence* of a traffic signal.

The performance of traffic signal optimization can be measured by different parameters. The most common measure is the average delay or the number of stops. A weighted combination of those has also been used. Sun et al. [19] show that the cycle time has great influence on the objective function when using these parameters. Long cycle times minimize
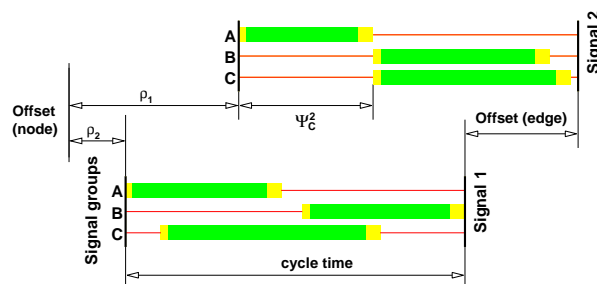
■ **Figure 1** Signal plans for two traffic signals. Signal 2 may belong to an intersection with two crossing streets. Signal group B and C have an interior offset of $\Psi^2_{B/C}$ with respect to a fixed time during the period. The green phase of C is extended, for example to enable left-turning. Signals 1 and 2 themselves have offsets ρ_1 and ρ_2. Therefore, the signal group A at signal 2 turns green exactly $\rho_2 - \rho_1$ time units after the signal group A at signal 1 does.

$n \in \{1, \ldots, N\}$	intersection indexes
$l \in \{1, \ldots, M\}$	link indexes
Γ	cycle time
ρ_n	(node) offset of intersection n
Φ_n	set of signal plans of intersection n
$A, B, C, \ldots$	signal groups
Ψ^n_A	intra-node offset of signal A at intersection n
f_l	link flow
c_l	capacity of link l
t_l	free speed travel time on link l
Θ	a set of commodities with source node, sink node, and demand

■ **Table 1** Notation for traffic lights and signals in a network model.

stops while short cycle times lead to less delay.

As we will simultaneously compute the traffic assignment problem and an optimal traffic signal coordination we will use the overall travel time of all road users as the measure of quality of the solution. Traffic participants may choose arbitrary routes. Thus, taking only stops and delay into account, a road user may choose a very long path through the network just to avoid stopping or waiting in front of a red traffic light. This is rather unrealistic as most road users are interested in the fastest way to their destination. Other measures of performance, like preferences for public transport and pedestrians, are not considered in this paper.

To define traffic assignment we need the notion of *flow* and *multi-commodity flow*. Due to space constraints, we omit a detailed definition here and refer the reader to a standard text book on network optimization like [2]. For short, a flow is a function $f : A \mapsto \mathbb{R}_{\geq 0}$ that has to fulfill capacity bounds and flow conservation constraints. We also require that the demands of all the commodities are satisfied. In the multi-commodity case capacities are shared by the different commodities. Furthermore, a travel time function is given for each link. A *traffic assignment problem* is the distribution of traffic flow in a street network satisfying demands of flow between origin and destination pairs of nodes. Assignment methods are

looking for an optimal way to distribute the traffic flow in the network according to different objective functions.

One option for this objective function is to minimize the overall transit time spent by all road users in the network, $\sum_{e \in A} f(e)t(f(e))$. Here $t(f(e)) : \mathbb{R} \mapsto \mathbb{R}$ is the travel time function describing the time a flow unit needs to traverse arc e in dependence on the amount of flow $f(e)$ on arc e. This solution is called the *system optimum*. While this solution is good for the system as a whole, it might be unfair to single users that could improve their own travel time by changing to a faster path. However, they will disturb traffic flow there. In contrast to the system optimum, a flow where no road user can find a faster route for himself is said to be satisfying *Wardrop's principle* and is called *user equilibrium*. This equilibrium is often used for modeling the behavior of traffic in street networks, where experienced users choose the path that minimizes their own travel time. Obviously, there is a gap between the value of the 'selfish' user equilibrium and the system optimum. This gap is often called the *price of anarchy*. A detailed survey on *selfish routing* can be found in [16].

The problem of finding the user equilibrium is referred to as equilibrium traffic assignment problem. For simplicity we will call this problem *traffic assignment problem* in the following. To calculate travel time on a given road, common static traffic models use *link performance functions* (or travel time functions) that depend on the load or flow on that particular road. To model both the time to traverse the road and the waiting time at the end of the road one uses appropriate monotone, convex functions, which are usually non-linear. Due to this non-linearity, the traffic assignment problem cannot be solved by simply applying standard network flow algorithms but requires more involved approaches.

Note that the described static traffic model does not capture any time-dependent behavior. For example (static) link performance functions can not handle platoons of traffic arriving at different points in time. But these platoons are essential for modeling inner-city traffic as they are formed naturally by all the traffic signals on a certain route. Furthermore, these platoons significantly depend on the traffic signal coordination.

Our model of traffic.

In our traffic model we make the following assumption. The travel time on a link in the network splits into two components: the free speed travel time that is needed to bridge the distance of the link, and the waiting time in front of the intersection, i.e., in front of a traffic signal. In particular, we assume the free speed travel time to be independent of the load of the street. Although this assumption is not justified on highways or in rural areas, it is appropriate in inner-cities, where the distance between consecutive intersections is comparable small and a strict speed limit is present. The speed of a single car does not differ much from the speed of a platoon of cars. The waiting time mainly depends on the offsets of the subsequent traffic signals and it depends linearly on the traffic load. Multiplying load and travel time yields a quadratic behavior of the overall travel time of all road users on this link.

3   Complexity of offset optimization

In this section we discuss the complexity of the *signal coordination problem*. The formulation and analysis essentially depends on the chosen model. In the literature various proofs of complexity for different approaches can be found. All of them have in common that the offset optimization problem for traffic signal coordination is $\mathcal{NP}$-*hard*. Network coordination is similar to the Periodic Event Scheduling Problem PESP. Serafini and Ukovich [18] proved

NP-completeness for the PESP by reducing the Hamiltonian Cycle Problem. In [21] Wünsch provides a reduction from PESP to his formulation of the Network Signal Coordination problem.

In the following we will sketch a $\mathcal{NP}$-completeness proof that fits our traffic model. At the end of Section 2 we motivated the use of constant transit times plus waiting times at the intersection. Building up on this assumption we define the signal coordination problem as follows.

▶ **Definition 1** (SIGNAL COORDINATION PROBLEM). INPUT: Network $G = (V, A)$ with capacities and travel times consisting of a constant transit time plus waiting time at the intersection; a set of traffic signals with arbitrary but fixed operating sequences at some of the intersections $v_n \in V$, $n \in \{1, \ldots, N\}$, the cycle time Γ is the same for all traffic signals; commodities $\theta \in \Theta$, $\theta = (s_\theta, t_\theta, d_\theta)$ with origin, destination and demand
OUTPUT: a set of offsets $\{\rho_1, \ldots, \rho_N\}$ which minimize the overall travel time

We reduce the SIGNAL COORDINATION PROBLEM to the 3-PARTITION PROBLEM which is $\mathcal{NP}$-complete [8]. The problem is to decide whether a given multiset $S = \{x_1, \ldots, x_{3m}\}$ of $3m$ integers can be partitioned into disjoint triples $S_1, \ldots, S_m$ that all have the same sum $C = \frac{1}{m} \sum_{i=1}^{3m} x_i$. The 3-PARTITION PROBLEM remains $\mathcal{NP}$-complete when $\frac{C}{4} < x_i < \frac{C}{2}$ $\forall i \in \{1, \ldots, 3m\}$.

Assume $3m$ roads, with x_i, $(i = 1, \ldots, 3m)$ flow units in each road, $C = \frac{1}{m} \sum_{i=1}^{3m} x_i$ and $\frac{C}{4} < x_i < \frac{C}{2}$ $\forall i \in \{1, \ldots, 3m\}$ as above. There is a traffic signal at the end of each road that has a cycle time of m time units and is green for exactly one time unit in each period. Let the roads be wide enough such that all flow units from one road can pass their traffic signal during one time unit. All intersections directly lead to a narrow road with a capacity of C cars per time unit. Figure 2 shows the network used for the reduction.
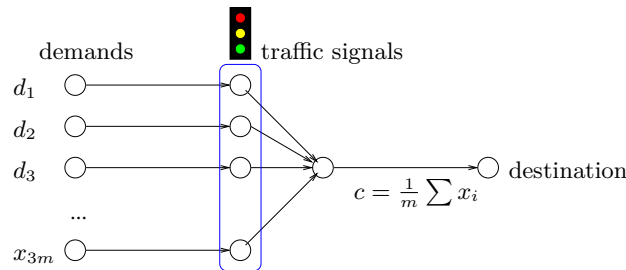
Figure 2 Traffic network for reduction to 3-PARTITIONING. The narrow road at the exits forces perfect coordination.

As the capacity can not be exceeded flow units will have to wait when this small road at the exit is congested. If too much traffic signals turn green at the same time, flow units will have to wait in front of a green traffic signal. Thus, if the related 3-PARTITION PROBLEM has a YES answer, we can use the exit road at maximum capacity all the time and we need m time units to empty the network. If the 3-partitioning has no YES solution, then there will be congestion in the network and flow units will have to wait for the next green cycle yielding a higher overall travel time. This directly leads to Theorem 2.

▶ **Theorem 2.** *Offset optimization of traffic signals is $\mathcal{NP}$-complete, even without fixed route choice and with travel times consisting of constant transit time plus waiting time.*

It is easy to see that there are more realistic traffic networks than the one in Figure 2

having similar properties. The main idea of the construction is to use traffic signals for sorting, assigning and splitting traffic.

4   A new model

In this section we describe the various parts of our model that are combined to solve the TSCTAP. These parts cover the cyclic time-expansion, expansion of crossings, and implementation of traffic signals. Building on these notions we will give a mixed integer programming formulation, study basic properties, and discuss various subsequent improvements of the model.
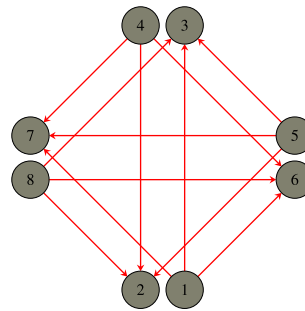
4.1   Modeling traffic signals

Cyclically time-expanded networks.

The notion of time-expanded graphs was introduced in the context of dynamic flows or flows over time. Unlike standard *static* network flows, flows over time capture flow units that travel through the network in a timely fashion. Flows over time were first studied by Ford and Fulkerson in their seminal work on network flow theory [7] (see [11] for further references). Already in this early work they introduced the concept of *time-expanded networks*. In this expanded network for every node several copies of this node are added to the graph, one for each desired time step. Then these nodes are connected by arcs, where the various copies of the vertices are connected according to the travel times of the original arcs. Figure 3 shows a simple network together with its time-expanded network. The big advantage of flows over time is the opportunity to capture the complete time-dependent behavior of the journey of a flow unit in a given network. In contrast, a static flow can only describe the time-independent behavior of this journey. Using time-expanded graphs one can not only model flows over time, but one can also solve various optimization problems of network flows efficiently (in the size of the time-expanded graph).
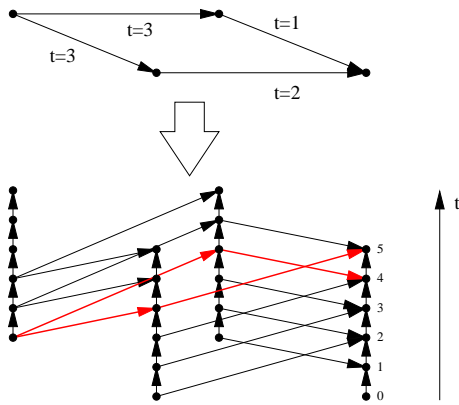
Figure 3 Simple example of a 4-vertex graph together with its time-expanded graph for time horizon t.

Figure 4 Expanded intersection with arcs for each turning alternative.

As explained earlier, for traffic signals and their coordination, a time-dependent model, capable of describing the time-offset between consecutive intersections is a vital ingredient. Hence, flows over time and time-expanded networks seem to be the suitable mathematical model. Yet, time-expanded networks are rather inefficient if the time horizon is large, since

this determines the number of network copies that have to be provided. However, since traffic signals have a periodical behavior it is not necessary to use a full time horizon expansion. Instead, we suggest a *cyclic time-expansion* where we use only $k \in \mathbb{N}$ time steps of size $t = \frac{\Gamma}{k}$ and add the arcs according to transit times modulo k with adjusted capacities. To model the ability to wait in front of an intersection, all copies of one node v are cyclically connected in chronological order with waiting arcs $v_i v_{i+1}$ to model a cyclically rolling horizon (see Figure 5). A cost of one time unit is assigned to each waiting arc.

Intersections.

To model intersections with different lanes, turning directions, and interior traffic signal offsets, we use a standard approach from traffic networks. Every intersection node is split up into several nodes for incoming and outgoing traffic; interior arcs connect the lanes (see Figure 4). Each of these arcs is assigned to one traffic light.
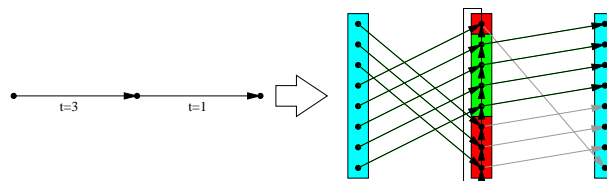
Traffic signals.

Now, the traffic signals themselves can be modeled by binary decision variables that switch the capacities of the interior arcs at an intersection on or off, depending on the signal plan and the corresponding time step (see Figure 5). For each traffic light a matrix $A_e \in \{0,1\}^{k \times k}$ is given; here $A_{eij} = 1$ means that the particular traffic light is green at time step j when using offset $i\frac{\Gamma}{k}$. So each row stands for a certain offset and determines the operating sequence of the traffic light for this specific offset. For each interior lane e of the not time-expanded graph we create such a matrix.

$$A_e = \begin{pmatrix} 0 & 0 & 1 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & \cdots & 0 \\ \multicolumn{7}{c}{\dotfill} \\ \multicolumn{7}{c}{\dotfill} \end{pmatrix}$$

For each intersection n we introduce k binary variables $\vec{b_n} = (b_{n_1}, \ldots, b_{n_k}) \in \{0,1\}^k$ with constraint $\sum_{t=1}^{k} b_{n_t} = 1$ that describe the solely chosen offset at the intersection. More precisely, $b_{n_i} = 1$ is equivalent to offset $\rho_n = i\frac{\Gamma}{k}$ at intersection n. By multiplying this characteristic vector $\vec{b_n}$ with A_e and the capacity c_e of a lane, we can switch the capacities of the links on or off and thus represent the green and red lights.

Using $f(e) \leq \vec{b_n} A_e c_e$ for all interior lanes of an intersection (where c_e is chosen in correspondence to the granularity of the time-expansion) we can map the complete dynamic behavior of traffic signals in our model.

Figure 5 Cyclic time-expansion of a traffic signal.

Putting together the various parts of the model we get the following definition of cyclic time-expanded networks.

▶ **Definition 3** (Cyclically time-expanded traffic network). A cyclically time-expanded traffic network is a network, that is obtained from a traffic network $G = (V, A)$ by (i) expanding the intersections with arcs for turning alternatives and lanes, (ii) cyclic time-expansion with respect to the cycle time, and (iii) expanding signal plans.

4.2   Modeling traffic assignment

Building up on the cyclically time-expanded network, we can now consider the traffic assignment problem within this framework. Although flow can be considered to travel in a time-dependent manner through the cyclically time-expanded network, one should rather see this model as a static model that just captures some time-dependent aspects of traffic flow. Due to the cyclic repetition of the vertex and arc copies, a flow particle traveling through this network can be seen as a representative of a whole set of temporally repeated particles at every multiple of the cycle time.

On the other hand, there is a closely related interpretation of static network flow for traffic networks. In this interpretation one considers a flow-carrying path in the static network as a mapping of a corresponding amount of flow particles traveling over time through the traffic network at the corresponding flow rate. In other words, a flow-carrying path in the static network represents a constant rate of flow on this path in the 'real' time-dependent traffic network.

This interpretation suggests how to put together the two models, the cyclically time-expanded network on the one hand and the static traffic assignment model on the other hand. Basically, the demands for the different commodities have to be subdivided to the number of layers/time steps in the cyclically time-expanded network. The precise construction is given in the following.

Given (static) commodities $\theta \in \Theta$, $\theta = (s_\theta, t_\theta, d_\theta)$ in the original (static) network (sending d_θ units of flow from node s_θ to node t_θ), one has to extend them to the cyclically expanded network. The demand has to be scaled to time Γ and can be divided uniformly among all copies of s_θ. The constraints for the sink nodes should be less restrictive, i.e., no uniform distribution is required. We propose to use a super-source and super-sink, connected by a backward arc, as shown in Figure 6.
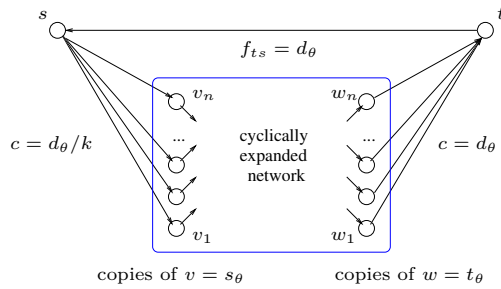
Figure 6 Demand splitting for commodities.

Let $G = (V, A)$ be a cyclically time-expanded traffic network and commodities $\theta \in \Theta$, $\theta = (s_\theta, t_\theta, d_\theta)$, capacities $c : A \to \mathbb{N}$, a set E of interior arcs at intersections with associated matrices A_e for each $e \in E$, travel times t_e for each link and flow values $f_\theta : A \to \mathbb{R}$ for each commodity. Now we can formulate a mixed integer program for TSCTAP. We extend a common multi-commodity min-cost circulation program for the cyclically time-expanded network by adding the binary variables and capacity constraints above. To keep the mixed

integer program in a readable format we present it here in a somewhat condensed form omitting some of the details and indices.

$$\min \quad \sum_{e \in A} \sum_{\theta \in \Theta} t_e f_\theta(e)$$

$$\text{s.t.} \quad 0 \le \sum_{\theta \in \Theta} f_\theta(e) \le c(e) \qquad\qquad \forall e \in A \setminus E \qquad\qquad (1)$$

$$\sum_{e \in \delta^+(v)} f_\theta(e) = \sum_{e \in \delta^-(v)} f_\theta(e) \qquad\qquad \forall \theta \in \Theta \quad \forall v \in V \qquad\qquad (2)$$

$$f_\theta((t_\theta, s_\theta)) = d_\theta \qquad\qquad \forall \theta \in \Theta \qquad\qquad (3)$$

$$\sum_{t=1}^{k} b_{n_t} = 1 \qquad\qquad \forall n \in \{1, \ldots, N\} \qquad\qquad (4)$$

$$f(e) \le \vec{b_n} A_e c_e \qquad\qquad \forall e \in E \qquad\qquad (5)$$

$$f_e \ge 0, \quad \vec{b_n} \in \{0,1\}^k$$

The constraints of type (1) fix the capacity bounds, type (2) implements the flow conservation and the constraints (3) force the circulation. Equation (4) ensures that exactly one offset is chosen at each intersection, and (5) permits flow only on arcs with respect to the chosen offsets.

4.3 Properties of the model

First, we examine the assignment problem for fixed traffic signals.

▶ **Theorem 4.** *Using the cyclically time-expanded network the traffic assignment problem for a fixed traffic signal coordination can be solved efficiently.*

Proof. With fixed traffic coordination, i.e., all binary decision variables of our model are fixed, we obtain a linear program for the traffic assignment problem. This LP can be solved efficiently.

Alternatively, one could just remove all arcs that are switched off. Any algorithm for the min-cost circulation problem can now be used for computing the traffic assignment problem. ◀

Assume flow values assigned to the network and a fixed coordination. Multiplying flow value and transit time and summing over all waiting arcs yields the overall waiting time. Increasing flow at a certain traffic signal will increase the waiting time, because more flow will be assigned to the waiting arcs. However, due to the bounded capacities, the flow units on the waiting arcs may not leave completely on the first 'green' outgoing arc. Instead, the flow will have to use more waiting arcs until the accumulated flow is reduced. Therefore, if the incoming flow is raised linearly on all copies of an arc, then the growth of the waiting time will be not linear but raher quadratic. More precisely the obtained function is piecewise linear, but converges to a quadratic function if the timestep of the expansion converges to zero. In Figure 7 we present the relation between flow and average waiting time on a single link in the cyclically expanded network. The traffic is equally distributed on all copies of this link, a traffic signal with a cycle time of 60 seconds and a red time of 20 seconds is put at the end of the road and a free speed travel time of 10 seconds is assumed. Additionally, a capacity reduction from two lanes to one lane at the traffic signal was used to demonstrate the capability of our waiting edge model.

In praxis, the flow will not be equally distributed on all copies of an arc. Different flow values on the copies can be interpreted as platoons of different sizes and densities.

Considering this platoons the behavior of the average waiting time with respect to platoon length and arrival time at the signal may change dramatically. But due to space limitations we have to omit a detailed analysis here.

Nevertheless, we can also conclude an interesting property of the optimal solution, even for arbitrary coordinations.

▶ **Proposition 1.** *An optimal solution (system optimum) of the proposed multi-commodity min cost flow problem is also an user equilibrium.*

Proof. All transit times are load independent. Therefore any route change of a single road user has no influence on the transit times. Consequently a system optimum is an user equilibrium. ◀
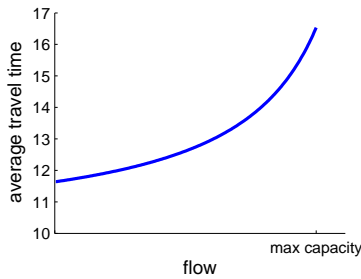
Figure 7 Average travel time with respect to flow on a single link in the cyclically expanded network. Incoming traffic is equally distributed over time.
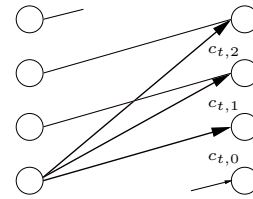
Figure 8 A fan-like expansion to model time-dependent travel times. Note, that the capacities are not equally distributed.

4.4 Improvements

In this section we propose some refinements of our model.

We introduced waiting arcs $v_i v_{i+1}$ between consecutive copies of a node v, but the capacities of these arcs have not been set. For modeling the restriction of small streets we propose choosing the capacity with respect to the length of the street. This bounds the queue at a traffic signal and allows backing-up of traffic over several consecutive intersections.

Often traffic signals are capable of switching between different programs or modes, e.g., different red-green splits. Our model can easily be extended to the case of several programs or modes. This can be done by adding these programs to the matrices A_e. Thus the binary variables not only choose the offset but also the operating sequence.

In reality, the speed within a platoon may vary from car to car. This is not yet captured in our model. Also our assumption of complete load independent travel times on streets between intersections may seem too extreme in some cases. For more realistic flow-dependent travel times we can make use of a model used for flows over time with flow-dependent transit times. In [10] a fan-like time-expanded network is suggested, that models different travel times for different flow-rates on a given arc. Consider two intersections v and u with free-speed travel time T on $e = (v, u)$. Instead of only adding one arc $e_t = (v_t, u_{t+T})$ for each time step of the expansion, several additional arcs $e_{t,i} = (v_t, u_{t+T+i})$ are added and the capacity is split among those copies. See Figure 8 for a visualization of such a fan. Therefore in a situation with low load the original arc e_t provides enough capacity and all road users can be routed with the free speed travel time. Considering a rush-hour scenario some flow units are assigned to the 'slower' copies $e_{t,i}$ yielding an increasing average travel time. This flow dependency does not touch the conclusion of Theorem 4.

5   Computational Results

In this section we evaluate the applicability of the presented model by discussing various simulation results. We apply two well-established microscopic traffic simulation tools: MATSim (TU Berlin) and VISSIM (ptv AG) for our studies.

5.1   Scenarios

Real-world data for offset optimization is always hard to get. One needs detailed information on the network, the traffic flow, and the signalization. Furthermore, we do not just need load or usage information on the roads, but rather commodities with origin-destination information.

   In the results presented here we concentrate on three different scenarios. Wünsch [21] uses parts of the inner-city of Denver and Portland. We chose the same scenario for Portland with 16 intersections in a 4×4-grid and a part of the Denver scenario with 36 intersections in a 6×6-grid which provides the opportunity of benchmarks. These two scenarios show a typical North-American inner-cities grid consisting of regularly arranged one way streets. For better evaluating the interplay between our traffic assignment and the signal coordination we were interested in a more complex street network. Therefore we also chose the inner-city of Cottbus, a German town of about 100.000 inhabitants south of Berlin, with 30 traffic signal controlled intersections to test our model on a more European shaped city. The most detailed data is available for Portland, but for our purposes the scenario is rather poor as the given commodities do not provide much route choice. Every scenario was examined with different load values reaching from nearly empty streets to nose-to-tail traffic.

   Our research partner ptv AG kindly provided us with data for the North American scenarios while we collected data for Cottbus ourselves. In the two larger scenarios traffic signals are influenced by a preference for public transportation. Therefore the comparison of simulated and observed traffic is limited.

5.2   Simulation

As mentioned before, for verifying and comparing our solutions to solutions of other optimization techniques we use two different simulation software tools. The tool VISSIM from ptv AG provides a microscopic traffic simulation with state-of-the-art longitudinal dynamics and lane change models. The software MATSim is mainly developed at TU Berlin and ETH Zurich and is a multi-agent transport simulation tool based on queue models. It is capable of simulating large scale traffic networks and computing traffic assignment using an iterative approach. See [14] for more details on this tool.

   Quite surprisingly both simulation tools VISSIM and MATSim, despite their completely different approaches, produce very similar results. Average travel times vary less than 5% and the overall picture of congestion is nearly identical. As a consequence, the coordinations obtained by our optimization seem to be robust to disturbances as both simulations rate them equally.

   The simulation process with these two tools shows the applicability of our model for offset optimization and traffic assignment. In our solutions 'green waves', i.e., progressive signal systems, are indeed created and traffic is assigned to well coordinated arterial roads. The optimized coordination features lower travel times and less congestion compared, for example, to random coordinations. This suggests that our model assumptions of flow independent transit times are not too simplifying.

coordination	travel time	rel. difference
average random	57900 s	-
best random	54600 s	−6%
optimized	44100 s	−24%

coordination	delay
average random	30.1
best random	21.9
present	16.6
OPTIMIZATION (10 s)	16.4
platoon model (1 s)	16.1
TRANSYT (800 s)	15.9

■ **Table 2** Coordination and traffic assignment in Cottbus.

■ **Table 3** Simulation results for Portland (VISSIM).

Since we have no coordination data of all of the considered cities that only use one uniform cycle time and are not influenced by public transport, we also compare the optimized set of offsets to sets of randomly chosen offsets. Out of these random coordinations we chose the simulation result of the best of these random offsets. This best simulation result with respect to the overall travel time gives an impression of what can be obtained in an arbitrary or at least not very involved signal coordination. For all scenarios our offset optimization yields an improvement of 10 to 30 % compared to this best guess. Table 2 shows a typical example of simulation results for Cottbus (10 commodities with 0.2 cars per second in average, cycle time 90 seconds, 100 random offsets, 20 simulation runs per coordination).

5.3   Comparison to established approaches

To compare our model to previous approaches we use the results of [21] for the Portland scenario. In this experiment route choice is disabled, i.e., we compare only the performance of coordination. The benchmark is set by TRANSYT, a genetic programming tool for coordination, and the platoon model of [21]. Table 3 summarizes the results. The running time of the particular algorithm is given in brackets.

For coordination without the assignment our approach yields competitive solutions. The variances most likely result from the differences of the traffic model. TRANSYT uses a model with dynamics similar to the model in the simulation; the platoon approach and our approach make much more simplifying assumptions. Furthermore, we did not use a fine-tuned calibration of travel times in our model yet. This will probably yield even less delay. The influence of overestimating or underestimating the travel times will be investigated in upcoming experiments.

5.4   Advantages of simultaneous traffic assignment

One of the main intentions of this model is feedback between traffic assignment and coordination during the optimization process. We demonstrate the impact of our approach by the help of a small example. Given two parallel routes between two intersections A and B with two additional intersections on each route, we aim to find an optimal assignment for two commodities (A to B and B to A) and an optimal coordination for the six traffic signals. We assume identical operating sequences for each traffic signal (cycle time 60 s, green for 27 s) and a travel time of 20 seconds between consecutive intersections. The network is presented in Figure 9.

For initially computing a static traffic assignment we assume that the same strictly convex link-performance function is given for each link in the network. A conventional approach
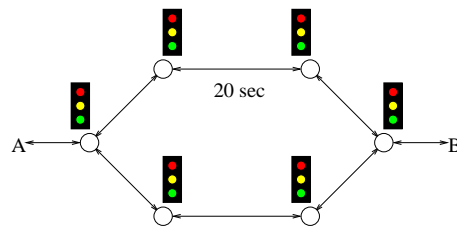
■ **Figure 9** Small example for simultaneous routing and coordination.

would first determine the assignment by the help of these functions. Convexity yields a fifty-fifty split for each commodity between upper and lower path in this simple network. Therefore, we fix this split and optimize the coordination with our model and compare this to the results of the simultaneous optimization.

For the fifty-fifty split there are a lot of conflicts with the traffic in opposite directions. The simulation yields an average waiting time of 25.2 seconds for the best coordination.

In contrast, the solution of the simultaneous optimization is very different. The commodity from A to B is completely assigned to the lower path, while the commodity from B to A is completely assigned to the upper path. This avoids all conflicts with opposite traffic and allows perfect 'green waves'. The average waiting time is reduced to 6.1 seconds and occurs only at the first traffic signal where cars are assumed to appear randomly. The simultaneous assignment and coordination reduces waiting time by 75 percent in this simple scenario. The effect of separating traffic flows to force 'green waves' is also observed in realistic scenarios like Cottbus, although the effect on the waiting times is not as dramatic as in the constructed network.

These completely different but high-performance assignments characterize the presented model. We believe that similar results can not be obtained without simultaneously considering assignment and coordination, i.e., a highly-adaptive traffic signal optimization can not find a competitive solution. These results also suggest that traffic signal coordination should be used to actively route or redirect traffic.

5.5   Solving the model

For practical applications a model should be quickly solvable. Due to the time-expansion and the decision variables, the MIP-formulation of our model is rather large. Table 4 shows the size of the Cottbus scenario with its 30 traffic signals controlled intersections, 14 commodities, and a cyclic expansion with 90 time steps ($\Gamma = 90$ seconds) and the corresponding mixed integer program (MIP). Additionally, the offset of one traffic signal is fixed to avoid symmetry.

Solving the mixed integer programs with CPLEX produces some expected and some quite surprising effects. Optimal solutions are obtained very fast. An optimal assignment and coordination for the regular shaped American cities is often calculated in less than 10 seconds (depending on commodities and load). Unfortunately it takes much longer to prove optimality, i.e., to close the gap to the dual bound. The dual bound increases slowly while the primal solution remains unchanged. In scenarios with many conflicts between commodities with opposite directions the gap can not be closed by CPLEX at all. In some scenarios a gap of up to 30 % remains even after hours of computation. Still, the obtained primal solutions show very good performance in the simulation. On the other hand even slight disturbances like changing the demand of a commodity or changing a signal plan a little bit may yield a

	Number of	
Network:	nodes	15570
	arcs	31410
	decision variables	2700
MIP:	variables	80211
	constraints	159821
	non-zeroes	625167

■ **Table 4** Size of the expanded Cottbus network and its MIP.

complete different behaviour of the solver and may significantly influence the running time.

Obviously it is very hard to compute good dual bounds. This observation is supported by data from the Cottbus scenario. The instance in Table 4 was solved using CPLEX. We obtained a primal solution with objective function of value 44100. The best dual bound stayed at 37025 after one day of computation. To rate the relevance of this dual solution we compare it to two trivial bounds. Calculating the length of a shortest path for each commodity times the demand of this commodity yields a dual bound of 33390. Calculating an assignment without traffic signals in the unexpanded network yields a dual bound of about 35180. Thus, the dual bound was not significantly increased by CPLEX. Therefore, our next steps will focus on the improvement of the formulation of the mixed integer program to speed up the computation of good lower bounds.

6   Discussion

The main advantage of our model is the simultaneous description of the traffic assignment problem and the traffic signal coordination problem in an uncomplex linear mixed integer programming formulation. The interactions between road users and the traffic signal coordination are taken into account and provides considerable potential for better solutions. Furthermore, our model allows for an efficient solution of the traffic assignment problem when the signal coordination is fixed.

The simulation results suggest the applicability of our model in real-world traffic optimization frameworks. In spite of the size of the MIP-formulation good solutions are obtained fast. Solving this mixed integer program with a common solver also yields a dual bound for the solution. This guaranteed maximal gap to optimality allows to evaluate the quality of the found solution and thus is a great advantage over heuristic approaches with genetic programming or neural networks. Our next steps especially focus on the improvement of the dual bounds. An improvement here will lead to better performance of the solver.

7   Acknowledgment

References

1   J. Adolph. *Regelung des Wagenverkehrs in Straßen mittels elektrischer Signallampen.* Reichspatentamt, 1925. DE 439255 A.

2   R. K. Ahuja, T. L. Magnanti, and J. B. Orlin. *Network flows.* Prentice Hall, 1993.

3   R. E. Allsop and J. A. Charlesworth. Traffic in a signal-controlled road network: An example of different signal timings inducing different routeings. *Traffic Eng. Control*, 18(5):262–265, 1977.

4   M. C. Bell and R. D. Bretherton. Ageing of fixed-time traffic signal plans. In *2nd International Conference on Road Traffic Control*, London, 1986.

5   M. G. H. Bell and H. Ceylan. Traffic signal timing optimization based on genetic alorithm approach, including drivers' routing. *Transportation Research Part B*, 38:329–342, 2004.

6   S. W. Chiou. *Optimization of Area Traffic Control for Equilibrium Network Flows.* PhD thesis, University College London, 1999.

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

8   M. R. Garey and D. S. Johnson. *Computers and Intractability; A Guide to the Theory of NP-Completeness.* W. H. Freeman & Co., New York, 1979.

9   N. H. Gartner, J. D. C. Little, and H. Gabbay. Optimization of traffic signal settings by mixed-integer linear programming, Part I: The network coordination problem. *Transportation Sciences*, 9:321–343, 1966.

10   E. Köhler, K. Langkau, and M. Skutella. Time-expanded graphs with flow-dependent transit times. In R. H. Möhring and R. Raman, editors, *Proceedings ESA 2002*, LNCS 2461, pages 599–611. Springer, 2002.

11   E. Köhler, R. H. Möhring, and M. Skutella. Traffic networks and flows over time. In J. Lerner, D. Wagner, and K. A. Zweig, editors, *Algorithmics of Large and Complex Networks*, LNCS 5515, pages 166–196. Springer, 2009.

12   S. Lämmer. *Reglerentwurf zur dezentralen Online-Steuerung von Lichtsignalanlagen in Straßennetzwerken.* PhD thesis, Technische Universität Dresden, 2007.

13   J. D. C. Little. The synchronizing of traffic signals by mixed-integer linear programming. *Operations Research 14*, pages 568–594, 1966.

14   MATSim Homepage. `http://www.matsim.org/`.

15   J. T. Morgan and J. D. C. Little. Synchronizing traffic signals for maximal bandwidth. *Journal of the Operations Research Society of America*, 12(6):896–912, 1964.

16   T. Roughgarden. *Selfish Routing and the Price of Anarchy.* MIT Press, 2005.

17   D. Schrank and T. Lomax. 2009 Urban Mobility Report. Texas Transportation Institut, 2009.

18   P. Serafini and W. Ukovich. A mathematical model for periodic scheduling problems. *SIAM Journal on Discrete Mathematics*, 2(4):550–581, 1989.

19   D. Sun, R. F. Benekohal, and S. T. Waller. Multiobjective traffic signal timing optimizing using non-dominated sorting genetic algorithm. In *IEEE IV2003 Intelligent Vehicles Symposium*, 2003.

20   F. Teklu, A. Sumalee, and D. Watling. A genetic algorithm approach for optimizing traffic control signals considering routing. *Computer-Aided Civil and Infrastructure Engineering*, 22:31–43, 2007.

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