wctr'04
4-8 july 2004 / istanbul-turkey

PROGRESSION OPTIMIZATION FEATURING ARTERIAL- AND ROUTE-BASED PRIORITY NETWORKS

Nathan H. Gartner and Chronis Stamatiadis
Department of Civil & Environmental Engineering, University of Massachusetts, Lowell, MA 01854, U.S.A.
Nathan_Gartner@uml.edu

Abstract

Arterial progression schemes based on the bandwidth criterion are widely used for traffic signal optimization. The schemes provide robust plans for traffic control as well as a variety of design options that can be tailored to specific network and traffic conditions. In recent years arterial progression optimization was also extended to grid networks. These models use advanced mathematical programming software packages which are computationally demanding when applied to large-scale networks. This paper describes procedures that dramatically improve the computability of such models and bring them into the realm of real-time application. The procedures are based on, first, selecting and optimizing an *arterial priority network* or a *route priority network*. These results are then used in a subsequent stage to determine an optimal plan for the entire network. The procedure is applicable to both uniform- and variable- bandwidth optimization and can accelerate computation by two orders of magnitude, *ceteris paribus*. This facilitates optimization of large-scale urban networks, provides a capability to analyze many design options and is also amenable for real-time implementation.

Keywords: Arterial priority network; Route priority network; Optimization
Topic Area: C3 Traffic Control

1. Introduction

Arterial progression schemes have been a perennial choice for signal optimization since the advent of co-ordinated traffic signal systems. Maximizing the width of the green progression bands provides the opportunity for continuous movement of platoons of vehicles through successive traffic lights on the arterial streets of the network. The schemes also offer a number of design features that are not readily available in other signal optimization models. Such features include optimal phase sequencing, advance queue clearance, progression speed adjustment and disaggregate bandwidth weighting. They enable to tailor the control plan to match the specific characteristics of each arterial street. Progression optimization has proven to be a flexible and robust design for traffic signal control (Gartner and Stamatiadis, 2002).

There are two principal categories of progression optimization: one produces uniform bandwidths, the other produces variable bandwidths. Two well-known programs in the first category are *MAXBAND* and *PASSER II*, both of which provide offset, split, cycle length and left-turn phase sequence optimization on individual arterial streets. *MAXBAND* (Little *et al*, 1981) uses a rigorous mathematical programming formulation for the maximization of a weighted combination of the bandwidths in the two directions of the arterial. The formulation contains integer decision variables and, therefore, mixed-integer linear programming (*MILP*) is used for solving the problem. The single arterial *maxband*[1] formulation was later extended to grid network optimization which considerably expands the size of the integer variable set (Chang *et al*, 1988a). *PASSER II* uses a search procedure to determine the combination of offsets that will result in the widest equal bands in both directions of an artery (Chang *et al*, 1988b).

Uniform bandwidth models have a basic limitation: they apportion the total available bandwidth along each arterial in proportion to the average volumes in each direction. Such models can provide optimal progressions only when the platoons maintain a constant size. In practice, however, traffic volumes may vary significantly along the arterial due to turn-in and turn-out traffic at each intersection. Consequently, the size of the platoon of vehicles traveling through a sequence of intersections is not, generally, constant in which case the basic assumption of conventional progression models does not hold. The effect of using average through-moving volume for apportioning the total bandwidth is that the green band may be either wasted at intersections with lower than average though-moving volume, or be deficient at intersections with higher than average though-moving volume.

The *multi-band* model, which is an extension of *maxband*, was designed to overcome this deficiency. It calculates a different bandwidth for each directional link of the arterial while maintaining main street platoon progression. The individual bandwidths depend on the actual traffic volumes that each link carries and the resulting signal-timing plan is tailored to the varying traffic flows along the arterial. The *multi-band* model is also formulated as a mixed-integer linear programming problem. A single arterial model was developed initially (Gartner et al., 1991), followed by a network version (Stamatiadis and Gartner, 1996). The *multi-band* design has been shown to provide significant benefits in terms of common performance measures such as delays, number of stops and fuel consumption over conventional uniform bandwidth models. The *multi-band* formulation increases the number of continuous variables and constraints by 20-60% for common size problems compared with the uniform band formulation. This increases running time proportionately.

The primary difficulty in solving these models, however, is due to the size and range of the integer variable set which is identical for both models. This paper presents an acceleration procedure that dramatically improves computability for both the uniform and the variable bandwidth network progression problem. The procedure is based on the traffic characteristics of the network. First, an *arterial priority network* (APN) or a *route priority network* (RPN) is determined and optimized. Results are then used in a second stage to determine an optimal plan for the entire network. This procedure facilitates the accelerated determination of the optimal values for the integer variables which permits the efficient optimization of larger-scale networks as can be found in larger cities and metropolitan areas. The solution that is obtained is equal to, or nearly equal to, the globally optimal solution obtained by solving the complete model. While the procedure can be applied to both the uniform- and variable-bandwidth optimization models, the latter has a clear advantage in terms of network performance. The speed and the quality of the developed plans allows also for on-line implementation of the method in suitably configured control systems (Gartner *et al*, 1995).

[1] Lower case italics denote a model; upper case italics denote a computer program.

2. Description of the model

This section presents the formulation of the network *multi-band* optimization model. The model consists of several blocks of constraints dealing with the individual arterials of the network as well as a set of network constraints that ensures that continuous bands are being produced on all the intersecting arterials. To explain the variable-band formulation, the basic uniform bandwidth optimization model is presented first.

The geometric relations for the uniform bandwidth model are shown in Figure 1. Consider a network with m arterials where each arterial j has n_j signalized intersections. Let S_{ij} denote the i^{th} signal on the j^{th} arterial of the network and L_{ij} denote the i^{th} link (between signals i and i+1) of the j^{th} arterial, with j = 1 ,...,m and i = 1 ,..., n_j. All time variables are defined in units of the cycle time. The following variables are defined:

C =	cycle time (sec);
b_j ($\bar{b}_j$) =	outbound (inbound) bandwidth on arterial j;
r_{ij} ($\bar{r}_{ij}$) =	outbound (inbound) red time at S_{ij};
w_{ij} ($\bar{w}_{ij}$) =	interference variables, time from right (left) side of red at S_{ij} to left (right) side of outbound (inbound) green band;
t_{ij} ($\bar{t}_{ij}$) =	travel time on link i of arterial j in the outbound (inbound) direction;
$\phi_{(ij),(kl)}$ ($\bar{\phi}_{(ij),(kl)}$) =	*internode* offsets, time from the center of the outbound (inbound) red at S_{ij} to the center of the outbound (inbound) red at S_{kl};
Δ_{ij} =	directional node phase shift, time from center of $\bar{r}_{ij}$ to nearest center of r_{ij};
τ_{ij} ($\bar{\tau}_{ij}$) =	queue clearance time for advancement of outbound (inbound) bandwidth at S_{ij} to clear turning-in traffic before arrival of main-street platoon;
v_{ij} ($\bar{v}_{ij}$) =	outbound (inbound) progression speed on link L_{ij} (ft/sec).

In the case of uniform bands, the objective function has the following form:

$$\text{Maximize} \quad \sum_{j=1}^{m}(b_j + k_j \cdot \bar{b}_j) \qquad (1)$$

where k_j is the target ratio of inbound to outbound bandwidth for arterial j.

The directional interference constraints ensure that the progression bands use only the green time and they do not cross through the red time. When the band has a fixed width throughout the arterial there is only one such constraint needed for each signal S_{ij} and each directional band:

$$w_{ij} + b_j \leq 1 - r_{ij} \qquad (2.a)$$

$$\bar{w}_{ij} + \bar{b}_j \leq 1 - \bar{r}_{ij} \qquad (2.b)$$

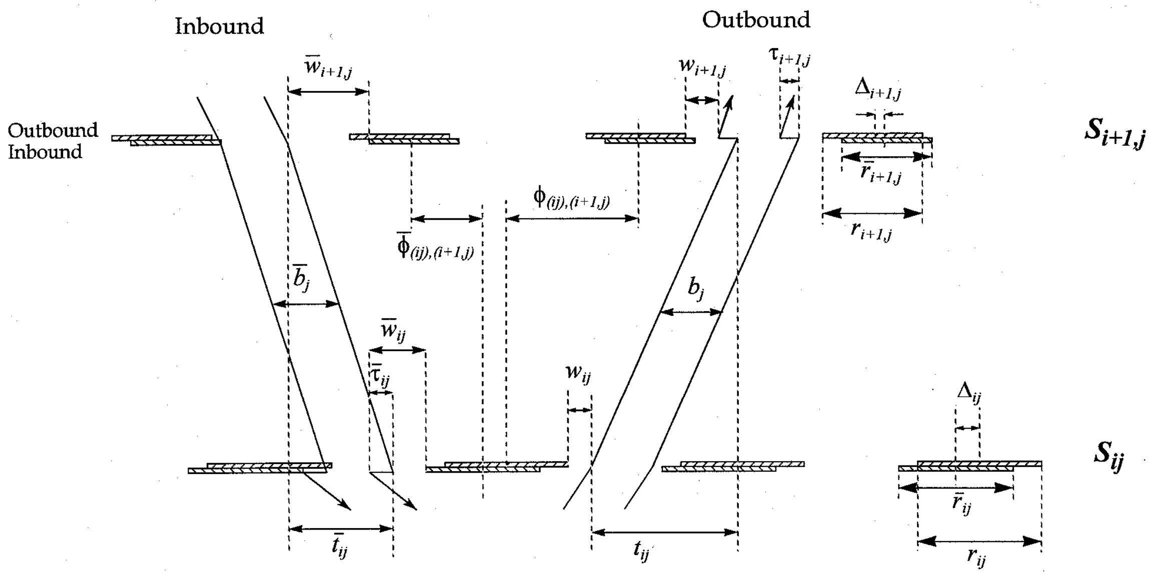


Figure 1: Time-space diagram for uniform bandwidth optimization

The arterial loop constraints result from the fact that all signals must be synchronized, i.e., that they operate with a common cycle time. In Figure 1 it can be seen that for each link L_{ij} the summation of the *internode offsets* and *directional node phase shifts* is an integer multiple of the cycle time as follows:

$$\phi_{(ij),(i+1,j)} + \bar{\phi}_{(ij),(i+1,j)} + \Delta_{i+1,j} - \Delta_{ij} = \kappa_{ij} \tag{3}$$

where κ_{ij} is an integer variable. The same principle of signal synchronization applies to closed loops of the network consisting of more than 2 links, resulting in the *network loop constraints*. For simplicity, we drop the arterial index in the notation of nodes and *internode offsets* and we define the *intranode offset* Φ_{ijk} as the time from the center of the red at S_j for traffic moving from S_i to S_j, to the center of red in the crossing direction at the same node for traffic moving from S_j to S_k (Figure 2.a). *The network loop constraints* specify that the summation of *internode* and *intranode offsets* around a loop of intersecting arterials must be an integer multiple of the cycle time (Figure 2.b):

$$\phi_{ij} + \omega_{ijk} + \phi_{jk} + \omega_{jkl} + \phi_{kl} + \omega_{kli} + \phi_{li} + \omega_{lij} = \mu_n \tag{4}$$

where μ_n is the integer variable of the n^{th} network loop. The number of network loop constraints and the choice of a fundamental set of loops are given by Gartner (1972a).

The cycle time C (sec) and the link specific progression speeds v_{ij} and $\bar{v}_{ij}$ are treated as decision variables. This feature introduces considerable flexibility in the calculation of the best progression scheme. Each variable must be constrained by upper and lower bounds as follows:

C_1, C_2 =	lower and upper bounds on cycle length;
(e_{ij}, f_{ij}), $(\bar{e}_{ij}, \bar{f}_{ij})$ =	lower and upper bounds on outbound (inbound) speed v_{ij} ($\bar{v}_{ij}$) (ft/sec);

$(g_{ij}, h_{ij}), (\bar{g}_{ij}, \bar{h}_{ij}))$ = lower and upper bounds on change in outbound (inbound) speed v_{ij} ($\bar{v}_{ij}$) (ft/sec).

The corresponding constraints will be:

$$C_1 \leq C \leq C_2 \tag{5}$$

$$e_{ij} \leq v_{ij} \leq f_{ij} \quad \text{and} \quad \bar{e}_{ij} \leq \bar{v}_{ij} \leq \bar{f}_{ij} \tag{6}$$

$$g_{ij} \leq v_{i+1,j} - v_{ij} \leq h_{ij} \quad \text{and} \quad \bar{g}_{ij} \leq \bar{v}_{i+1,j} - \bar{v}_{ij} \leq \bar{h}_{ij} \tag{7}$$

An important decision capability that can is afforded by the *MILP* formulation of the problem is the determination of the optimal left-turn phase sequence with respect to the through green at any signal S_{ij}. A left-turn green can be chosen to lead or lag the through green, whichever results in the most total bandwidth. Figure 3 shows four possible patterns of left-turn phases, where l_{ij} ($\bar{l}_{ij}$) is the outbound (inbound) green time for left-turning traffic at S_{ij}. The *intranode offsets* can be expressed in terms of l_{ij} and $\bar{l}_{ij}$ as follows:

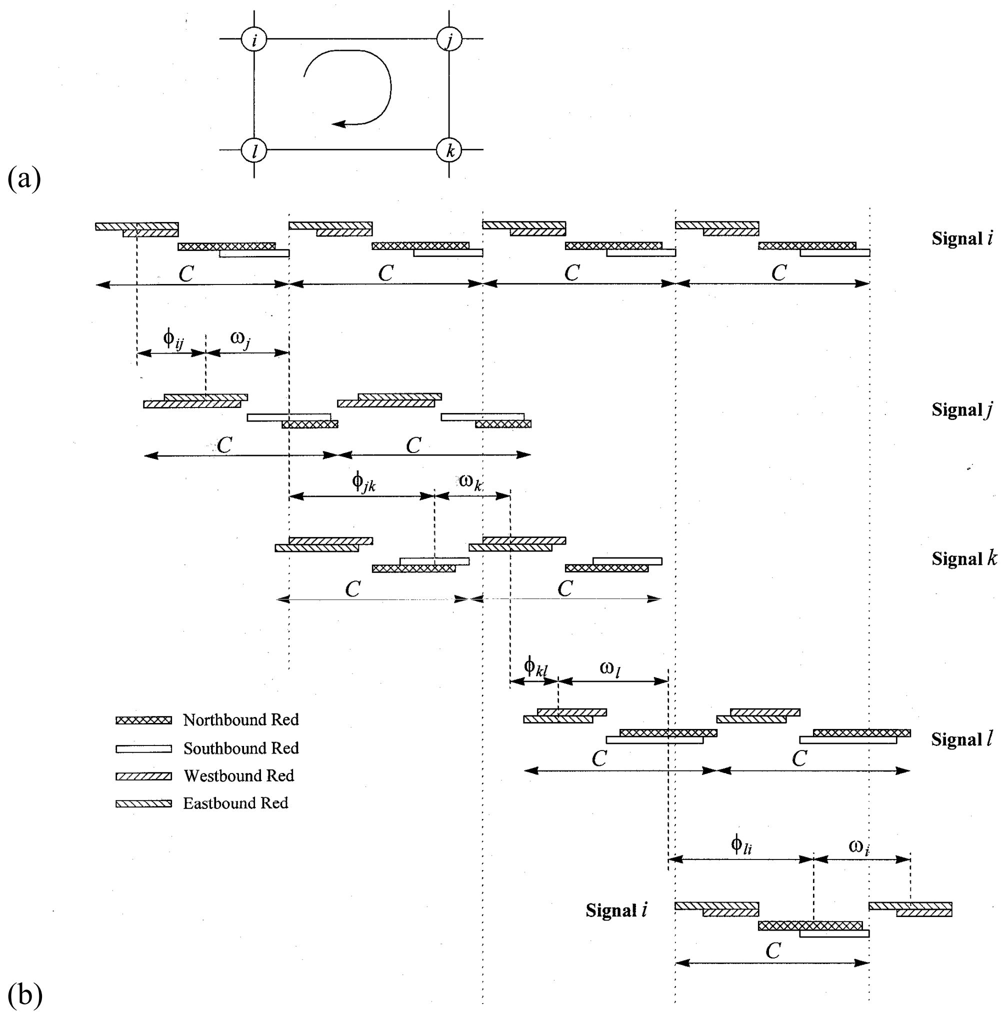


Figure 2: (a) Closed loop of four intersecting arterials and (b) the geometry of the network loop constraint.

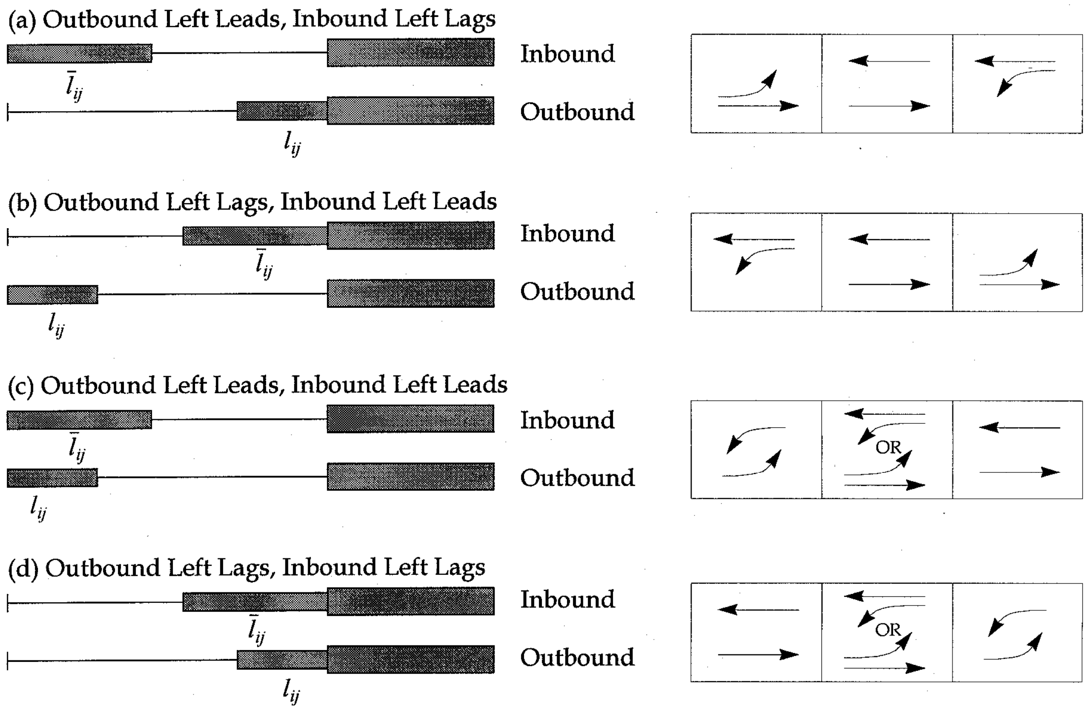


Figure 3: The four different phase sequences.

Pattern 1: $\Delta_{ij} = -(l_{ij} + \bar{l}_{ij})/2$

Pattern 2: $\Delta_{ij} = (l_{ij} + \bar{l}_{ij})/2$

Pattern 3: $\Delta_{ij} = -(l_{ij} - \bar{l}_{ij})/2$

Pattern 4: $\Delta_{ij} = (l_{ij} - \bar{l}_{ij})/2$

By introducing two binary decision variables δ'_{ij} and δ''_{ij} we can coalesce the four patterns into a single equation:

$$\Delta_{ij} = ((2\delta'_{ij} - 1) \cdot l_{ij} - (2\delta''_{ij} - 1) \cdot \bar{l}_{ij})/2 \qquad (8)$$

The binary variables δ'_{ij} and δ''_{ij} are defined as follows:

Pattern	δ'_{ij}	δ''_{ij}
1	0	1
2	1	0
3	0	0
4	1	1

The traffic engineer may specify that only some of these patterns are allowable, in which case additional constraints are imposed on the combinations of allowable values of δ'_{ij} and δ''_{ij}.

In the *multi-band* model the width of the directional bands may differ from link to link. The bandwidth can be individually weighted with respect to its contribution to the overall objective function. By introducing weights that are computed based on the directional volume on each link, a method is obtained that is sensitive to the varying traffic conditions along each arterial of the network. The link specific bands generated by *MULTIBAND* are symmetric about the centerline of the arterial progression band. The geometry of the *multi-band* model is shown in Figure 4. The bandwidths and the interference variables are redefined as follows:

$b_{ij}\ (\bar{b}_{ij}) =$ outbound (inbound) bandwidth of link i on arterial j; there is now one specific band for each link *Lij*;

$w_{ij}\ (\bar{w}_{ij}) =$ the time from right (left) side of red at *Sij* to the centerline of the outbound (inbound) green band; the reference point at each signal has been moved from the edges to the centerline of the band.

The objective function now has the form:

$$\text{Maximize} \quad \sum_{j=1}^{m} \frac{1}{n_j - 1} \sum_{i=1}^{n_j - 1} (\alpha_{ij} \cdot b_{ij} + \bar{\alpha}_{ij} \cdot \bar{b}_{ij}) \tag{9}$$

where □$_{ij}$ and $\bar{\alpha}_{ij}$ are the link specific weighting coefficients for the outbound and inbound directions respectively. The weighting coefficients in Eq. (9) can be chosen to fulfill desirable performance criteria. The coefficients currently used are as follows:

$$\alpha_{ij} = \left(\frac{V_{ij}}{s_{ij}}\right)^p \quad \text{and} \quad \bar{\alpha}_{ij} = \left(\frac{\bar{V}_{ij}}{\bar{s}_{ij}}\right)^p \tag{10}$$

where . $V_{ij}\ (\bar{V}_{ij}) =$ outbound (inbound) directional flow rate on link L_{ij}; either the total or the through volume can be used;

$s_{ij}\ (\bar{s}_{ij}) =$ saturation flow rate outbound (inbound) for link L_{ij}; either the total flow rate or the through flow rate can be used;

$p =$ integer exponent; values of 0, 1, 2 and 4 are used

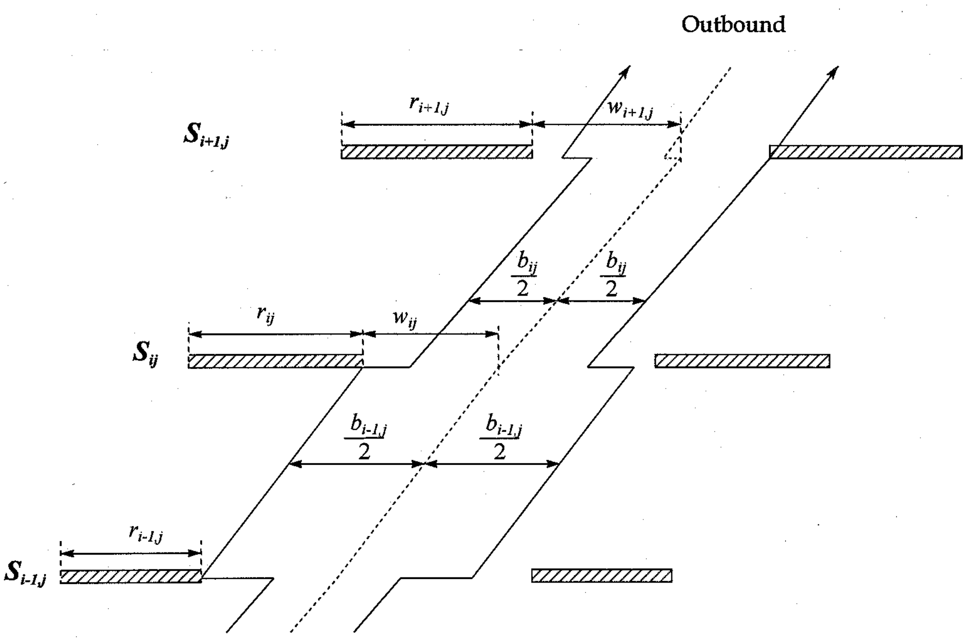
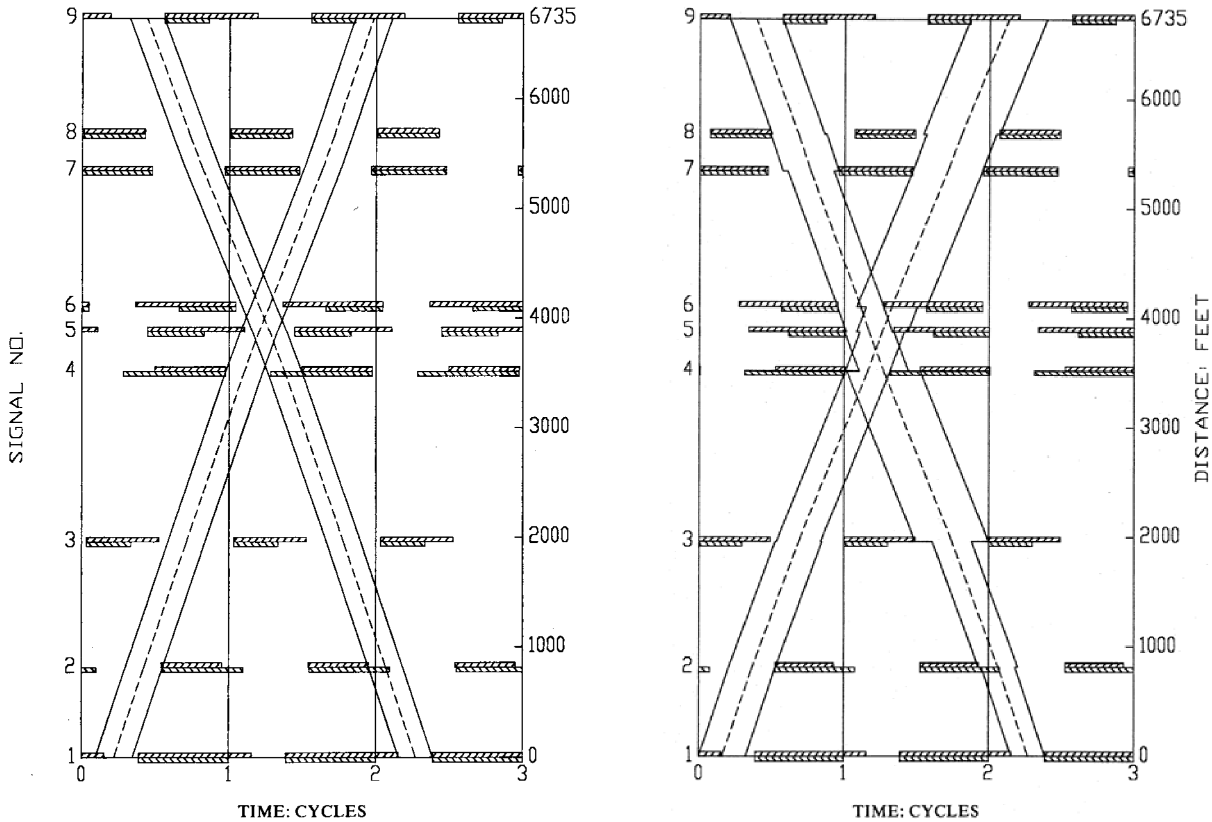


Figure 4: Geometric relations for the variable bandwidth optimization model.



(a) (b)

Figure 5: Uniform (a) and variable (b) bandwidth schemes.

In case of variable bandwidths, each band segment must be constrained from both sides so that neither edge of the band crosses the red time. For each signal S_{ij} and for each link-specific directional band there are two interference constraints, as follows:

$$w_{ij} + b_{ij}/2 \leq 1 - r_{ij} \quad \text{and} \quad w_{ij} + b_{ij}/2 \geq 0 \tag{11.a}$$

$$\overline{w}_{ij} + \overline{b}_{ij}/2 \leq 1 - \overline{r}_{ij} \quad \text{and} \quad \overline{w}_{ij} + \overline{b}_{ij}/2 \geq 0 \tag{11.b}$$

The same relationship must be valid at both ends of the band, i.e., at signals S_{ij} and $S_{i+1,j}$:

$$w_{i+1,j} + b_{ij}/2 \leq 1 - r_{ij} \quad \text{and} \quad w_{i+1,j} + b_{ij}/2 \geq 0 \tag{11.c}$$

$$\overline{w}_{i+1,j} + \overline{b}_{ij}/2 \leq 1 - \overline{r}_{ij} \quad \text{and} \quad \overline{w}_{i+1,j} + \overline{b}_{ij}/2 \geq 0 \tag{11.d}$$

The constraints given by equations (3) - (8) are unaffected in the variable bandwidth reformulation.

By introducing increased flexibility in the design of progression schemes, the *multi-band* approach results in significant improvements in all performance measures. It produces optimal variable progression schemes tailored to both the demand and the capacity of each individual road section along each arterial street, while simultaneously optimizing progressions on the crossing arterials as well. Table 1 shows results of simulation studies for the downtown network

Table 1: Comparison of Network Performance for *MULTIBAND* vs. *MAXBAND*

Model	**Weighting Coeff.**	**Avg. Delay (sec./veh.)**	**Avg. % of Stops**	**Avg. Speed (m.p.h)**
MAXBAND	1	26.64	53.20	9.40
	AVR[a]	25.52	53.04	9.89
MULTIBAND	V/C [b]	22.93	50.02	10.23
		(-10.1%)	(-5.7%)	(3.4%)
	$(V/C)^2$	22.93	50.02	10.23
		(-10.1%)	(-5.7%)	(3.4%)
	$(V/C)^4$	23.59	51.41	10.14
		(-7.6%)	(-3.1%)	(2.5%)

a. Average directional volume ratio;
b. Volume to capacity ratio.

of Ann Arbor, Michigan comparing *MAXBAND* and *MULTIBAND* designs (Stamatiadis and Gartner, 1996). Different bandwidth weighting coefficients were examined for both models and in all cases the variable bandwidth model improved considerably the network performance. Examples of uniform- and variable-bandwidth progressions for an arterial street are illustrated in Figure 5. The examples demonstrate the ability of the programs to optimize phase sequences in a way that maximizes the widths of the bands in each direction. This is a unique feature of *MAXBAND* and *MULTIBAND* that is made possible by the discrete decision capabilities of the integer programming codes. The Ann Arbor network will be used again as an example for the application of the network acceleration procedure and is shown in Figure 8 below.

3. Network acceleration procedure

Both the *maxband* and *multi-band* models use mixed-integer linear programming for determining the optimal solution. While there are now available quite efficient codes for solving large-scale problems off-line, it is incumbent on all modelers, especially in traffic engineering applications, to devise the most efficient solution procedure for a given problem. More efficient computational procedures result in the ability to obtain improved solutions and, ultimately, lead to improved performance of the traffic network. They can be as important as more accurate traffic models are, especially in an era of increasing real time applications. In the case of progression optimization in networks the advanced procedures enable to:

- use more economical mathematical programming codes which are more affordable and more easily accessible;
- optimize larger-scale networks than would otherwise be possible;
- assure more reliable convergence to an optimal solution, i.e., the procedure would not fail as often;
- analyze a larger number of alternatives at a much reduced cost (the design of progression schemes involves multiple data sets, choice of coefficients and parameter ranges); and

- use the codes in real time, e.g. in a multi-level RT-TRACS system, due to the accelerated decision capabilities of the codes (Gartner *et al*, 1995).

Most *MILP* codes use branch-and-bound strategies for determining the optimal values of the discrete variables (Nemhauser and Wolsley, 1988). Branch-and-bound is a strategy of "divide and conquer": the objective is to partition the feasible region into more manageable subdivisions and, in this way, to fathom the tree of integer solutions. The strategy is designed for general purpose mixed-integer programs and does not take advantage of special characteristics that a particular problem may have. Solving the complete progression optimization problem by general purpose branch-and-bound is a formidable task. On the other hand, heuristic methods which quickly lead to a good, though not always optimal, solution are often preferable. Heuristics have an intuitive justification motivated by an intimate familiarity with the particular problem characteristics. This is especially pertinent in the case of the traffic signal optimization problem and forms the basis for the acceleration procedure described in this paper.

For many *MILP* problems it is advantageous to partition the original problem into smaller sub-problems in order to reduce the number of integer variables that have to be considered in each sub-problem and to restrict as much as possible the allowable range for each variable. Several authors have proposed solutions for the uniform bandwidth problem along these lines. Mireault (1991) proposed to solve more efficiently an early version of the *MILP* progression model by carefully restricting the range of the integer variable set. Chaudhary *et al.* (1991) devised two decomposition procedures for the network version of MAXBAND. The procedures are based on dividing the integer variable set into two or three sub-sets consisting of (1) the arterial two-way loop variables $\square_{ij}$, (2) the network loop variables $\square_{n}$, and (3) the phase sequencing variables δ'_{ij} and δ''_{ij}. The procedures are based on calculating one of the sets while relaxing the integrality requirements on the other two. The results are then used to fix the values of the first set and to calculate, in turn, integer values for the other two sets. Six alternative feasible integer solutions are kept from which the best final solution is chosen. Figure 6 outlines the steps involved in the two heuristic methods developed by Chaudhary. Pillai *et al.* (1994) proposed a two part "greedy" heuristic for the uniform bandwidth network problem which, similar to Chaudhary's, involves partial relaxation of some integer variables coupled with a depth-first search of the branch-and-bound tree.

Two Step Heuristic

Optimize N with $\delta_i \in R^+$;

↓

Save κ_l, μ_L ; ($\kappa_l, \mu_L \in I^+$) for best six solutions

↓

Re-optimize N by setting κ_l, μ_L

↓

Select optimal solution

Three Step Heuristic

Optimize N with $\delta_i, \mu_L \in R^+$;

↓

Save κ_l; ($\kappa_l \in I^+$) for best six solutions

↓

Re-optimize N with $\delta_i \in R^+$ and by setting κ_l

↓

Save μ_L; ($\mu_L \in I^+$) for best six solutions

↓

Re-optimize N by setting κ_l, μ_L

↓

Select optimal solution

Figure 6: The two- and three-step heuristic developed by Chaudhary for MAXBAND

Both approaches are based on an arbitrary partition of the integer variable set and are not motivated by traffic-related considerations. By relaxing the integrality of some of the integer variables, intermediate solutions are not feasible and cannot be implemented in an actual signal system. The acceleration procedure that is described in this paper does not merely exploit the mathematical structure of the mixed-integer problem, it is inherently based on the specific traffic characteristics of the network. In addition, it does not require a modification of the mathematical programming code as is required by other approaches.

An *arterial priority network* (APN) or a *route priority network* (RPN) consisting of an arterial (or route) tree is selected from the original traffic network based on its geometry and on the prevailing traffic flows in the network. The priority sub-network is optimized first and the results are then used for the solution of the entire network. The goal in selecting a priority sub-network is to provide the best progression opportunities for the widest possible demands, i.e. the maximal sum of (artery-bandwidth-meters)x(vehicles/hr) for all routes in the sub-network. Thus, the selection of a priority sub-network is striving to include arterials and/or routes that carry the highest volumes. Two approaches can be employed. The first is to use a graph-theoretical approach in which the links that form the arterials and routes are selected according to the above criterion. This is similar to the approach used by Gartner (1972b and 1975) in determining priority networks for offset optimization. This approach can be labeled the *objective approach*, since it does not require familiarity with the network and, especially, can be employed when analyzing a future network which is still in planning stages. The second approach is to select arterials and routes that, similarly, strive to meet the above criterion but is based on familiarity with the network and on the experience of the user. This is a *subjective approach* that is commonly practiced by the traffic engineering community and is akin to the ‘expert system’ approach. This approach works well when the networks are small and the number of alternative choices is limited, or in larger networks with clearly identifiable priority routes. The second approach is employed in the examples described below.

Alternative sub-networks can be selected in an iterative manner if further improvements are desired. Most importantly, the prioritization scheme does not require relaxation of any integer variables during the optimization. Any intermediate solution, which is an optimal solution for the priority sub-network, is also a feasible solution for the entire network. Considering the fact that the priority sub-network carries the bulk of the traffic volumes in the network, this solution is in itself a good, close to optimal solution for the entire network that can be implemented if so desired. Moreover, this is only the first stage of a two-stage process that enables the achievement of faster and better solutions for the entire network. Experience shows that an optimal solution is obtained in the majority of cases during the first iteration. The procedure is described in detail below.

Let N be the original network including all the arterials/routes and N_i be a sub-network including only a subset of arterials/routes of *N*. We define the following two optimization problems:

- *P1* is a progression optimization problem for a sub-network N_i, such that the included arterials/routes form a “tree” without creating any network loops. *P1* contains arterial loop integer variables but no network loop integer variables;
- *P2* is a progression optimization problem for the entire network *N*, obtained by freezing a subset of the arterial loop integer variables to predetermined values.

Then network prioritization scheme is shown in Figure 7, and is described below:

- **Step 1**: Identify a new priority sub-network $N_i \Box N$
- **Step 2**: Optimize *P1* for N_i, and save the resulting values of the arterial loop integer variables $\Box_{ij}^{*}$ for all the links loops of N_i.
- **Step 3**: Optimize *P2* by setting the integer variables calculated in step 2 ($\Box_{ij} \Box \Box_{ij}^{*}$: $\Box_{ij} \Box P1$).
- **Step 4**: Calculate the objective function. If it is better than the previous solution, save it.
- **Step 5**: Stop if all priority sub-networks have been considered; otherwise go back to step 1.

The arterials (routes) contained in the “tree” should include the principal arterials (routes) of the network and can be chosen based on the following criteria:

1. Choose the principal arterial (or route) of the network to be the trunk of the tree and include only crossing arterials (or routes) in the sub-network;
2. The resulting tree should consist of the maximum number of arterial (or route) two-way links without forming any network loops.

The solution of both *P1* and *P2* can be obtained very quickly due to the reduced number of integer variables. Table 2 shows the number of integer variables in the original problem and in the two sub-problems *P1* and *P2* of the network prioritization approach for an $m \Box n$ closed grid network ($m \Box n$ intersections and $m+n$ arterials).

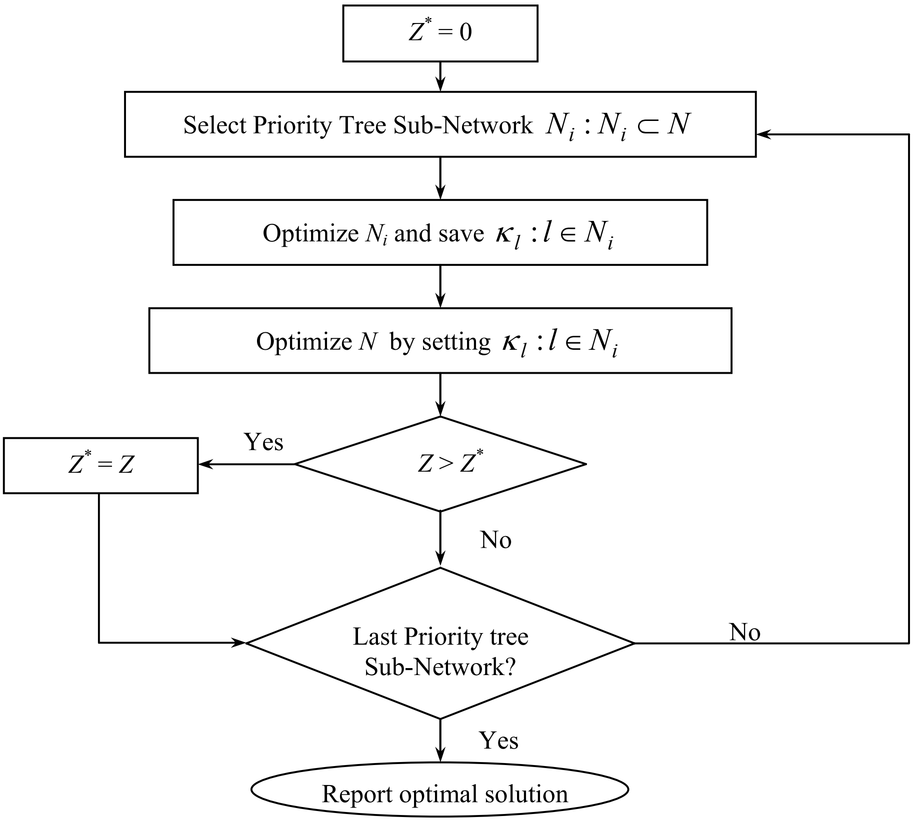


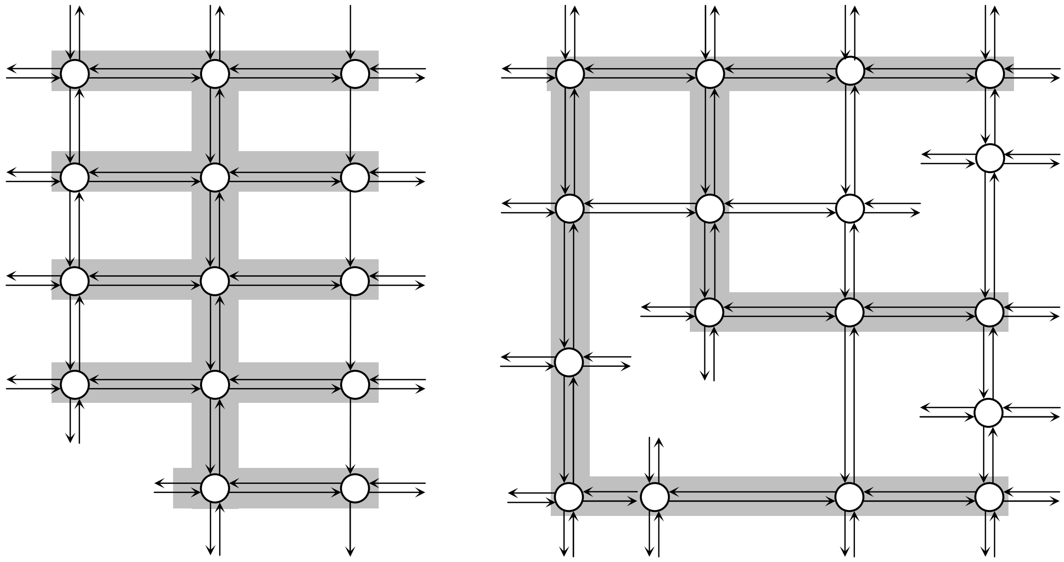
Figure 7: The network decomposition procedure

Table 2: Size of *MILP* Problem for an (*m* □ *n*) Closed Grid Network.

Variable	***MAXBAND***	***MULTIBAND***	**Network Acceleration Procedure**	
			***P1*: Priority Net.**	***P2*: Complete Net.**
$b, \bar{b}$	2(m+1)	2(m(n-1)+n(m-1))	2(mn-1)	2(2mn-m-n)
Z	1	1	1	1
$w, \bar{w}$	2(2mn-m-n)	2(2mn-m-n)	2(mn-1)	2(2mn-m-n)
□	(2mn-m-n)	(2mn-m-n)	mn-1	mn-m-n+1
□	(m-1)(n-1)	(m-1)(n-1)	0	(m-1)(n-1)
Total Integers	3mn-2m-2n+1	3mn-2m-2n+1	mn-1	2(mn-m-n)
Example (no. of integers)				
4x6 network	53	53	23	30
3x7 network	44	44	20	24

4. Compuational results

Application of the prioritization procedure is illustrated for three urban street networks. The first, in downtown Ann-Arbor, Michigan is a 3 × 5 grid containing 14 signals and 8 arterials as shown in Figure 8a. The second, in downtown Memphis, Tennessee is a 4 × 4 grid with 17 signals and also 8 arterials as shown in Figure 8b. The grids are not complete, as there are no



(a) (b)

Figure 8: The (a) Ann-Arbor, Michigan and the (b) Memphis, Tennessee networks.

Signals at each of the intersecting nodes. The shaded arterials indicate the arterial priority sub-networks utilized in this analysis. The computational results (objective function value and execution times) are given in Table 3. The values of the objective function of *MAXBAND* and *MULTIBAND* are not directly comparable. Execution times are reported for comparison purposes and are based on a 200MHz Pentium processor. The procedure was applied to both the uniform and the variable bandwidth models. Execution times were significantly reduced by as much as 1:263 compared with the original times. For the uniform bandwidth case the prioritization procedure calculated an optimal solution – there are multiple possible optimal solutions– for both sample networks. In the case of the *multi-band* model, the procedure located the globally optimal solution for only one of the two test networks. This is due to the fact that only one iteration was done in each case. Further iterations of the procedure of Fig. 7 are needed to approach the optimal solution. Nevertheless, in terms of traffic performance, the 77% *multi-band* solution is generally superior to the 100% *maxband* solution since the bandwidth criterion is not strictly commensurate with common traffic performance measures such as delay, travel time, etc., it only acts as a proxy.

The third network, shown in Figure 9, consists of an 11-node section of downtown Aachen (Germany) and illustrates the application of *route priority sub-networking*. The progressions that are being sought do not comprise only single arterials but may follow the prevailing path flows in the network along several arterials. This requires a modification of the optimization model to account for continuous progressions along routes in the network. Establishment of optimal progressions along routes (as opposed to arterials) is particularly advantageous when O-D demands are considered in a combined assignment and control problem and when rerouting of traffic is an option (see Gartner and Stamatiadis, 1998; Keller and Ploss, 1987; Ploss *et al*, 1990). In this way, improved travel times can be achieved for demands from origin to destination, rather than only along selected arterials. The development of route-based optimization models, in conjunction with origin-destination demand estimation, is the subject of much current work in ITS and is an essential building block of many advanced system designs.

Table 3: Objective Function Values and Execution Times of the Original and Accelerated Solutions.

Network/ *Model*	**Size**	**Original *MILP* Problem**		**Acceleration Procedure**	
	Art./Nodes	**Obj. Value**	**Exec. Time**	**Obj. Value**	**Exec. Time**
Memphis, TN	8/17				
MAXBAND		3.4682	6,735sec	3.4682 (100%)	50sec (1/135)
MULTIBAND		7.9381	15,012sec	7.9381 (100%)	57sec (1/263)
Ann Arbor, MI	8/14		*		
MAXBAND		2.9381	4,235sec	2.9381 (100%)	44sec (1/96)
MULTIBAND		4.8930	9,995sec	3.7680 (77%)	51sec (1/196)

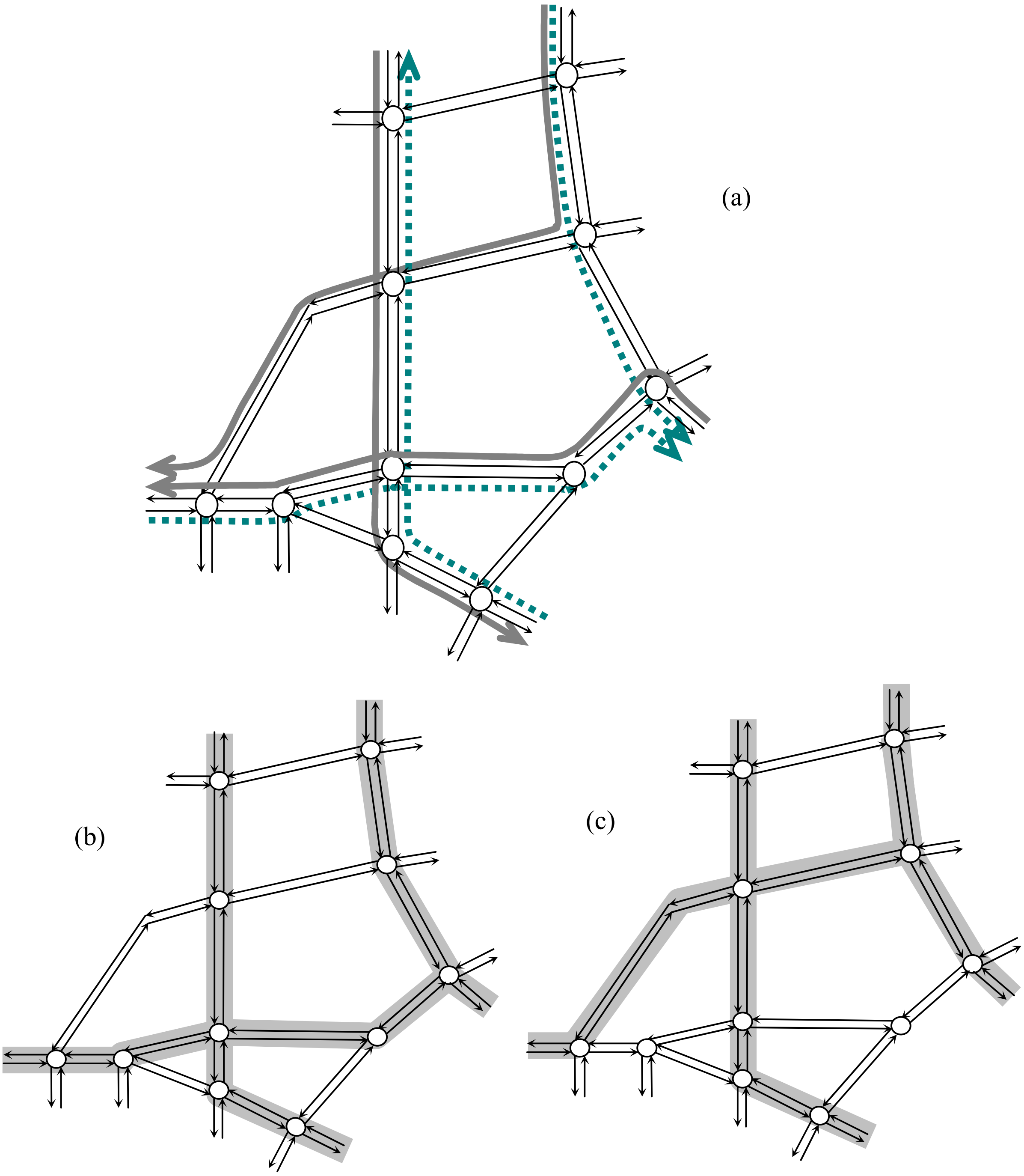


Figure 9: (a) Section of the downtown Aachen network indicating travel patterns; (b, c) Two possible route priority sub-networks.

5. Conclusions

This paper describes a new network prioritization procedure for the efficient solution of network progression problems using both uniform and variable bandwidths. The procedure is based on the traffic characteristics of the network and uses arterial as well as route priority sub-networks. It produces intermediate solutions that are feasible and practical and can be implemented in the control system at any stage of the computation. This is of particular benefit when the procedure is to be used in a dynamic traffic management system. The procedure yields considerable reductions in computational effort with little degradation in the quality of the results. Improvements in computation times range from 1:100 to 1:300 compared with the original formulation with the most dramatic reductions occurring in the case of the more complex problem of variable bandwidth optimization. By achieving these results one can obtain more easily optimal solutions for large-scale networks, analyze a larger number of alternatives, as well as implement this strategy in an on-line system. More efficient computational procedures result in the ability to obtain improved solutions and, consequently, lead to improved performance of the traffic network. Of particular importance is the ability to integrate this procedure with an O-D demand prediction model within a dynamic traffic management system.

References

Chang E., Cohen, S., Liu, C., Chaudhary, N., Messer, C., 1988a. MAXBAND-86: A program for optimizing left-turn phase sequence in multiarterial closed networks. Transportation Research Record 1181, 61-67.

Chang E., Lei, J.C., Messer, C.J., 1988b. Arterial Signal Timing Optimization Using PASSER II-87 Microcomputer User's Guide. TTI Research Report 467-1. Texas A&M University.

Chaudhary, N.A., Pinnoi, A., Messer, C., 1991. Proposed enhancements to MAXBAND-86 program. Transportation Research Record 1324, 98-104.

Gartner, N., 1972a. Constraining relations among offsets in synchronized signal networks. Transportation Science 6, 88-93.

Gartner N., 1972b. Algorithms for Dynamic Road-Traffic Control. Procs. Fifth IFAC World Congress, Theme 12: Transportation. Paris, France, June 1972. Published by Instrument Society of America, Pittsburgh, PA, USA.

Gartner N., 1975. Priority Network Formation and Optimization. Bull. Operations Res. Soc. of America, 13(1), Spring 1975.

Gartner, N.H., Assmann, S.F., Lasaga, F., Hou, D.L., 1991. A multi-band approach to arterial traffic signal optimization. Transportation Research 25B(1), 55-74.

Gartner, N.H., Stamatiadis, C., Tarnoff, P.J., 1995. Development of advanced traffic signal control strategies for IVHS: A multi-level design. Transpn. Res. Record 1494, 98-105.

Gartner, N. and Stamatiadis C., 1998. Integration of Dynamic Traffic Assignment with Real Time Traffic Adaptive Control. Transpn. Research Record 1644, TRB, 150-156.

Gartner, N. and Stamatiadis C., 2002. Arterial-Based Control of Traffic Flow in Urban Grid Networks. Mathematical and Computer Modelling 35, 657-671.

Keller, H. and Ploss, G., 1987. Real-Time Identification of O-D Network Flows from Counts for Urban Traffic Control. Transportation and Traffic Theory (Eds., N.H. Gartner and N.H.M. Wilson), Elsevier, New York.

Little, J.D.C., Kelson, M.D., Gartner, N.H., 1981. MAXBAND: A program for setting signals on arteries and triangular networks. Transportation Research Record 795, 40-46.

Mireault, P., 1991. A branch-and-bound algorithm for the traffic signal synchronization problem with variable speed. Procs. TRSTAN I, Universite de Montreal, Canada.

Nemhauser, G.L. and Wolsley, L.A., 1988. Integer and Combinatorial Optimization. Wiley-Interscience, New York.

Stamatiadis, C., Gartner, N.H., 1996. MULTIBAND-96: A program for variable bandwidth progression optimization of multi-arterial traffic networks. Transportation Research Record 1554, 9-17.

Pillai, R.S., Rathi, A.K., Cohen, S., 1994. A restricted branch-and-bound approach for setting the left turn phase sequence in signalized networks. Transportation Research 32B(5).

Ploss, G., Philipps, P., Inaudi, D. and Keller, H., 1990. Motion – A New Traffic Control Concept based on Real Time O-D Information. Transportation and Traffic Theory (Ed., M. Koshi), Elsevier, New York.