1. Report No. FHWA/TX-03/0-4020-P2	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle GUIDELINES FOR SELECTING SIGNAL TIMING SOFTWARE		5. Report Date September 2002
		6. Performing Organization Code
7. Author(s) Nadeem A. Chaudhary, Vijay G. Kovvali, and S. M. Mahabubul Alam		8. Performing Organization Report No. Report 0-4020-P2
9. Performing Organization Name and Address Texas Transportation Institute The Texas A&M University System College Station, Texas 77843-3135		10. Work Unit No. (TRAIS)
		11. Contract or Grant No. Project No. 0-4020
12. Sponsoring Agency Name and Address Texas Department of Transportation Research and Technology Implementation Office P. O. Box 5080 Austin, Texas 78763-5080		13. Type of Report and Period Covered Design Guidelines: September 2000 – August 2002
		14. Sponsoring Agency Code

15. Supplementary Notes
Research performed in cooperation with the Texas Department of Transportation and the U.S. Department of Transportation, Federal Highway Administration.
Research Project Title: Improvements to Signal Timing Software

16. Abstract
This report contains guidelines for selecting software for use in arterial signal timing projects. These guidelines are based on an extensive comparison of four signal timing programs using several real arterials and a range of traffic conditions. These programs include: TRANSYT 7F, Synchro, PASSER II, and PASSER V. PASSER V is a new program developed in this project. It contains many of the best features of the other programs. These features include: PASSER II's bandwidth maximization algorithm, a traffic model that is similar to TRANSYT 7F's link-based model, and a genetic algorithm to develop timings for minimizing delay or maximizing bandwidth. We found that Synchro produces best solutions when delay-minimization is the primary objective, and PASSER V produces the best timings when bandwidth maximization is the primary objective. In addition PASSER V's delay-minimization tool produces timings similar to Synchro.

17. Key Words Signalized Arterial, Delay Minimization, Progression, Bandwidth Maximization, Traffic Signal Coordination	18. Distribution Statement No restrictions. This document is available to the public through NTIS: National Technical Information Service 5285 Port Royal Road Springfield, Virginia 22161		
19. Security Classif.(of this report) Unclassified	20. Security Classif.(of this page) Unclassified	21. No. of Pages 28	22. Price

Form DOT F 1700.7 (8-72)   Reproduction of completed page authorized

GUIDELINES FOR SELECTING SIGNAL TIMING SOFTWARE

by

Nadeem A. Chaudhary, P.E.
Research Engineer
Texas Transportation Institute

Vijay G. Kovvali
Graduate Research Assistant
Texas Transportation Institute

and

S. M. Mahabubul Alam
Graduate Research Assistant
University of Texas at El Paso

September 2002

TEXAS TRANSPORTATION INSTITUTE
The Texas A&M University System
College Station, Texas 77843-3135

DISCLAIMER

The contents of this report reflect the views of the authors, who are solely responsible for the facts and accuracy of the data, the opinions, and the conclusions presented herein. The contents do not necessarily reflect the official view or policies of the Texas Department of Transportation (TxDOT), the Federal Highway Administration (FHWA), the Texas A&M University System, the Texas Transportation Institute (TTI), the University of Texas System, or the University of Texas at El Paso (UTEP). This report does not constitute a standard or regulation, and its contents are not intended for construction, bidding, or permit purposes. The use and names of specific products or manufacturers listed herein does not imply endorsement of those products or manufacturers. The engineer in charge of the project was Nadeem A. Chaudhary, P.E. # 66470.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF FIGURES

1. INTRODUCTION

BACKGROUND

Engineers install traffic signals to provide safe right of way to competing traffic movements. When two or more traffic signals are located in close vicinity, traffic flow on links joining the two signals becomes dependent on timings at these signals. This dependency may be strong or weak depending on a number of factors. These factors include:

- type of facility,
- distance between signals,
- link speed,
- traffic volumes, and
- traffic distribution and origin/destination patterns.

Typically, the first two of these factors remain unchanged for many years. However, the other three may change from one day to the next and several times within a day. At certain times of a day, arterial traffic flow may be balanced in both directions or be predominant in one direction. Coordinating a pair of traffic signals improves traffic flow when there is high flow dependency between pairs of signals. A signalized arterial may contain as few as two signals or as many as a dozen or more signals. Most urban arterials have predominant traffic flow in at least one flow direction during a significant part of each day. During these times, traffic flow can be significantly improved by coordinating all traffic signals on the arterial.

Coordinating two or more signals on a signalized arterial requires the determination of the following four signal-timing parameters to achieve the desired results or objectives:

- cycle length,
- green splits,
- phase sequence or order, and
- offsets.

Providing or maintaining safe flow of traffic and pedestrian traffic at each signal in the system is very important. Engineers achieve this objective by selecting phase clearance times that satisfy minimum requirements based on operational needs and driver expectancy. In addition, engineers can coordinate signals to achieve one or more of the following objectives:

- minimizing delay,
- minimizing number of stops,
- maximizing progression efficiency,
- minimizing queue size at approaches, and
- maximizing system throughput.

All of the above objectives may not apply under a given set of geometric and traffic conditions. Even if they do all apply, it may not be possible to fully achieve all objectives simultaneously.

Delay to vehicles at a link, for instance, is a function of how much time vehicles spend traveling on the link and the time they spend stopped in queue at a signalized approach. Neither of these delays can be completely eliminated. Thus, engineers desire to minimize this delay. They can minimize time spent waiting during a red phase by using a smaller cycle length, which produces less red time and shorter cycle-by-cycle queues. However, since a smaller cycle length also produces smaller green time, the number of stops to vehicles may increase. In addition to using a smaller cycle length, engineers can minimize delay by timing the lights such that the bulk of vehicles arrive during green. In minimizing delay, priority is given to the most significant traffic stream (through or cross street) flowing from an upstream signal to the downstream signal.

Maximizing progression, on the other hand, gives priority to arterial through traffic. Thus, it minimizes stops and delay to through traffic at the expense of cross-street traffic, so it may not result in the lowest possible total delay. Signal timings providing maximum through progression are easily noticed and appreciated by drivers, especially in Texas. The reason is that these drivers generally do not mind extra delay at minor approaches, but they do not like a situation where they have to stop many times while traveling through on the arterial. On the other hand, drivers cannot easily notice differences in delay.

Minimizing the size of a queue of vehicles becomes a necessity under many geometric and traffic scenarios, listed below:

1. a queue of vehicles in a turn bay starts to interfere with, or spills back into, the adjacent lane;
2. a queue in a through lane extends to, or beyond, the entrance to an adjacent turn bay; or
3. a link queue reaches near, or spills back into, the upstream signal.

In the less severe of these cases, the saturation flow reduces. In the worst scenarios, traffic flow is forced to halt. In such an event, the available capacity of a signal phase and, thus, the signal capacity go to waste. The situation when a signal is green but vehicles cannot use this opportunity is referred to as starvation. In coordinated signal systems, only a few signals may be critical because they are operating at or near capacity. The other signals usually have slack capacity, and starvation at these signals happens by design. At critical signals, however, engineers must make all possible attempts to eliminate starvation. The reason is that starvation causes loss in much-needed capacity and, as a result, reduces system throughput or productivity. Throughput (in vehicles per hour) of a signal system is its ability to move vehicles through the system. Thus, higher throughput is better. Reduction in throughput occurs not only in congested systems but in pseudo-congested systems as well. A signal system is pseudo-congested when it has the ability to provide capacity to satisfy traffic demand, but it does not provide sufficient capacity because of non-optimal signal timings (green splits and offsets). Shorts links (less than 500 ft) are often sources of pseudo-congestion in a signal system.

For an arterial experiencing certain traffic conditions, different sets of values can be selected for various signal-timing parameters. There could be millions of possible combinations of these parameters; however most of these possibilities will not achieve any of the above objectives in a satisfactory manner. A small number of these possible sets will result in the achievement of a subset of the objectives described above. Fortunately, a number of computer programs are

available for use by engineers to time signals on arterials. All of these programs are based on simplified models of reality, and their effective use requires some knowledge of how they work. In addition, users of these programs must be familiar with the principles of traffic engineering, which will allow better use of these computerized tools. In the following subsections, we provide this basic information.

CONSIDERATIONS IN SELECTING SIGNAL TIMING PARAMETERS

Selection of good cycle length cannot be overemphasized. Maintaining stable flow of traffic from one signal through an adjacent signal on an arterial implicitly requires that all signals in a coordinated system operate under a common cycle length. As a result, some restrictions have to be placed on cycle lengths of individual signals. These restrictions can be established using Webster's theory for isolated traffic signals facing undersaturated traffic conditions (*1*). According to this theory, there exists an absolute minimum (critical) cycle length for an undersaturated traffic signal. It is:

$$C_c = \frac{L}{1-Y} \qquad\qquad \text{Equation 1}$$

Where:
$\quad C_c\quad$ = Critical or minimum cycle length,
$\quad L\quad$ = Total lost time for all critical phases, and
$\quad Y\quad$ = Sum of flow (volume/saturation flow) ratios for all critical phases.

The flow ratio for a movement is equal to volume divided by saturation flow rate. Saturation flow rate is a measure of how much traffic will move if the signal remains green continuously for one hour, and it depends on lane assignment and a number of other factors. The critical cycle provides saturated green times for all critical movements. However, when traffic arrivals are random and vary from cycle to cycle, critical cycle will not be able to clear a queue if one develops due to higher demand during a signal cycle. Thus, it is desirable to use a slightly larger cycle length. According to Webster's theory, the minimum delay cycle length for a pretimed signal is:

$$C_m = \frac{1.5L - 5}{1-Y} \qquad\qquad \text{Equation 2}$$

The reader should note that C_m is approximately 50 percent higher than C_c. Note that these equations will result in unusually large cycle lengths when Y approaches one. They do work well for undersaturated conditions. In practice, one needs to set a practical upper limit on cycle length. A variation of Equation 2 for application to actuated signals (*2*) is as follows:

$$C_m = \frac{1.3L - 5}{1-Y} \qquad\qquad \text{Equation 3}$$

This equation accounts for shortening of cycle length in the field due to phase gap-out and phase skipping. In actuated-coordinated signals, the cycle length remains unchanged because any slack

time gets allocated to the coordinated phases. Thus, it is appropriate to use the original formula (Equation 2) for pretimed signals.

Figure 1 illustrates a hypothetical delay-versus-cycle length curve for an isolated signal. This figure also identifies critical and minimum-delay cycle lengths. The location of this curve depends on traffic volumes and saturation flow rates. For fixed saturation flow rates, increase in traffic volumes will shift this curve to the right and up. However, the following general characteristics apply to all situations:

- Delay increases sharply for decreases in cycle lengths below C_c.
- The curve is flat in the vicinity of C_m, and deviation from this cycle length will not adversely affect delay as long as this deviation is within 85 and 120 percent.
- Delay curve on the right of C_m has a lower absolute slope than that on the left of C_m.
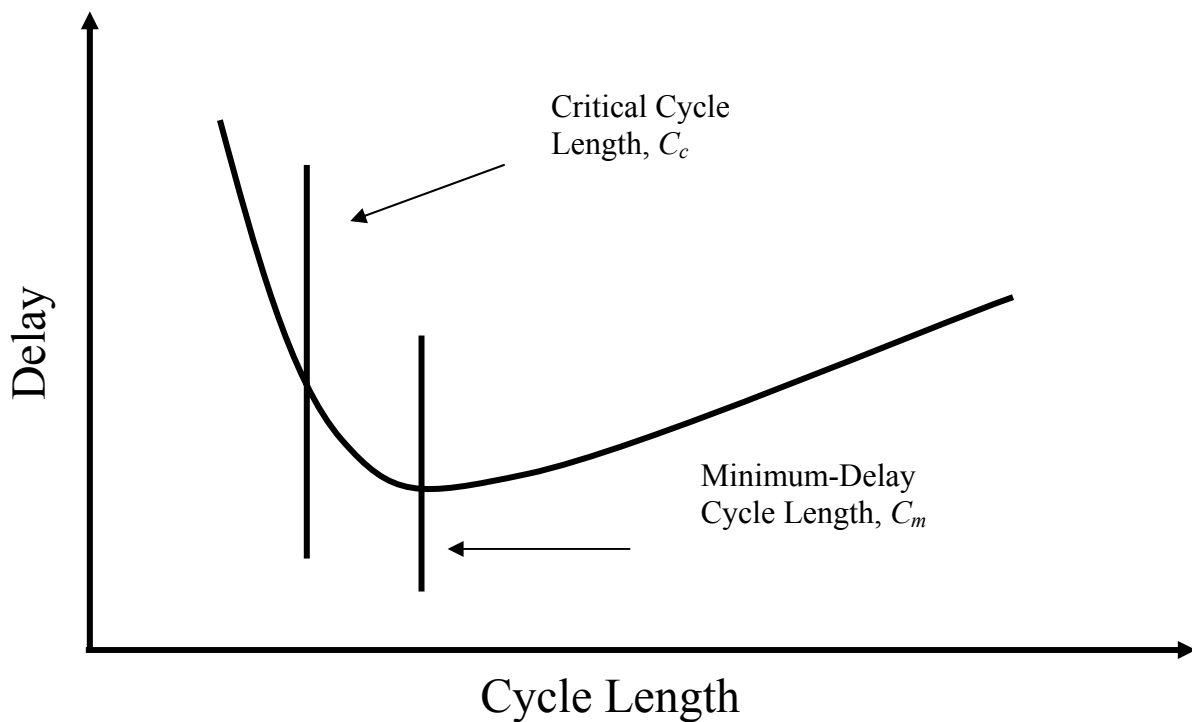
Figure 1. Shape of a Typical Delay-versus-Cycle Length Curve for an Isolated Signal.

In a system of signals under consideration for coordination, minimum delay cycle length for each signal will be different from other signals. This difference will require selection of a cycle length that may compromise the operation of some signals but is necessary to provide good coordination that will benefit all signals. The selection of a common cycle length should ensure that no signal in the system causes excessive delays to motorists or causes queue spillbacks. We illustrate the initial process of cycle length selection for one possible scenario using Figure 2. This figure shows hypothetical cycle length versus delay curves for three signals with different minimum delay cycle lengths. In this situation, the signal with the maximum minimum-delay

(*Maximin*) cycle length could be used to establish the system cycle length. The reason is that this selection does not result in an unusually large cycle length with respect to the other two signals.
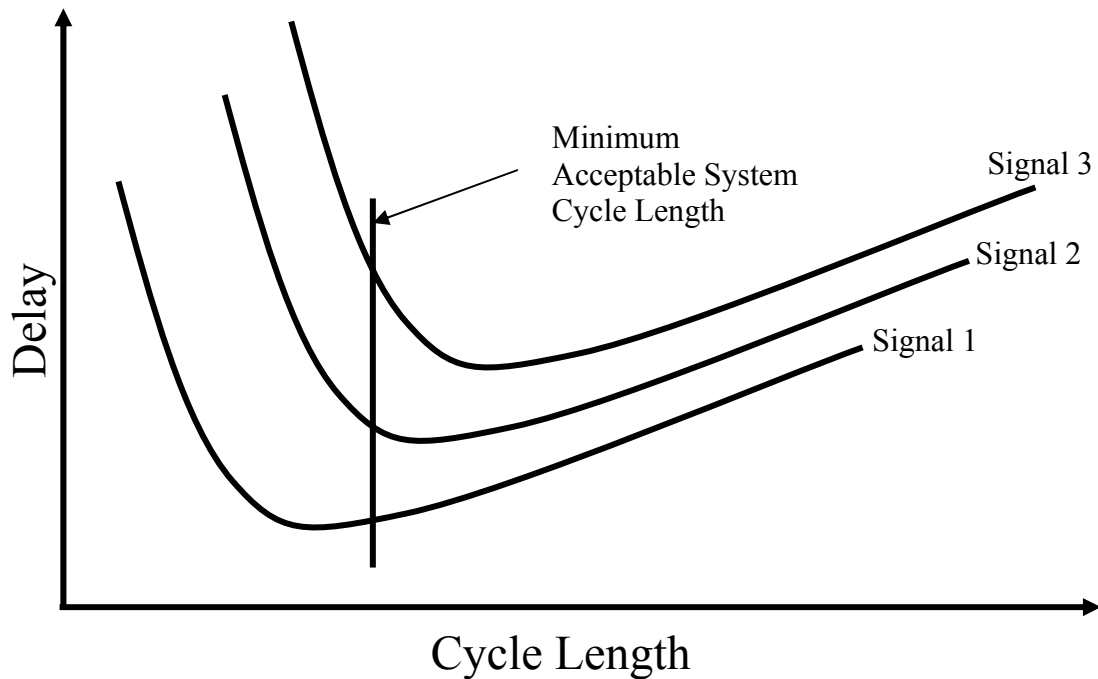

Figure 2. Constraints on System Cycle Length.

Figure 3 illustrates another common scenario where the minimum delay cycle length for at least one signal (Signal 1) in the system is significantly different from those for other signals. In such cases, the analyst can divide the system or coordinate using half/double cycling. One can implement half/double cycling by operating a group of signals having similar characteristics using a common cycle length, and each of the remaining signals using a cycle length that is a half or double of the common cycle length. Link length and traffic volumes play a major role in deciding whether to coordinate such a signal with other signals. For the case illustrated in Figure 3, assume that the *Maximin* cycle (corresponding to Signal 3) is 100 seconds, the minimum-delay cycle length for Signal 1 is 55 seconds, and that Signals 1, 2, and 3 are located on an arterial in that order. In this case, one has the following distinct options:

1. If traffic flow between Signal 1 and Signal 2 is light or the distance between signals is large, operate Signal 1 as isolated.
2. Double-cycle Signal 1 and coordinate with the other signals.
3. Operate all signals using a common cycle length constrained by the *Maximin* cycle length.

The most common method of calculating green splits is to allocate the selected signal cycle to signal phases in proportion to volume-to-saturation flow ratios for the critical movements served by these phases (*1*). Most programs for timing traffic signals use this method. However, the user must enter data to ensure that calculated splits meet some minimum constraints based on pedestrian requirements and motorists' expectations. Furthermore, safety concerns for motorists

may require precluding some phase patterns. An example of a safety concern is the use of split phasing at intersection approaches where possibility of collision exists for opposing left-turn traffic.
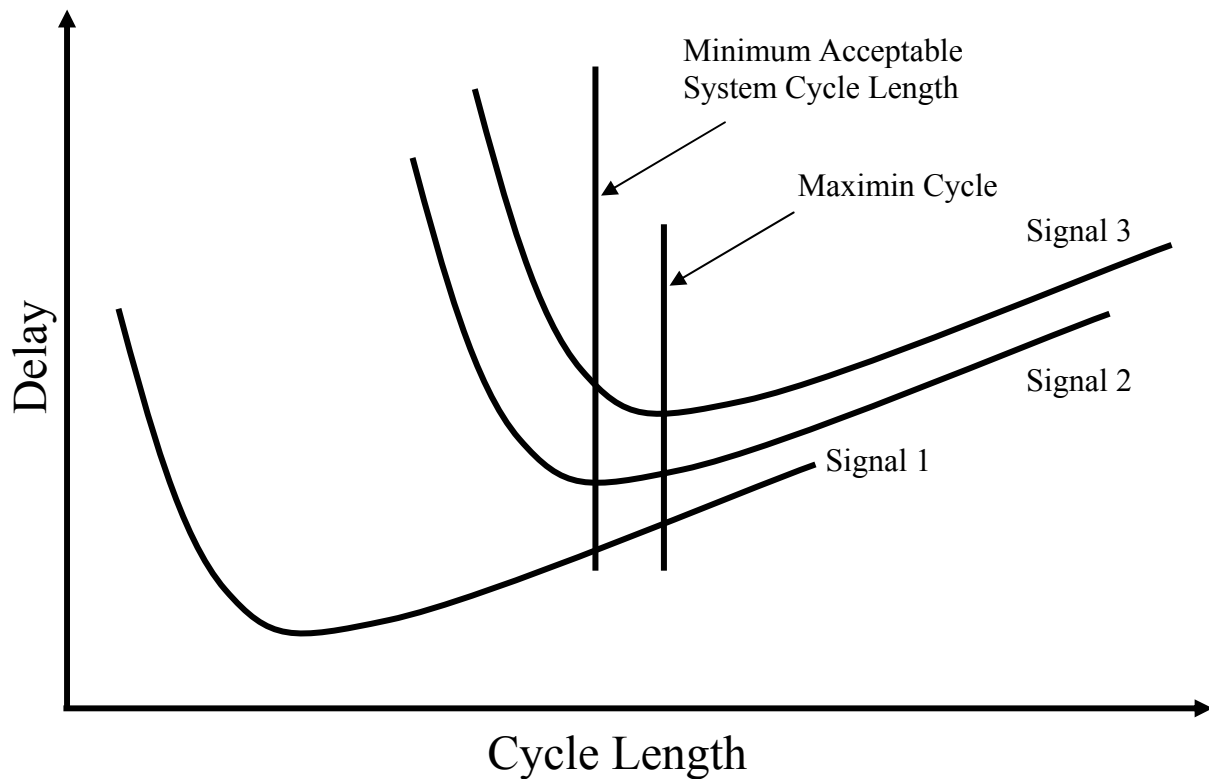

Figure 3. Possible Warrant for Double Cycling a Traffic Signal.

Earlier, we listed several objectives of signal coordination and mentioned that all listed objectives may not apply to a given situation. Even when multiple objectives are applicable, selecting a timing plan to achieve one objective may not guarantee the achievement of other objectives. For instance, minimizing systemwide delay generally results in smaller queues but not necessarily maximum progression bandwidths and vice versa. Maximizing progression bandwidth, on the other hand, will generally result in the least number of stops as well. However, it is possible to optimize signal timings to achieve one primary objective and then fine-tune some signal-timing parameters to improve a secondary objective without sacrificing the primary objective. Thus, the user should select one primary objective and several secondary objectives of coordination. In any situation, cycle length selection is extremely important.

MODELS USED BY TRAFFIC SOFTWARE

A number of computer programs are available to assist in the analysis and coordination of traffic signals on an arterial. All of these tools are based on the abstraction of reality and have their inherent weaknesses and strengths. In this section, we provide a brief description of commonly used models in popular signal timing and analysis software. These models could belong to the following two classes:

1. a traffic simulation model, also called a traffic model; and
2. an optimization model.

The following subsections provide a brief discussion of key concepts related to the above two types of models. We restrict this discussion to models for signalized arterials.

Traffic Signal Analysis Models

A traffic model takes traffic volumes, geometric information for the facility, and the complete description of a traffic control plan as input. Then, it simulates the described scenario and outputs measures-of-effectiveness (MOEs). Typical MOEs include: average or total delay, stops, fuel consumption, total bandwidth, bandwidth efficiency, average or maximum queues, throughput, and starvation. Most traffic models provide an estimate of several, if not all, MOEs. One method of classification of a program is the primary MOE used by it for selecting the best signal timings. The two main types of models are: delay-based and bandwidth-based. The level of detail or abstraction used by a model is another classification method.   There are three common types of traffic models based on the latter classification: microscopic, mesoscopic, and macroscopic.

Microscopic Traffic Models

Microscopic traffic models provide the most detailed analysis by simulating the behavior (acceleration, deceleration, car-following, etc.) of individual vehicles in the traffic stream. In general, these models are also stochastic in nature and rely on a random number generator that uses a seed value to generate values of various parameters during simulation. To obtain another sample, the user must change the seed value and re-run the simulation. Running the simulation with different random number seeds is equivalent to collecting random samples of data, similar to collecting data for a peak period during many consecutive days. Due to the level of detail simulated, these models require the maximum amount of data and are the most computationally intense. CORSIM (*3*) uses a macroscopic traffic model.

Mesoscopic Traffic Models

These models simulate traffic flow in specified time steps, and they are deterministic in nature. The time step can be 1 second, 2 seconds, or larger. For each time step, these models estimate the flow of traffic entering a link, traveling downstream, stopping due to a red light, and moving again when the light turns green. Some of these models also account for platoon dispersion as vehicles travel from one point to a downstream point in space. Mesoscopic models can be further classified as link-based or time-based. Link-based models simulate traffic flow one link at a time for all time steps in a signal cycle. These models treat a queue of vehicles at the signal approach as an upward stack. As a result, all vehicles arriving during red travel to the stop-bar and join a vertical (upward stack) queue. Link-based models cannot account for queue spillback because they do not keep track of the back of the queue. In addition, they may allow more vehicles to stack in a queue than a link's storage capacity. Thus, these models are not suitable for congested conditions or for short links where sub-optimal timing may cause queues to block flow from the

upstream signal. Step-based models, on the other hand, simulate traffic flow on all links at each time step. These models can accurately account for the behavior of queued traffic and traffic flow interactions between adjacent links, and they are better suited for all types of traffic conditions in signal systems. The down side is that they require more computer time than the link-based models. Also, the accuracy of these models may depend on the number of cycles simulated, which in turn depends on the number of cycles needed to achieve stable flow.

Macroscopic Traffic Models

Models in this category simulate the cycle-by-cycle behavior of platoons of traffic at each link in the system. They are deterministic in nature. These models may or may not account for platoon dispersion. Macroscopic models treat a queue of vehicles at an approach as an upward stack. Thus, they are accurate only for undersaturated flow conditions. Because of their simplistic nature, macroscopic models are the most efficient from a computational point of view.

Optimization Models and Search Algorithms

As mentioned earlier, traffic models simulate a given set of traffic and control conditions. In other words, they are able to only tell how good or bad a given scenario is. Optimization and search algorithms are techniques that systematically generate multiple scenarios, compare their fitness or objective function value (i.e., delay, bandwidth efficiency, throughput, etc.) obtained from a simulation model, and select the best scenario based on a predetermined criteria. For instance, if delay minimization is the desired objective, the primary fitness value will be the delay to motorists resulting from a specific scenario. Such an optimization model will evaluate the delay value for each alternate timing plan and select the timing plan that results in the least amount of delay. In other words, search algorithms are wrappers around traffic simulation models to provide the optimization function. Search algorithms can be simple or extremely sophisticated. The following subsections describe some of the common search algorithms.

Exhaustive Search Algorithms

As the name implies, these algorithms calculate and compare the fitness values for all possible signal-timing scenarios. As mentioned previously, there can be millions of such combinations of signal-timing parameters, depending on the size of the facility and how many variables are to be optimized simultaneously. Thus, exhaustive search may require hours of computer time. Unless a model is designed for small facilities, the shear number of possible scenarios usually requires the use of a divide-and-conquer strategy. For instance, computational time can be drastically reduced by stage-wise optimization of each variable instead of all variables simultaneously. Such a strategy increases computational efficiency by compromising accuracy. The positive feature of exhaustive search algorithms is that full information is available for each scenario. Most signal-timing programs use some level of exhaustive search combined with other search algorithms.

Hill-Climbing Algorithm

A hill-climbing (or valley descent) algorithm starts with one (base) scenario, either specified by the user, selected by the program using a fixed criterion, or selected randomly. Then, it selects

one or more variables (i.e., offset, cycle length, etc.). Then, it creates two additional scenarios for each variable, one by increasing the values of that variable and the other by decreasing the value. The values of the selected variables are increased or decreased by specified amounts called step sizes. Following this, the algorithm uses a traffic simulator to calculate the fitness values for each of the two new scenarios and compares them with the base scenario. This evaluation identifies the two best scenarios and, consequently, a direction of further search. For instance, if increasing the value of the selected variable resulted in a better fitness value, the search algorithm will mark this new scenario as the current best and continue in the direction of increasing values of the variable. In the next iteration, the search algorithm generates a new scenario by increasing or decreasing the values of the selected variable in the selected search direction, calculating the new fitness value, and comparing it with the two current best values. The algorithm continues in this manner until the fitness value for the new scenario ceases to be better than the current best. Hill-climbing methods guarantee optimal solution only when the function to be optimized is unimodal (that is, it has one peak or valley). For multi-modal functions, the hill-climbing method may terminate with a sub-optimal solution depending on how good the base scenario is. Most implementations of hill-climbing algorithms use sophisticated techniques, such as a variable step size, to speed up the search process.

Mathematical Programming Techniques

Mathematical programming techniques, such as linear- and integer-programming, require a compete specification of the objective (fitness) function, along with all the applicable constraints of the traffic model in mathematical form (equations and/or or inequalities). These techniques are based on systematic procedures (called programs) that are designed to search a small subset of all possible scenarios in an intelligent manner. Mathematical programming techniques are applicable only when a closed-form mathematical model exists. When applicable, these techniques also guarantee the best solution. Further discussion of these techniques is beyond the scope of this report.

Genetic Algorithms

A genetic algorithm (GA) starts with a subset of scenarios (some members of a population) and applies principles of natural selection (mating, gene mutation, etc.) to generate a new or revised set of scenarios (called the next generation). A GA-based optimization model uses a specified traffic simulation model to evaluate the fitness of each member (i.e., a signal-timing scenario) in the current population. Then, it generates a new population by combining the characteristics of (that is, by mating) selected pairs of scenarios (members) with the best fitness values. The principles of natural selection ensure that the characteristics of the fittest members (i.e., those with higher bandwidths or lowest delays, depending on the objective of optimization) have a high probability of transmission to the next generation. A GA terminates when either no more improvement occurs or a certain number of user-specified generations are complete, whichever occurs first. GAs are different from all previously described search algorithms in that they utilize codings of variables rather than the values of variables. Given a large enough population and sufficient number of generations, a GA can provide the global optimum. GAs have this ability because they perform simultaneous optimization of all variables. GAs can be applied to all types of optimization problems, even those that cannot be described in closed forms. Their

effectiveness depends on the scheme used for coding the variables and the details of the natural selection process used. Conceptually, an exhaustive optimization algorithm is a GA that uses all members of a population and applies the initial generation of the optimization algorithm only.

2. SIGNAL TIMING OPTIMIZATION PROGRAMS FOR ARTERIALS

TRANSYT 7F

TRANSYT 7F (*4*) uses a mesoscopic-deterministic model for analyzing and optimizing signal timings on arterials and networks. TRANSYT 7F uses a combination of exhaustive, hill-climbing, and GA-based optimization methods. Furthermore, it uses a delay-based traffic model. In other words, it is primarily designed to select signal timings that produce minimum system delay and stops. In addition, it provides a capability to select several secondary objectives, including minimization of stops and maximization of progression opportunities. During its optimization process, TRANSYT 7F generates second-by-second flow profiles of vehicles on all links in the network. Then, it analyzes these profiles to determine MOEs. TRANSYT 7F has two delay-based traffic models: link-based and step-based. The original link-based model works well for undersaturated traffic conditions. It has been extensively validated by users all over the world and accepted as a sound model. Recently, the program developers added the step-based model to remove the limitations of the first model. This model takes into consideration the formation and dissipation of queues in space. In addition, it accounts for flow interactions on adjacent links through a step-by-step analysis of all links in the system. Conceptually, this model is better suited for the analysis and optimization of congested (oversaturated) facilities. Consequently, it also requires more computation time. TRANSYT 7F treats actuated signals as equivalent pretimed signals. It also has the ability to half/double cycle traffic signals. Until recently, the main deficiency of TRANSYT 7F has been its inability to optimize signal phase sequences. In the latest version of TRANSYT 7F (Version 9) released earlier this year, developers removed this deficiency through the addition of a GA-based optimization algorithm. Unfortunately, Version 9 was not available in time for use in this project.

TRANSYT 7F performs exhaustive search for cycle length. For each cycle, it starts by calculating equal saturation splits and applying a hill-climbing method to optimize signal offsets and splits. For this reason, its final results depend on the base timing plan supplied by the user. Although it contains a good delay-based traffic model, TRANSYT 7F's bandwidth analysis model is not very good. Finally, learning to use TRANSYT 7F requires considerable effort. For these reasons, many practitioners dealing with signalized arterials, especially in Texas, prefer not to use it.

SYNCHRO

Synchro (*5*) is a delay-based program for analyzing and optimizing timing plans for arterials and networks. Its objective function also minimizes stops and queues by applying penalties for these MOEs. Synchro's traffic model is similar to the link-based model in TRANSYT 7F. Synchro uses an exhaustive search technique to optimize signal timings. To reduce the number of scenarios analyzed for a coordinated system, it relies on the divide-and-conquer principle. To optimize timings for an arterial, the program requires the user to apply several manual steps (cycle length optimization followed by offset and phase sequence optimization) in a specific order. It optimizes cycle length by analyzing all cycles in the defined range. Optimization of offsets is divided in multiple stages, during which the step sizes depend on the optimization level selected by the user. For instance, if the user requests extensive offset optimization, Synchro

simulates all offsets in 4-second increments, followed by a search using 2-second increments. Finally, it performs another search using 1-second increments in the vicinity of the best offset from the second stage. Unlike TRANSYT 7F, Synchro's traffic model does not consider platoon dispersion. However, it recommends when to coordinate two adjacent signals by calculating a *coordinatability* factor using link distance, travel time, and traffic volumes as input.

Unlike other programs, Synchro generates optimal signal timings for each signal by averaging the analysis results of five volume scenarios for that signal. For this purpose, it assumes that a volume entered by the user is the mean and variance of the real traffic volume having a Poisson distribution. Then, it applies *Z* factors from a normal distribution to generate four additional volume scenario representing −2 (10th percentile), −1 (30th percentile), 1 (70th percentile), and 2 (90th percentile) standard deviations from the mean. In this scheme, user-supplied volumes are treated as 50th percentile volumes. Once these five percentile volume scenarios have been determined, the program calculates cycle length and green splits (using Webster's method) and delay (using HCM method) for each of the five volume scenarios for a signal. The program averages these values for all five scenarios. In Synchro terminology, timing and delay calculation based on this averaging method is the *percentile delay method*. Using this method, Synchro incorporates a method to model phase gapping and skipping behavior for actuated and actuated-coordinated signals. This method skips or gaps a phase if the number of vehicles estimated to arrive per cycle is less than 0.69. Lastly, Synchro accounts for upstream metering in its delay calculations. Synchro optimizes all signal-timing parameters for pretimed and actuated signals, and it applies internally calculated progression adjustment factors for progressed movements. It can also handle double and half cycling of signals. For each cycle length, the program summary report includes numerous MOEs. This list, however, does not include measures showing the quality of through progression. Thus, a user must manually inspect the time-space diagram to determine if a timing plan provides acceptable progression.

Synchro has an excellent user interface that provides features to easily fine-tune a timing plan. Furthermore, it provides for data conversion to other popular software. Due to this, Synchro popularity has grown at a phenomenal rate since its initial availability during the mid 1990s. Because of its ease of use, many engineers use it as an input processor for TRANSYT and CORSIM.

PASSER II

PASSER II (*6*) is a bandwidth-based program for optimizing signal timings for signalized arterials. Originally developed for TxDOT about 30 years ago, it is one of the most popular programs in its class. The heuristic signal-timing optimization model of PASSER II is based on a graphical technique, which is simple, efficient, and powerful. PASSER II has passed the test of time and is known to produce good signal-timing plans, even when some level of congestion exists on an arterial. PASSER II can determine all four signal-timing variables described earlier. It selects the plan that maximizes arterial progression. Because of its simplicity, it is also the most computationally efficient program in its class.

PASSER II performs exhaustive search over the range of cycle lengths provided by the user. It starts by calculating equal saturation splits using Webster's method. Then, it applies a hill-

climbing approach to adjust splits to minimize delay. Finally, it applies a bandwidth optimization algorithm using the pre-calculated splits for a specific cycle length as input to that model. At the optimization stage, it finds offsets and phase sequences that produce maximum two-way progression. At this stage, PASSER starts by calculating offsets for providing perfect one-way progression in the A (arbitrarily selected) direction. Then, it minimizes band interference in the B (opposite) direction by adjusting phasing sequences and offsets. The maximum total band calculated by the program is as follows:

$$Total\ Band = G_A + G_B - I \qquad\qquad \text{Equation 4}$$

where:

G_A = Least green in A-direction;
G_B = Least green in B-direction; and
I = Minimum possible band interference.

After achieving the best band (minimum interference) in the B direction, the program adjusts the two bands according to user-desired options for directional priority. The reader should note here that the interference minimization algorithm intelligently searches a very small subset of all possible combinations of signal timings. Finally, the program calculates delays, bandwidth efficiency, and attainability.

The program performs delay calculations for internal through movements using a macroscopic traffic model that explicitly considers platoon dispersion. Delay calculations for all other movements use the Highway Capacity Manual (HCM) method. Efficiency and attainability measure the quality of through progression provided by a timing plan. Efficiency for a direction is the percent of cycle used for progression. Attainability is the percent of bandwidth in a direction in relation to the minimum green split in the same direction. Theoretically, the maximum bandwidth in a direction can be no more than the smallest through green split in that direction. PASSER II uses the following formulas to calculate combined efficiency and attainability for the two arterial directions:

$$Progression\ Efficiency(\%) = \frac{(Arterial\ Band_A + Arterial\ Band_B)}{2*Cycle\ Length}*100 \qquad \text{Equation 5}$$

$$Progression\ Attainability(\%) = \frac{(Arterial\ Band_A + Arterial\ Band_B)}{(Min.Green_A + Min.Green_B)}*100 \qquad \text{Equation 6}$$

The reader should note that while bandwidth generally increases with an increase in cycle length, efficiency may increase, decrease, or remain constant. Figures 4 and 5 illustrate this concept using a real arterial.
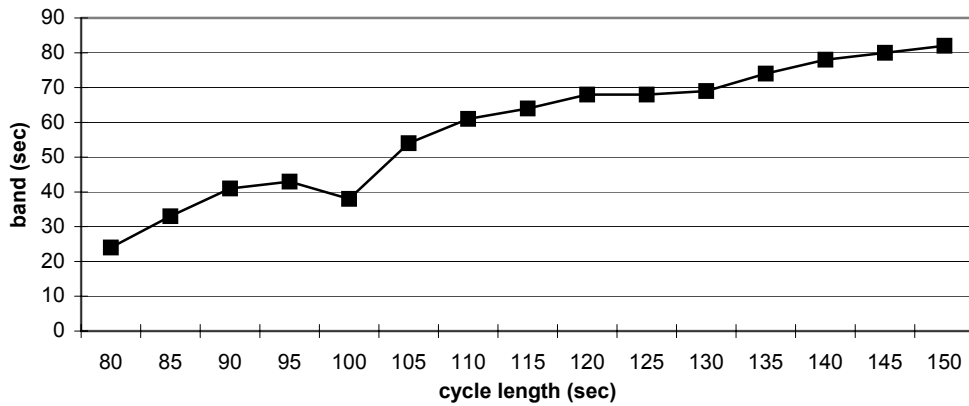
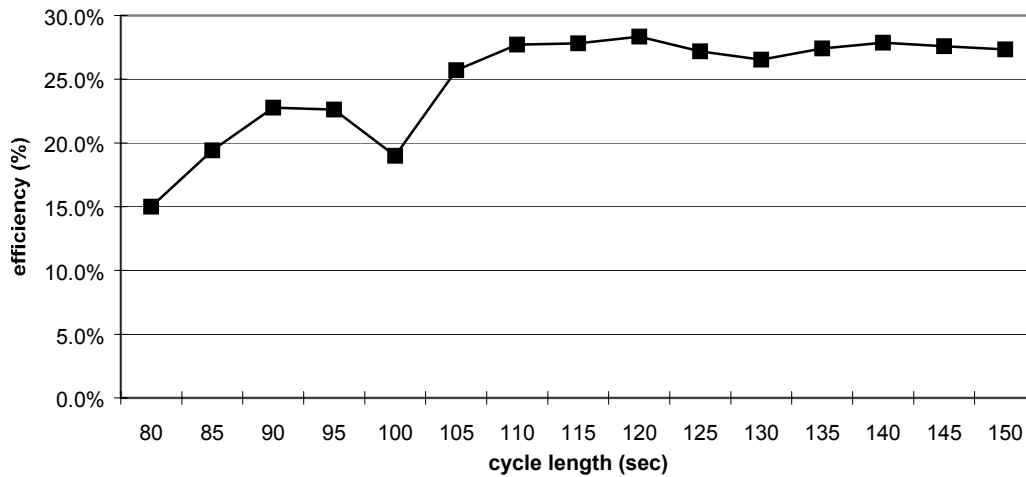**Figure 4. Cycle Length versus Progression Band.**

Figure 5. Cycle Length versus Band Efficiency for the Case in Figure 4.

In the example problem illustrated in Figures 4 and 5, 120-second cycle length provides the best efficiency. Cycle lengths above 120 seconds provide larger bands, but their efficiencies are smaller. The user should select a timing plan that provides high, if not the largest, efficiency and lower delay. An attainability value of 100 percent identifies the best possible timing plan for a given cycle-split combination. Thus, the user-selected timing plan also should have high attainability. The user could also select other criteria from a policy point of view. One of these timing criteria would be to set a maximum upper limit on the cycle length. Another might be to select a timing plan that provides at least a specified minimum bandwidth. The latter criterion is important because a small band (for instance, less than 15 seconds) may not produce efficient progression from a practical point of view, even if the corresponding timing plan provides maximum efficiency.

PASSER V

PASSER V (*7*) is a new program the researchers developed in this research project. We designed it to provide analysis and optimization of arterial signal timings. It contains many of the best features of the programs described above. PASSER V has a graphic user interface that provides access to several tools for analyzing and optimizing arterial signal timings. These tools are described below.

The *PASSER II* tool uses a re-engineered version of the interference minimization algorithm used by previous versions of PASSER II programs. This version of algorithm contains the following changes:

1. It computes splits for each cycle using Webster's method and utilizes these splits without modification, and
2. It applies the interference algorithm for both directions and chooses the best timing plan.

The above modifications resulted in two improvements. First, the tendency of earlier PASSER II programs to select large cycle lengths has been eliminated. Second, it finds better progression bands than PASSER II software for the same set of splits. In addition, we have added features to automatically fine-tune offsets to minimize delay. For fine-tuning offsets, it uses a new traffic model that also applies to congested conditions. Similar to previous PASSER II programs, this tool uses an exhaustive search method for determining the best cycle length. This tool is for arterials that do not contain Texas diamond interchanges using TTI 4-Phase or 3-Phase strategies.

The *PASSER III* tool provides a capability similar to the PASSER III program (*8*) for timing diamond interchanges. It uses an exhaustive search method like PASSER III.

The *GA Optimizer* uses a genetic algorithm, and it is capable of simultaneously selecting all signal-timing parameters. It can be used to optimize progression bandwidth or delay. It uses Webster's method for calculating green splits. For bandwidth optimization, it uses a new bandwidth analysis routine. For delay optimization, it uses a new traffic model that is mesoscopic in nature. This model is similar to the step-based model in TRANSYT 7F. Because of its ability to model queues in space, this delay model applies to all types of traffic conditions. The GA-based optimizer can also time arterials that contain Texas diamond interchanges.

The *Volume Analysis* tool provides a simple method for analyzing the throughput capacity of facilities. This model identifies critical approaches and provides accurate results when queue spillback or blocking does not occur at interior approaches of the coded network. Therefore, its results are accurate for most diamond interchanges using TTI 4-Phase strategy. In addition, this tool can be used for analyzing the spacing-versus-capacity issue for such interchanges. The *Volume Analysis* tool uses an exhaustive search method.

The *Delay/Cycle Analysis* tool uses the new mesoscopic traffic model in PASSER V to analyze delay for a specified range of cycle length. It provides the user options to keep the existing offsets unchanged or to change them in proportion to changes in cycle length relative to the

current system cycle. It also provides a graph comparing its delay to delay calculated using HCM procedures assuming random arrivals at all signals. This tool is especially useful for analyzing the effects on delay of small changes in system cycle length.

The *TS-Diagram* tool generates a graphical time-space diagram that allows the user to manually fine-tune offsets.

Thus, PASSER V provides many of the best optimization and analysis features of other programs under a common user interface. These features are above and beyond those provided by previous versions of PASSER II and PASSER III. In addition, PASSER V's traffic model is similar to that in TRANSYT 7F, and it can accurately model effects of congestion.

3. SOFTWARE SELECTION GUIDELINES

SOFTWARE COMPARISON RESULTS

In this project, we compared the performance of several popular programs for optimizing the timings at signalized arterials. In the following subsections, we provide a summary of findings about these programs.

Synchro 4 and 5

Synchro, a delay-based program, has one of the best user interfaces. It minimizes a function of delay, stops, and queues. We found the following about this program:

- Its cycle length selection algorithm is very good.
- Even though its delay model does not consider effects of queue spillback and blocking, in most cases Synchro found the best timing plan (lowest delay, stops and queues) for real arterials. These arterials had a mix of congested and uncongested links.
- For small arterials, Synchro also produces good progression bands, but this ability severely degrades for larger arterials. The reason is that the program does not explicitly attempt to find progression bands and, thus, it may select a timing plan that does not provide any through bands. Also, the signal timing summary produced by the program does not contain any information about progression. Thus, the user must inspect the time-space diagram for a plan to determine if it provides good progression.

TRANSYT 7F

TRANSYT 7F is a delay-based program. Its default objective is to minimize a function of delay and stops. In this research, we tested Version 8.2 of the program. This version has a very good traffic model. We found that this model accurately simulates all congested and uncongested traffic conditions. However, TRANSYT 7F's performance for optimizing signal timings was the worst of all programs studied. We found that its cycle length selection algorithm is not very good. One reason may be that this version cannot optimize signal phase sequences, and our use of leading left-turn phases for all scenarios may be a cause of less-than-optimal results. In addition, the program is difficult to use as compared to other programs. In Release 9 of this software, many of the above deficiencies have been removed, but this version was not available in time for use in the project.

PASSER II-02

PASSER II is designed to produce a timing plan with the best progression bands. In general, it provides larger progression bands, higher throughput, and fewer stops than the above programs. However, its solutions may also produce larger delay than these programs. Additional studies showed that PASSER II can produce solutions with lower delay than Synchro if the user specifies a good cycle length range based on the *Maximin* criterion. In our studies, it outperformed Synchro 60 percent of time when we used a range based on this criterion. PASSER II's traffic model, however, applies only to undersaturated conditions.

PASSER V-03

PASSER V is a new program developed in this research. It provides many of the best features of all the above programs under one easy-to-use graphic user interface similar to that of Synchro. Comparison of different tools in PASSER V with other programs resulted in the following findings:

- Its bandwidth optimization algorithm produces the largest progression bands without selecting the largest cycle length.
- Its delay optimization algorithm produces timing plans comparable to those produced by Synchro.
- It produces better signal timings for congested interchanges than PASSER III.

Summary

In summary, Synchro, PASSER II, and PASSER V produced good timing plans depending on the objective of coordination. In the next section, we present guidelines for selecting signal timing software.

GUIDELINES FOR SOFTWARE SELECTION

The guidelines presented here are for selecting software for timing signalized arterials. The choice of software depends on the objective of coordination. These are addressed in the following subsections.

Maximum Arterial Progression

When providing maximum arterial progression bandwidth is the primary objective, we recommend the use of PASSER V. PASSER V provides two tools (*PASSER II* and *GA-based*) to assist engineers in achieving this objective. Bandwidth optimization also results in maximum throughput and minimum stops. In order to ensure that the program also produces a timing plan with lower delay, we recommend using the *Delay/Cycle Analysis* tool to select an appropriate cycle length range for input to the *PASSER II* tool. Also, the user can request MOEs (band, efficiency, delay, queue spill back, and starvation) for all cycles analyzed by the *PASSER II* tool and use these MOEs to select a timing plan. When using this tool, the user can also request the program to fine-tune bandwidth-based offsets to minimize delay.

Minimum System Delay

When a user wishes to implement timings that produce minimum delay, the user can select Synchro. It is the best delay-based program for undersaturated conditions. However, it should be used with caution because its traffic model does not account for queue spillback, especially for short links. When using Synchro, the user should force coordination of all signals on the arterial and select extensive offset optimization. As an alternate, the user can select PASSER V and use the GA-based delay optimizer. This tool selects timing plans with results very similar to Synchro.

REFERENCES

1. Webster, F.V. and B.M. Cobbe. Traffic Signals, *Road Research*, Technical Paper No. 56, Her Majesty's Stationery Office, London, England, 1966.
2. *Traffic Control Systems Handbook*, Federal Highway Administration, Washington, D.C., February 1996.
3. *CORSIM, Version 1.03 User's Manual*. Kaman Sciences Corporation, Colorado Springs, Colorado 80933-7463, April 1997.
4. *TRANSYT 7F User's Guide*. Methodology for Optimizing Signal Timing: MOST Volume 4. University of Florida, Gainesville, Florida, March 1998.
5. Husch, D., and J. Albeck. *Synchro 4.0 User Guide*. Albany, California, 1999.
6. Chang, E.C., and C.J. Messer. *Arterial Signal Timing Optimization Using PASSER II-90, Program User's Manual*, Report 467-2F, Texas Transportation Institute, College Station, Texas, June 1991.
7. Chaudhary et al. *Software for Timing Signalized Arterials*. Research Report 4020-1, Texas Transportation Institute, College Station, Texas, September 2002.
8. Venglar, S., P. Koonce, and T. Urbanik II. *PASSER III-98 Application and User's Guide*, Texas Transportation Institute, Texas A&M University System, College Station, Texas, 1998.