

Developing optimal timing plans for cyclic traffic along arterials using pre-timed controllers

N. T. Ratrout

Department of Civil Engineering, King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia

Abstract

Many countries are still using pre-timed signal controllers. Although this type of controller is technically inferior compared to actuated and computerized controllers, they provide an acceptable level of performance, especially for cyclic traffic. They are also cheaper and reliable since they do not depend on traffic sensors. The efficient use of pre-timed controllers depends primarily on selecting the optimum number of signal timing plans within a day and the starting and ending time of each plan which is commonly known as Time-of-Day (TOD) breakpoints. An optimal timing plan is then developed for each time period using optimization software such as Synchro. When the traffic pattern is cyclic and repeats itself from one day to another, these plans can be used for several days. This paper suggests a procedural framework for selecting the optimum number of timing plans and TOD breakpoints through the use of K-means statistical clustering and traffic simulations. This study also intends to reduce “outliers” of clusters. Outlier means that the statistical clustering technique proposes a short cluster interval between two relatively long intervals. K-means clustering technique based approach to find out optimal TOD breakpoints usually suffers from “outliers” of clusters. In this proposed procedure, Z-score of approach volume, in lieu of raw approach volume, is used simultaneously with time variable to solve the outlier problem. The recommended procedure was tested on a major arterial in Al-Khobar city, Saudi Arabia. It eliminated the outlier problem and provided a systematic and reasonably efficient procedure for timing pre-timed signal controllers.

Keywords: optimal signal plans, K-means clustering, pre-timed controllers, TOD breakpoints, cyclic traffic.



1 Introduction

Most of signalized intersections along arterials in the Kingdom of Saudi Arabia are equipped with pre-timed controllers. Although such controllers are technically inferior when compared to actuated and computerized controllers, they provide an acceptable level of performance for cyclic traffic. In such situation, the traffic is known to have well-defined characteristics which are repeated from one day to another. The pre-timed controllers do not rely on traffic sensors which makes them reliable [1]. The use of pre-timed controllers is not uncommon also in the USA [2]. Timing such controllers involves dividing the day into a number of time periods (intervals) where each period has its own distinctive traffic characteristics. An optimal timing plan is then developed for each time period. Since the traffic pattern along major urban arterials is usually cyclic and repetitive from one day to another, these optimal plans can be used for all weekdays. The literature is overwhelmed with techniques and software for developing optimal timing plans for a given traffic condition. However, there are few techniques for dividing the typical day into an optimum number of periods for which optimal timing plans can be developed [3]. One technique for accomplishing such task is the Time-of-Day (TOD) mode which is based on the concept of segmentation of time series data [2]. In the TOD mode, the day is divided into a pre-specified number of intervals such as AM peak period, PM period, off-peak period, night time period and so on in which the traffic characteristics are similar [1].

Selecting the optimal number of intervals is not a straightforward issue. Theoretically speaking, increasing the number of intervals will decrease the variation of traffic within each interval and the corresponding optimum traffic signal plan will better serve the traffic. Nevertheless, the frequent change in signal plans will be disruptive to the arterial operation [4]. For a given number of intervals, statistical clustering is usually used to determine the optimum TOD breakpoints (starting time) of each interval [1, 5, 6].

One of the commonly used statistical clustering techniques in this context is the K-means clustering. This technique will determine the TOD breakpoints of a pre-specified number of K clusters [7]. Subtractive clustering is a modified version of the K-means clustering which can also be used for this purpose [8]. In this method, clustering can start without predefining number of clusters. The method will suggest number of clusters and the TOD breakpoints. The number of clusters determined by subtractive clustering is not necessarily the optimum number of clusters but is usually considered as a good start point. Hierarchical clustering is another clustering technique that is used for such problem [9]. Ratrout [10] discussed all of the above clustering techniques and their possible utilization in TOD.

The main objective of this paper is to recommend a procedural framework to find out optimal number of timing plans (intervals) and optimal TOD breakpoints for each of these plans for a typical arterial in the study area taking into consideration the cyclic nature of traffic along urban arterials.



The study was conducted for a major arterial in Al-Khobar city, Saudi Arabia. The arterial consists of three major pre-timed coordinated signalized intersections, each of which consists of four protected phases. The traffic data was collected during summer season. A similar study was conducted recently on the same arterial but during the spring season [10]. The traffic volume along the arterial is higher in the spring season and has different characteristics from the summer season, especially in terms of percentage of turning volume. In this study, the issue of cyclic traffic is addressed more explicitly at the beginning of the proposed procedure contrary to the previous study conducted in the spring season [10].

2 Proposed methodology

The proposed procedural framework for finding TOD breakpoints aims to reduce the number of transitions from one timing plan to another and ensure better performance with respect to selected measures of effectiveness (MOEs). This paper also intends to reduce “outliers” of clusters. If some clusters jump around the adjacent TODs such that the TOD has to be changed frequently, they can be termed as “outliers” of clusters [6]. Simply this means that the statistical clustering technique proposes a short cluster (*i.e.* interval) between two relatively long intervals. Such short clusters are usually impractical in implementation. Hierarchical and K-means clustering techniques based approach to find out optimal TOD breakpoints suffers from “outliers” of clusters [1, 5]. In this proposed procedure, the approach volume, Z-score of approach volume, and time are used to solve the problem or at least reduce the number of transitions (*i.e.* outliers). The adopted methodology of this paper can be divided into five steps as shown in fig. 1.

2.1 Feature selection

In the proposed procedure, the transition cost is considered and emphasized in finding the optimum TOD breakpoints. If the use of certain traffic parameter (*i.e.* feature) such as approaches volume, results in 8 numbers of transitions corresponding to only 4 numbers of clusters, then it is causing 5 extra transitions compared to the absolute minimum number of transitions. Therefore, the parameters are selected based on their performance with respect to the number of transitions. This performance measure is considered for both hierarchical and K-means clustering techniques. If each clustering technique indicates different traffic parameters (parameters which provide less number of transitions), then both traffic parameters should be considered for the next phase. In this stage, only 2 to 12 numbers of clusters are considered to investigate the performance of the parameters. The approach traffic volume, Z-score of the traffic volume, and time variable are considered as the inputs of clustering technique. Initially, the representative 15-min approach traffic counts of intersections for the road network have to be prepared. Typically, the traffic counts are available in 15-min

interval which is widely used in many kinds of traffic engineering projects. The Z-score can be calculated by using the following formula:

$$Z\text{-score} = \frac{x - \bar{x}}{\sigma} \tag{1}$$

where x = 15-min traffic counts of a given approach, \bar{x} = mean of the 15-min traffic counts over 24 hours of all approaches for all studied intersections, and σ = standard deviation of the 15-min traffic counts over 24 hours of all approaches for all studied intersections.

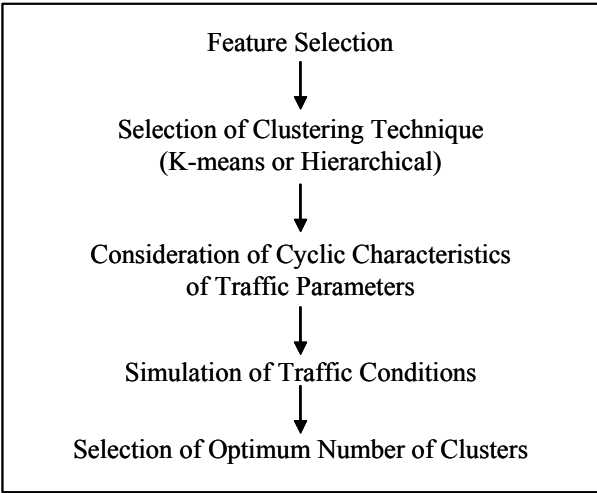


Figure 1: General procedural framework for determining optimum TOD breakpoints.

2.2 Selection of clustering technique

Hierarchical and K-means clustering techniques are evaluated with respect to silhouette value. The silhouette value for each point of a cluster is a measure of how similar that point is to the points in its own cluster compared to the points in other clusters, and it ranges from -1 to +1. The mean silhouette value of each cluster is considered for the sake of comparison. Higher silhouette values indicate better clusters. If one technique fails to consistently perform better than other technique, then both techniques can be considered for the next step of the procedure. For both clustering techniques, the same traffic features are used as input.

2.3 Consideration of cyclic characteristics of traffic parameters

In TOD breakpoints determination problem, the clustering technique should consider the cyclic characteristics of the traffic from one day to the next. The available clustering techniques do not consider the early period data of a day for

possible clustering with the late night traffic parameters. In reality, these traffic parameters are occurring consecutively and usually have very similar characteristics. To solve this problem, the 15-min traffic flow of an average day is considered twice successively to form continuous traffic data over 48 hours. Subtractive clustering algorithm is applied to provide optimum number of clusters and the starting and ending time of the specific cluster, which includes the time duration between midnight and the next morning periods. This specific cluster is then fixed in all further processing. The remaining part of a day is considered for a range of different numbers of clusters using the clustering technique that is selected from the previous step.

2.4 Simulation of traffic conditions

The selected traffic features are used to determine TOD breakpoints for certain practical numbers of clusters. In this stage, 2 to 5 numbers of clusters are considered. The range of clusters is selected considering the prevalent traffic condition and existing traffic controllers in Saudi Arabia. However, the proposed methodology can be easily extended for increased number of clusters. Optimum timing plans are developed for each TOD. For example, say an analysis reveals two TOD breakpoints at 07:00 and 13:00 which correspond to time T-1 and T-2 in fig. 2. In this case, three different optimum traffic signal plans, namely Timing Plan 1, Timing Plan 2, and Timing Plan 3, should be used, as shown in fig. 2.

Timing Plans 1, 2, and 3 are the optimum timing plans which can be developed by Synchro (or a similar kind of software) based on the design traffic

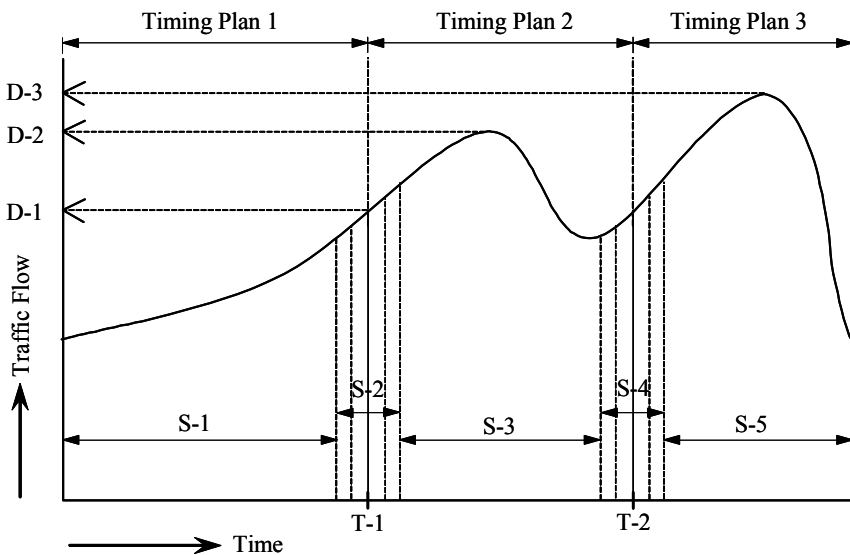


Figure 2: Graphical presentation of TOD breakpoints and relevant information [10].

flows D-1, D-2, and D-3, respectively, as shown graphically in the same figure. It is assumed that the effect of transition between successive traffic signal timing plans is limited to the duration of one hour. The implementation of new timing plan will usually disturb the coordination and progression between successive intersections. The cost of such disturbances is referred to in this study as the transition cost. SimTraffic software is used to evaluate the performance of the transition periods in this proposed methodology. The SimTraffic manual explains the procedure followed by SimTraffic for this analysis [11]. To find out the transition cost at a certain time such as 07:00 which is referred to as T-1 in fig. 2, the observed traffic counts from 06:30 to 06:45 and from 06:45 to 07:00 should be simulated using SimTraffic under Timing Plan 1. The observed traffic counts from 07:00 to 07:15 and from 07:15 to 07:30 are simulated under Timing Plan 2. The time interval from 06:30 to 07:30 is symbolized by S-2 in fig. 2.

The remaining periods of each TOD such as S-1 and S-3 will be similarly simulated under Timing Plan 1 and Timing Plan 2, respectively. The performance measures (MOEs) of all time periods over 24 hours including transition periods (S-1, S-2, S-3, S-4 and S-5 in fig. 2) should be added.

2.5 Selection of optimum number of clusters

The optimum number of clusters and TOD should be selected by comparing the values of a selected MOE for the investigated clusters. A range from 2 to 5 clusters is reasonable and practical for the local traffic conditions in the study area. The number of clusters which will ensure the best values of MOEs will be considered as the optimum number of clusters and will govern the selection of TOD. The recommended procedure suggests using multiple microscopic simulation runs in SimTraffic by adopting different random seeds to enhance the reliability of the results.

3 Implementation

The methodology described above was implemented on the studied arterial as discussed in the following subsections.

3.1 Feature selection

In order to determine the appropriate feature of the clustering, approach volume, Z-score of volume, and time variable expressed in hours were all investigated. Table 1 shows the number of transitions for different numbers of clusters using approach volume, approach volume with time, and Z-score of approach volume with time. It is clear that the Z-score along with time ensured the absolute minimum number of transitions for the whole range of clusters investigated (*i.e.* 2–12). Table 2 shows the TOD breakpoints obtained using the approach volume for clusters 8 to 12, as an example. It was observed that some of the clusters lasted less than an hour, which made those TOD breakpoints impractical. The same behaviour was also observed for clusters 3 to 8. Adding the time variable to

the approach volume did not improve the situation. On the other hand, using the Z-score of approach volume with time produced very reasonable TOD breakpoints. With these two variables, all clusters lasted for more than one hour as can be seen in table 3 for clusters 8 to 12. The same behaviour was also noticed for all other clusters less than 8. Therefore, Z-score of approach volume and time were used as the features to perform the clustering technique. It should be mentioned that all the above results did not significantly change when hierarchical clustering was used in lieu of the K-means clustering technique. The use of the Z-score of approach volume with time is still the best and practical choice even if hierarchical clustering is used. These results are consistent with the previous study conducted on this arterial during spring season [10].

3.2 Selection of clustering technique

In this research, hierarchical and K-means clustering techniques were studied. In order to determine the better performing technique, silhouette values were used. The silhouette value for each point of a cluster is a measure of how similar that point is to the points in its own cluster compared to the points in the other clusters. Clusters 2 to 12 were studied and the mean silhouette value of each cluster was considered for comparison purposes. For both clustering techniques, the same traffic features (Z-score of volume and time variable expressed in hours) were used as input. It was observed that when the number of clusters is 2, the silhouette values of both clustering techniques are equal. For all other investigated clusters, the silhouette value of K-means clustering was consistently larger than that of hierarchical clustering. On average, it was about 70% larger.

Table 1: Number of transitions obtained using different features in K-means clustering technique.

Number of clusters	Number of transitions using volume only	Number of transitions using volume and time	Number of transitions using time and Z-score
2	1	1	1
3	8	8	2
4	12	12	3
5	15	15	4
6	16	16	5
7	14	14	6
8	23	23	7
9	22	22	8
10	27	24	9
11	24	24	10
12	27	26	11

Table 2: The time of transitions obtained using volume in K-means clustering technique.

	Number of Clusters				
	8	9	10	11	12
Start Time of Transition	1.5	1.5	1.5	1.5	1.5
	5	5	5	5	5
	5.75	5.75	5.75	6	6
	6.75	6.75	6.75	6.75	6.75
	7.5	7.5	7.5	7.5	7.5
	8.75	9	9	9	9
	10.75	10.75	10.75	10.75	10.75
	11	11	11	11	11
	11.25	11.25	11.25	11.25	11.25
	11.75	11.75	11.75	11.75	11.75
	13	13	12.5	12.5	12.5
	15.5	15.5	12.75	12.75	12.75
	15.75	15.75	13	13	13
	16	16	14.5	15.5	13.25
	16.5	16.75	15.25	15.75	13.5
	17	17	15.5	16	14
	17.75	18	15.75	16.75	15.5
	18	18.75	16	17	15.75
	18.75	19.25	16.75	18	16
	19.25	20.25	17	18.75	16.75
	20.25	21.75	18	19.25	17
	21.75 ^a	23.25	18.75	20.25	18
	23		19.25	21.75	18.75
			20.25	23	19.25
			21.75		20.25
			23.25		21.75
			23.50		23.00

Note: ^a21.75 indicates time 09:45 PM.



Table 3: The time of transitions obtained using Z-score of volume with time variable in K-means clustering technique.

Start Time of Transition	Number of Clusters				
	8	9	10	11	12
	2.5	2.5	2.5	2	2
	5.75	5.75	5.75	4.75	4.75
	7.5	7.5	7.5	6	6
	11.75	10.75	9	7.5	7.5
	15.5	13	11.25	9	9
	18.75	15.5	13.25	11.25	11.25
	21.75	18.75	15.5	13.25	13.25
		21.75	18.75	15.5	15.5 ^a
			21.75	18.75	17
				21.75	19
					21.75

Note: ^a15.5 indicates time 03:30 PM.

This implies that K-means ensured better homogeneity of features within clusters and increased disparity among clusters which are the desirable characteristics of a cluster technique. Therefore, K-means was selected as the cluster technique to determine TOD breakpoints for the studied arterial.

3.3 Consideration of cyclic characteristics of traffic parameters

The results obtained in this study using Z-score and time variable solved the problem of “unclean” clusters. This clustering approach, however, will not consider the early period data of a day for possible clustering with the late night traffic flow data of the previous day. In reality, the traffic flows during these two periods are occurring consecutively and usually have similar characteristics. In order to solve this problem, the Z-score of approach volume during an average day was considered twice successively to form continuous data string over 48 hours. Subtractive clustering technique was then applied on the 48 hours data. The results indicated that the optimum number of clusters is seven. The seven clusters are graphically shown in fig. 3. This clustering effort identified the cluster which encompasses the time period between late night and continued to the next morning. This cluster was between 22:00 PM and 06:00 AM, which is shown as cluster number 3 in fig. 3. This obtained cluster was then fixed and used over the original 24-hour data set. The features of the remaining part, between 06:00 AM and 22:00 PM, were subjected to K-means clustering with different numbers of clusters. It was previously assumed that the practical numbers of clusters in the study area are 2 to 5. One cluster (between 22:00 PM

and 06:00 AM) is now fixed; therefore, the remaining part of the day was subjected to 2, 3, and 4 clusters.

3.4 Simulation of traffic conditions

The optimum traffic signal plans were determined by using Synchro for all studied clusters. If the number of clusters is two, then two optimum timing plans were determined by using the highest hourly volume (*i.e.* design volume) for each cluster. For three clusters, three signal timing plans were determined and so on. Then the whole 24-hour period was simulated. The effect of transition period was also considered during simulation as described previously.

3.5 Selection of optimum number of clusters

SimTraffic was used to evaluate the performance of the selected number of clusters with respect to travel time (hr), total delays, total stops, and fuel used. The performance of each number of cluster was evaluated for the three studied intersections, and the traffic was simulated for 24 hours.

Table 4 shows the aggregated performance of the studied arterial with different numbers of clusters. The values of the second column are not reliable and should not be compared with the values of other columns. During the simulation with two clusters, vehicles experienced significant congestion and around 500 vehicles failed to enter the network. SimTraffic is not typically

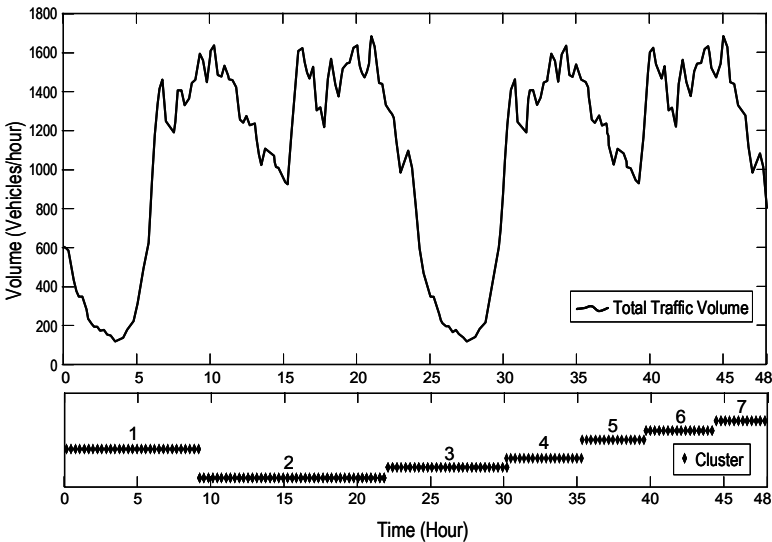


Figure 3: TOD breakpoints obtained considering time and Z-score of volume data for two consecutive days.



intended to simulate such congested traffic condition. Consequently, results for two clusters should be discarded. In this case study, it is clear that the optimum number of clusters is 5 based on the values shown in table 4. The duration of the TODs corresponding to 5 numbers of clusters is shown in table 5.

Table 4: Performance evaluation of different numbers of clusters using SimTraffic.

Number of clusters	2	3	4	5
Travel Time (hr)	11,796.6	14,297.8	13,259.1	11,945.4
Total Delay (hr)	7,585.9	10222.1	9,048.1	7,731.7
Total Stops	335,649	399,915	361,936	343,053
Fuel Used (l)	73,622.6	77,554.9	76,711.8	73,681.7

Table 5: Duration of TODs for optimum five clusters.

Duration of TODs				
TOD-1	TOD-2	TOD-3	TOD-4	TOD-5
00:00 to 06:00 and 22:00 to 24:00	06:00 to 11:30	11:30 to 15:45	15:45 to 19:00	19:00 to 22:00

4 Conclusions

This paper provides a systematic objective methodology for determining the optimum number of timing plans for major urban arterials serving daily cyclic traffic in Al-Khobar city, Saudi Arabia. Adopting Z-score of traffic volume with time variable expressed in hours resolved the “unclean clusters” problem typically encountered in statistical clustering of traffic data. The issue of cyclic characteristics of traffic flow is explicitly addressed in the early part of the proposed methodology through subtractive clustering conducted on 48 hours traffic data. For the studied arterial and the range of clusters investigated, it was found that 5 is the optimum number of clusters over 24 hours traffic data. Consequently, 5 different optimum timing plans are required to optimize the traffic operation along the studied arterial. The proposed methodology can be smoothly used over a broader range of clusters and more complicated road network. The results of this paper are very comparable to a previous study conducted on the same arterial under different traffic conditions.

Acknowledgements

The author would like to acknowledge the financial support provided by King Fahd University of Petroleum & Minerals (KFUPM) and also the support of the Department of Civil Engineering at KFUPM in carrying out this study.

References

- [1] Wang, X., Cottrell, W. & Sichun Mu, Using K-means clustering to identify time-of-day breakpoints for traffic signal timing plans. *Proc. of 8th Int. IEEE Conf. on Intelligent Transportation Systems*, Vienna, Austria, pp. 586 591, 2005.
- [2] Gordon, R.L., Reiss, R.A., Haenel, H., Case, E.R., French, R.L., Mohaddes, A. & Wolcott, R., Traffic control systems handbook, No. FHWA-SA-95-032, U.S. Dept. of Transportation, USA, 1996.
- [3] Hauser, T., Scherer, W.T. & Smith, B.L., Signal system data mining. Report, Center for Transportation Studies, University of Virginia, USA, 2000.
- [4] Hua, J. & Faghri, A., Dynamic traffic pattern classification using artificial neural networks. *Transportation Research Record: Journal of Transportation Research Board*, No. 1399, pp. 14 19, 1993.
- [5] Smith, B.L., Scherer, W.T. & Hauser, T.A., Data-mining tools for the support of signal-timing plan development. *Transportation Research Record: Journal of Transportation Research Board*, No. 1768, pp. 141 147, 2001.
- [6] Park, B. & Lee, J., A procedure for determining time-of-day breakpoints for coordinated actuated traffic signal systems. *KSCE Journal of Civil Engineering*, **12(Issue 1)**, pp. 37 44, 2008.
- [7] Chiang, M. M-T. & Mirkin, B., Experiments for the number of clusters in K-means. *Proc. of 13th Portuguese Conf. on Progress in Artificial Intelligence*, Portugal, pp. 395 406, 2007.
- [8] Chiu, S., Fuzzy model identification based on cluster estimation. *Journal of Intelligent & Fuzzy Systems*, **2(3)**, Sept. 1994.
- [9] Park, B., Santra, P., Yun, I. & Lee, D., Optimization of time-of-day breakpoints for better traffic signal control. *Transportation Research Record: Journal of Transportation Research Board*, No. 1867, pp. 217 223, 2004.
- [10] Ratrout, N., Subtractive clustering based K-means technique for determining optimum time-of-day breakpoints. Accepted for publication in *Journal of Computing in Civil Engineering*, ASCE, 2010.
- [11] Husch, D. & Albeck, J., *Trafficware Synchro Studio 7 User Guide*, 2006.

