Simulation Optimization of Urban Arterial Signals via Simultaneous Perturbation Stochastic Approximation (SPSA)

Ning Yang (ningyang@umd.edu) and Yue Liu
Department of Civil and Environmental Engineering
University of Maryland, College Park
Maryland, 20742, USA

ABSTRACT

In this paper, we develop a simulation optimization procedure for optimizing the urban arterial traffic signal timings including a bunch of sequential intersections. The system performance is estimated via a stochastic discrete-event meso-scopic traffic simulator, and a gradient-based search algorithm on stochastic approximation is applied to give the optimal signal timings. Simultaneous perturbation analysis is used to derive both left-hand and right-hand gradient estimators of the system performance with respect to the cycle lengths, green splits, and green offsets for those intersections within the arterial. Numerical experiments show that the meso-scopic traffic simulator provides reasonable system performance in much less running time if properly calibrated, compared with a widely-used commercial traffic microscopic simulation program CORSIM. In particular, for all scenarios designed, the optimizer converges to optimal signal timing plans which significantly increase the system performance.

INTRODUCTION

The vehicular delay at signalized intersections, which increases the travel time as well as reduces speed and reliability, is an obstacle that has a detrimental effect on cost-effectiveness of transportation system (6). Therefore, it has been the traffic engineer's endeavor to quantify delay and optimize the signal system to increase the operational efficiency of the urban traffic system.

Traffic simulation is an important tool for modeling the operation of dynamic traffic systems and helps analyze the causes and potential solutions of traffic problems such as congestion and safety. Various simulation models and optimization techniques have evolved and aided traffic engineer in the optimization process (1-6). The level of detail in simulation models ranges from macroscopic via meso-scopic to microscopic. Most of the existing traffic signal optimization programs, such as SYNCHRO, TRANSYT-7F, and PASSER-II(1), rely on deterministic and macroscopic simulation programs. One drawback of such applications is that the simulation program does not reflect real-world conditions (e.g. the left-turn bay capacity constraint). Microscopic simulation programs, such as CORSIM and VISSIM (6), can emulate traffic at signalized intersections in details. However, car-following and lane-changing logics are complicated to simulate and integrating signal optimization with this class of simulation is quite time and cost consuming. Meso-scopic models, which fill the gap between the aggregate level approach of macroscopic models and the individual tracking approach of the microscopic ones, can also simulate signal timing by translating signal states to road segment capacities. Examples are DynaMIT, DYNASMART, MITSIM and METROPOLIS (9). When the precise level more than macroscopic simulation is desirable and the detail of microscopic simulation is infeasible due to a large network or resources available are limited, meso-scopic models might be better choices.

To solve the optimization problem through the simulation approach, perturbation analysis and stochastic approximation algorithms could be used. SPSA is an iterative technique for optimization of complex, multidimensional, stochastic systems in areas such as statistical parameter estimation, discrete-event systems and feedback control. Different from the Finite Difference Stochastic Approximation (FDSA) which has the same central idea coming from the classical definition of a derivative, SPSA requires only two objective function evaluations per gradient estimate. Thus it is not surprising that SPSA is extremely attractive for high dimensional problems. For stochastic simulation models, multiple simulation runs should be conducted for each evaluation under a given scenario to obtain statistically significant evaluation results. Information of interest is extracted and average from the outputs of individual simulation runs.

SIGNAL OPTIMIZATION MODEL

The signal timing optimization problem is formulated as a minimization problem. The objective function is to minimize the expected total vehicle delay in the system for a certain time period. We try to simultaneously optimize the signal plans for all intersections controlled by one central signal controller. The inputs of the system are: (1) the cycle length for all coordinated signals; (2) the offset of each signal, (3) the green splits for each phase for each traffic signal. For the convenience of solution generation, some variable which has a value between 0 and 1 are generated and the decision variables are specified through a transformation step. Due to safety and operation reasons, some constraints, such as maximum cycle length, should be considered.

The objective function and constraints are expressed in the following equations:

Minimize $J(\hat{\theta}_k) = E[L(\hat{\theta}_k)]$

Subject to:

$\hat{\theta}_k = [\theta_1, \hat{\theta}_2, \hat{\theta}_3]$

$\theta_1 = \theta_1$

$\hat{\theta}_2 = [\theta_{2,1}, \ldots.., \theta_{2,i}, \ldots..., \theta_{2,S}]$ $\qquad i = 1,2,\ldots S$

$\hat{\theta}_3 = [\hat{\theta}_{3,1}, \ldots.., \hat{\theta}_{3,i}, \ldots..., \hat{\theta}_{3,S}]$ $\qquad i = 1,2,\ldots S$

$\hat{\theta}_{3,i} = [\theta_{3,i,1}, \theta_{3,i,j}, \ldots..., \theta_{3,i,Ni}]$ $\qquad i = 1,2,\ldots S \quad j = 1,2,\ldots N_i$

$0 < \theta_{1,\min} \le \theta_1 \le 1$

$0 \le \theta_{2,i} \le 1$

$0 < \theta_{3,\min} \le \theta_{3,i,j} \le \theta_{3,\max} < 1$

where,

$J(\hat{\theta}_k)$ = the system performance with respect to the decision variable vector $\hat{\theta}_k$;

θ_1 = a scalar equal to the actual cycle length / the maximum cycle length;

$\theta_{2,i}$ = a scalar equal to the actual offset of signal *i* / the actual cycle length;

$\theta_{3,i,j}$ = a relative value of the green time for the movement *j* at signal *i*;

$\theta_{1,\min}$ = constraint on the minimum cycle length;

$\theta_{3,\min}, \theta_{3,\max}$ = constraint on the relative value of $\theta_{3,i,j}$ for obtaining a reasonable green split.

The transformation steps are shown as follows:
Cycle length: $C = C_{\max} \cdot \theta_1$
Offset of the signal *i*: $C_{\max} \cdot \theta_1 \cdot \theta_{2,i}$
Split for movement *j* at signal *i*: $G_{ij} = C_{\max} \cdot \theta_1 \cdot \theta_{ij} / \sum_j \theta_{ij}$

MESOSCOPIC TRAFIC SIMULATOR

In this study, a meso-scopic traffic simulator is developed to emulate the performance of a traffic system and designs of traffic signal timing plans. In the meso-scopic simulation model, individual vehicles are tracked but they are grouped into platoons which control their behaviors. Vehicles can enter and leave platoons when needed. The vehicle platoons move over roadway segments according to predefined speed-density relations instead of complicated car-following and lane-changing logic. The speed of the vehicles is determined by the platoons, not by the individual drivers' decisions.

The main structure of the proposed simulation model is an iteration of functions at specified frequencies (time-based). The movements of vehicles and the update of system status are processed at a specific time interval.

The simulation starts with the loading of the road network description, travel demand information, the simulation parameters, and signal timing plans. Those inputs provide the operational environment under which the system will be evaluated. Once the simulator is initialized, an iterative procedure begins. Firstly, new vehicles which would enter the system during the time interval are generated. The vehicle characteristics, including the vehicle type, the origin and destination of the vehicle, the choice of the route, etc., are assigned to each new vehicle according to input travel demand information and user-defined distributions. New vehicles will be appended to a virtual queue in the entry links and loaded into the system with FIFO criteria. Then, the vehicle movements will be simulated in two phases: an update phase and an advance phase. In the update phase, capacities for each segment and turning movements at intersections are updated based on the traffic signal settings and the volume and composition of approaching traffic flows. The speeds of traffic platoons are calculated according to a speed-density model (9). Individual vehicles' speeds are calculated by interpolation according to their positions in the platoon. The advance phase moves individual vehicles downstream and updates their speeds and positions. As spacing between vehicles changes, platoons are split and combined. Vehicles that arrive at their destinations will be removed from the system.

The simulator architecture is shown in Fig. 1.

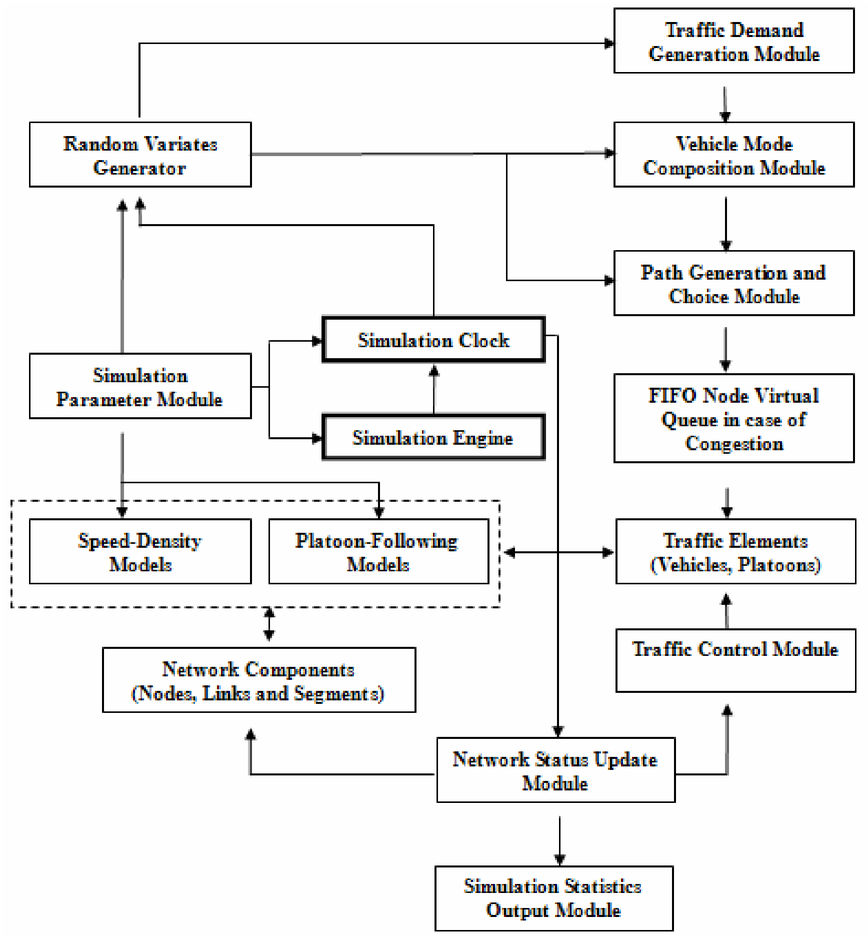

Fig. 1 Simulator Architecture

OPTIMIZATION VIA SPSA

In this section, firstly, simultaneous perturbation analysis (SP) is used to derive both left-hand and right-hand gradient estimators of the system performance from the meso-scopic traffic simulator with respect to the cycle lengths, green splits, and green offsets. Then, a gradient-based search algorithm on stochastic approximation (SA) is applied to give the optimal signal timings.

Simultaneous Perturbation (SP)

We performed both left and right-hand perturbation evaluation for the system performance function. The left-hand evaluation is:

$$E[L(\hat{\theta}_k)]^- = E[L(\hat{\theta}_k - c_k\Delta_k)]$$

And, the right-hand estimator is:

$$E[L(\hat{\theta}_k)]^+ = E[L(\hat{\theta}_k + c_k\Delta_k)]$$

Where, $c_k = c/(k+1)^\gamma$, and Δ_k is a p-dimensional random perturbation vector, whose component is to use a Bernoulli ± 1 distribution with probability of 1/2 for each ± 1 outcome.

Stochastic Approximation (SA)

SA is a gradient-based stochastic optimization algorithm, where the "best guess" of the optimal parameter is updated iteratively based on the estimate of the gradient of the performance measure with respect to the parameter (7).The general form of SA is:

$$t_{k+1}^{*} = \prod_{(0,t_p)} (t_k^{*} - a_k \nabla J_k)$$

Where, t_k^{*} is the parameter value at the beginning of iteration n, an is a positive sequence of step sizes, ∇J_k is the estimate of $\nabla J_k(t_k^{*})$, the gradient of J_k at parameter value t_k^{*} and, $\prod_{\Omega}$ is the projection onto Ω. $\prod_{\Omega}$ keeps the parameter within the valid range of values. In this implementation of the SA algorithm, we have:

$$\hat{\theta}_{k+1} = \hat{\theta}_k - a_k g_k(\hat{\theta}_k)$$

Where, $g_k(\hat{\theta}_k)$ is the gradient estimator for the performance function at the iteration k with:

$$g_k(\hat{\theta}_k) = \frac{E[L(\hat{\theta}_k)]^{+} - E[L(\hat{\theta}_k)]^{-}}{2c_k} \begin{bmatrix} \Delta_{k1}^{-1} \\ \Delta_{k2}^{-1} \\ \Delta_{kp}^{-1} \end{bmatrix}$$

In summary, the SPSA algorithm is implemented as follows, as shown in Fig. 2:

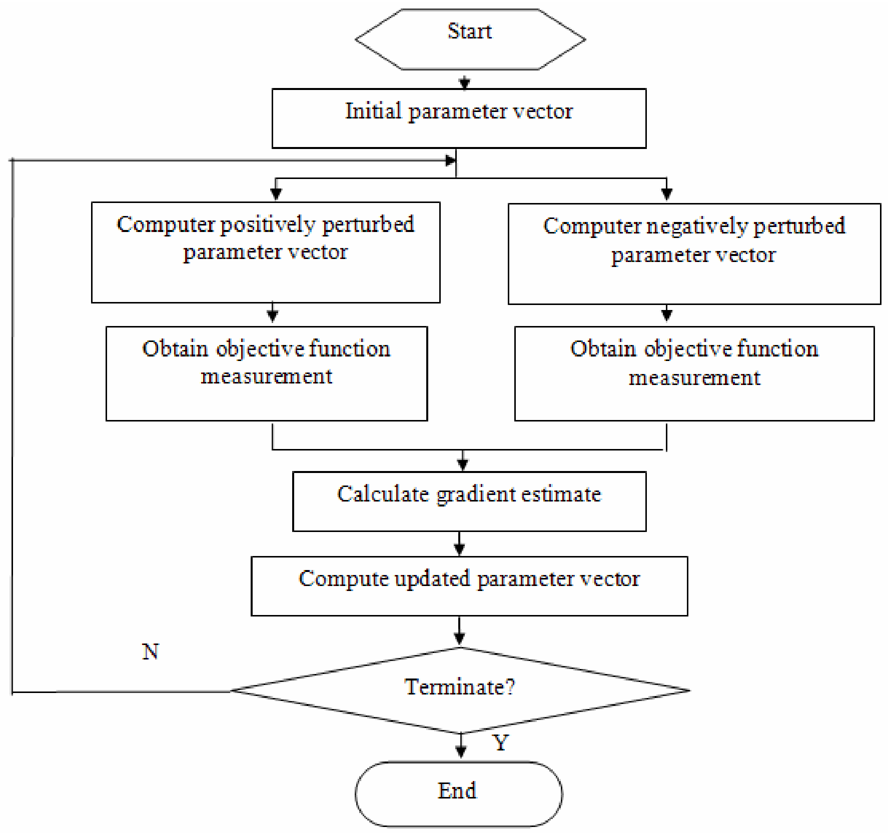

Fig. 2 SPSA Algorithm

Step 1: Pick initial guess $\hat{\theta}_0$ and non-negative coefficients a, c, A, α, and γ in the SPSA gain sequences $a_k = a/(A+k+1)^{\alpha}$ and $c_k = c/(k+1)^{\gamma}$. Practically, efficient values for α and γ are set to

0.602 and 0.101 respectively. The values of a, A and c may be determined based on the practical guidelines given in (8);

Step 2: Generate a p-dimensional random perturbation vector Δ_k, whose component is to use a Bernoulli ± 1distribution with probability of 1/2 for each ± 1outcome;

Step 3: Obtain two estimation of the performance function $E[L(\hat{\theta}_k - c_k\Delta_k)]$and $E[L(\hat{\theta}_k + c_k\Delta_k)]$;

Step 4: Generate the approximation to the unknown gradient $g_k(\hat{\theta}_k) = \frac{E[L(\hat{\theta}_k)]^+ - E[L(\hat{\theta}_k)]^-}{2c_k}\begin{bmatrix}\Delta_{k1}^{-1}\\ \Delta_{k2}^{-1}\\ \Delta_{kp}^{-1}\end{bmatrix}$;

Step 5: Update $\hat{\theta}_k$ to a new value $\hat{\theta}_{k+1} = \hat{\theta}_k - a_k g_k(\hat{\theta}_k)$, if satisfy the convergence criteria, then End; else go to step 2;

End.

NUMERICAL EXPERIMENT

Test Network

A hypothetical network including three intersections is built for simulation study in this numerical experiment (see Fig. 3). This network includes one major arterial (O1D1) and three minor streets. Thus, there are 3 intersections in this network. The free flow speed along the major arterial is 40 mph and that along the minor streets is 30 mph. The length of each link is set to 700 ft. Each link has 2 lanes. To simplify the optimization problem, all internal streets are one-way roads and it is assumed that no left-turn and right-turn movements would be permitted at each intersection. Therefore, there are only two phases for each pre-timed traffic signal. Four O-D pairs are considered. Traffic volumes are loaded from O1, O2, O3 and O4.

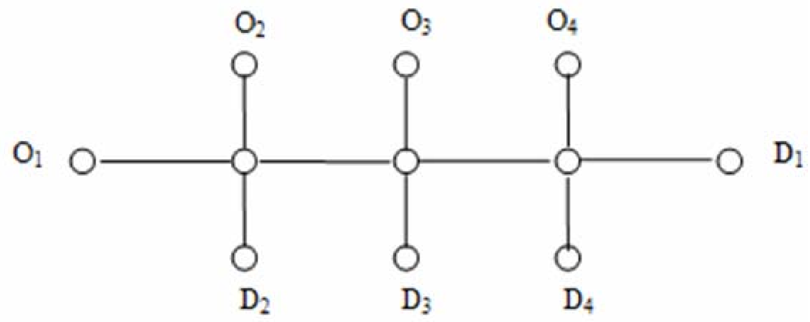

Fig. 3 Test Network for Numerical Experiment

Experiment Design

Seven scenarios with incremental increasing volumes (200vph) on major arterial and fixed volumes on minor streets are designed to test the proposed simulation-based optimization method. The volumes for each O-D pair in scenario 1 are shown in Table 1. Initial signal plans were generated using the Webster Formula for all the 7 scenarios (offsets were set to 0s, and splits were assigned according to the saturation degree). All three signals in the network use the same initial timing plan. Constraints on signal timing parameters are summarized as follows: $c_{\max} = 150s$, $\theta_{1,\min} = 0.3$, $\theta_{3,\min} = 0.1$, and $\theta_{3,\max} = 0.9$. For each scenario, one hour system operation will be simulated.

TABLE I OD matrix for scenario 1 (in vph)

	D_1	D_2	D_3	D_4
O_1	1200	0	0	0
O_2	0	600	0	0
O_3	0	0	600	0
O_4	0	0	0	600

Simulation Validation

Before applying the meso-scopic simulation model to optimization, it is necessary to check for any inconsistency in the model behavior and executed logic. This is done through a number of simulation runs using artificial data to ensure that the model behaves correctly and the logic is executed as designed. Following this verification, we will check whether the proposed model can output reasonable results. Since the test network is hypothetic, no field data is available. For validation purpose, the test network is coded in both the proposed model and a widely-used commercial microscopic simulation program CORSIM. For the initial signal timing plan, the results from meso-scopic model and those from CORSIM for all 7 scenarios are compared with each other with 10 simulation replications for each scenario. Table 2 summarizes the comparison of the total delays from the two models.

TABLE 2 Difference of total delay between meso-scopic model and corsim under 10 replications (in veh hr)

Scenario #	1	2	3	4	5	6	7
Meso-scopic Model	23.69	25.35	27.96	31.20	36.67	61.56	141.85
CORSIM	26.87	28.94	31.21	33.80	36.97	41.56	47.45
Mean Difference	-3.17	-3.58	-3.25	-2.60	-0.30	20.00	94.40
Half Length 95% C.I.	0.17	0.39	0.23	0.43	0.24	0.60	0.45

From the comparison results, we can observe that:

(1) Under light and medium traffic conditions, the total delays obtained from the two models are close to each other. Although the meso-scopic model provides slightly lower values, the differences are still in acceptable range.

(2) Under heavy traffic conditions, the total delays from the meso-scopic model are significantly higher than those from CORSIM. Two possible reasons may cause this situation: (a) CORSIM will neglect the vehicles which cannot enter the system due to congestion while the meso-scopic model can store these vehicles in a virtual queue. Therefore, more delays will be recorded; (b) With the default parameters, CORSIM and the meso-scopic model may

get different roadway and intersection capacities. If the capacity calculated in the meso-scopic model is lower than that simulated in CORSIM, delay will sharply increase in the meso-scopic model. Further calibration of model parameters using filed data may be need in further study.

Optimization Results

The optimization process via the propsed SPSA algorithm based on meso-scopic model is performed for all the 7 designed scenarios. All of them converge to an optimized solution within less than 50 iterations, shown in Fig. 4.

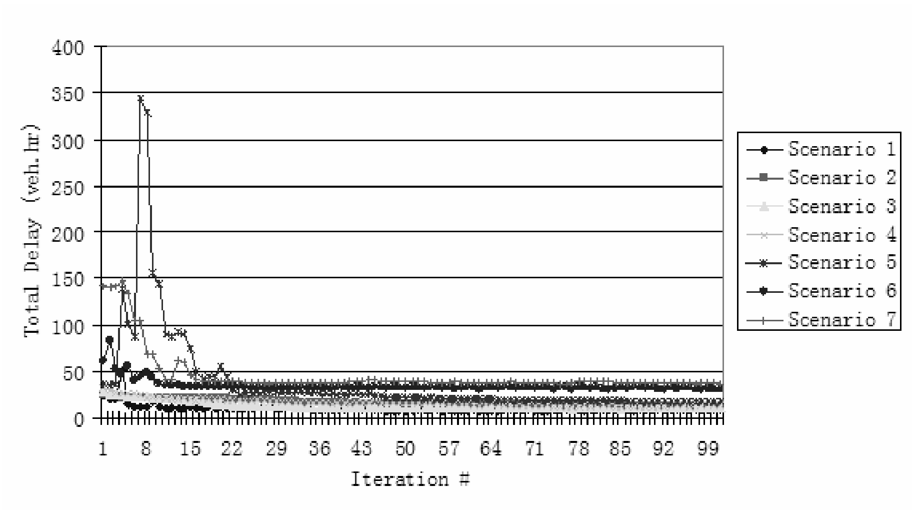

Fig. 4 Optimization Paths for All 7 Scenarios Using SPSA

The optimized signal timing plans for each scenario are listed in Table 3. It shows that:

TABLE 3 Optimized Signal Timing Plans for All 7 Scenarios (in secs)

Signal#	Timing Plan	S1	S2	S3	S4	S5	S6	S7
1	Cycle length	45	51	45	51	50	102	116
	Offset	9	0	2	1	4	1	1
	Major arterial green	32	36	32	37	37	78	86
	Minor street green	13	15	13	14	13	24	30
2	Cycle length	45	51	45	51	50	102	116
	Offset	31	2	7	1	0	6	1
	Major arterial green	25	36	33	37	37	76	86
	Minor street green	20	15	12	14	13	26	30
3	Cycle length	45	51	45	51	50	102	116
	Offset	1	7	16	17	47	3	1
	Major arterial green	25	36	33	36	36	76	86
	Minor street green	20	15	12	15	14	26	30

(1) Green times for the major arterial increases with the increase of traffic demand along this direction.

(2) The cycle length increases when traffic flows become higher in the system.
(3) Offsets become less important when the traffic condition gets congested.
(4) The SPSA algorithm converges to the optimal results quickly and stably.
(5) The optimal signal timing plans from the proposed method can significantly decrease the total delays of the system compared to the initial timing plans, as shown in Fig 5.

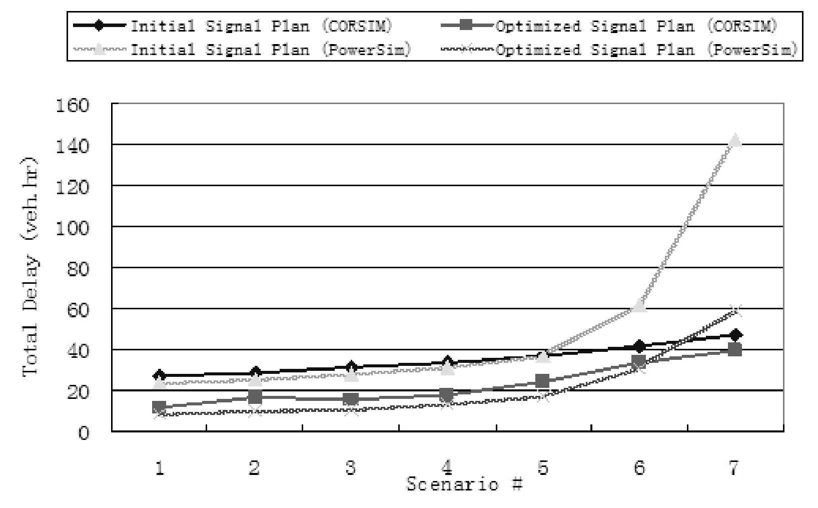

Fig. 5 Comparison of Initial Signal Plan and Optimized Signal Plans for All 7 Scenarios

CONCLUSION AND FUTURE STUDIES

In this paper, a meso-scopic simulation model which can deal with signalized traffic network is developed. The model has the ability to represent the stochastic urban traffic environment in a moderate detail level. The meso-scopic simulation model can output measures of effectiveness close to microscopic simulation results while in shorter running times. An optimization method combining the proposed meso-scopic simulation and the SPSA is proposed to optimize the signal timings and coordination for urban arterials. The applicability and effectiveness of the methodology is tested in a numerical experiment. Satisfactory optimal signal timing plans are obtained for the test network for different scenarios.

Several works can be performed in future studies:
(1) Extensive calibration and validation of the proposed meso-scopic simulation model shoud be performed.
(2) The proposed meso-scopic simulation model and the optimization algorithm would be tested in a real traffic network or more complex traffic conditions.
(3) Sensitivity analysis of the optimization algorithm on different traffic patterns, simulation parameters and optimization parameters are expected in the future.

REFERENCES

(1). Burghout, W. “Hybrid Microscopic-mesoscopic Traffic Simulation”. Ph. D. Dissertation, Royal Institute of Technology, Stockbolm, Sweden, 2004.
(2). Hewage, K. N. and Ruwanpura, J. Y. “Optimization of Traffic Signal Light Timing Using Simulation”, in Proc. of the 2004 Winter Simulation Conf., pp. 1428-1433.

(3). Kamarajugadda, A. and Park, B. B. “Stochastic Traffic Signal Timing Optimization Research Report No. UVACTS-15-0-44, Center for Transportation Studies, Department of Civil Engineering, University of Virginia, 2003.
(4). Li, M. T. and Gan, A. C. “Signal Timing Optimization for Oversaturated Networks Using TRNSYT-7F”, Transportation Research Record, No. 1683, pp. 118-126, National Research Council, Washington, D.C., 1999.
(5). Park, B., Messer, C. J. Urbanik, T. “Traffic Signal Optimization Program for Oversaturated Conditions: Genetic Algorithm Approach”, Transportation Research Record, No. 1683, pp. 133-142, National Research Council, Washington, D.C., 1999.
(6). Park, B., Rouphail, N. M. and Sacks, J. “Assessment of a Stochastic Signal Optimization Method using Micro-simulation”, Transportation Research Record, No. 1748, pp. 40-45, National Research Council, Washington, D.C., 2000.
(7). Panayiotou, C. G., Howell, W. C. and Fu, M. “Online Traffic Light Control through Gradient Estimation Using Stochastic fluid models”, Proc. of the 16th IFAC World Congress, Prague, 2005.
(8). Spall, J.C. “Implementation of the Simultaneous perturbation algorithm for stochastic optimization”, IEEE Transaction on Aerospace and Electronic Systems, Vol.3, No.3, pp. 817-823, 1998.
(9). Yang, Q. “A Simulation Laboratory for Evaluation of Dynamic Traffic Management Systems”, Ph. D. Dissertation, Massachusetts Institute of Technology, 1997.