SOLVING THE INTERACTION BETWEEN SIGNAL CONTROL AND TRAFFIC ASSIGNMENT: GLOBAL, LOCAL AND ITERATIVE SEARCHES

Chungwon Lee
Research Associate, Ph.D.
Department of Urban Transportation
Seoul Development Institute
4-5 yejang chung Seoul 100-250 Korea
Fax: 1-512-475-8744
chungwon@mail.utexas.edu

Randy B. Machemehl
Professor, Ph.D.
Department of Civil Engineering
University of Texas at Austin
ECJ 6.9, UT-Austin, Tx 78712, US
Fax: 1-512-475-8744
rbm@mail.utexas.edu

ABSTRACT

Two stochastic global optimization methods, simulated annealing (SA) and a genetic algorithm (GA), are applied to the combining traffic signal control and assignment problem seeking globally optimal signal settings and an equilibrium flow pattern. The two methods aim at overcoming the nonconvexity of the problem. A local search algorithm is implemented. An iterative approach to perform assignment and control optimization sequentially finds mutually consistent points. Link performance is described by the Webster curve. OD matrix is assumed fixed and green time ratios are decision variables. For the total time minimization policy constrained by user equilibrium, the four approaches are implemented. Numerical comparison is performed on different size networks at different OD levels. The demand level and network size affect the relative performance of the approaches. While local and iterative searches converge fast, SA and GA reduce the risk of stalling in low quality solutions.

1. INTRODUCTION

From the transportation planning perspective, traffic assignment models are used to forecast network flow patterns, commonly assuming that capacities decided by network supply parameters such as signal settings are fixed during a short time period with a given particular origin-destination matrix (OD). On the other hand, from the transportation engineering perspective, network flow patterns are commonly assumed fixed during a short time period and control parameters are optimized in order to improve some performance index for prevailing flow patterns. The input flow patterns must either be observed or forecast through traffic assignment.

The two processes, traffic assignment and signal optimization, are usually dealt with separately, however, the processes mutually influence each other. This mutual interaction can be explicitly considered by effective integration of these two processes, producing the so called combined control and assignment problem. When drivers follow Wardrop (1952)'s first principle, i.e., user equilibrium (UE) flow, the problem is called the equilibrium network traffic signal setting problem, which is normally nonconvex and obtaining explicit gradient information for any gradient based algorithm application is difficult.

Changing signal settings may stimulate drivers to adjust route choices; however, changing flow may suggest re-setting signals. Allsop (1974) first noted the necessity of combing signal calculations and traffic assignment by pointing out that network traffic routing according to Wardrop's first principle is dependent on signal timings and should ideally be regarded simultaneously with signal timing. Gartner (1976) supported the same point. Allsop suggested an iterative procedure to solve such a problem, which decomposed the problem into two well-researched subproblems as in Figure 1. The assignment uses link performance functions derived by the signal optimization subproblem. Signal optimization is performed with flow patterns provided through the assignment subproblem. In the literature, this is called the Iterative Optimization and Assignment Procedure, or simply

Iterative Approach. The procedure continues until it converges to a solution, which is called mutually consistent because the flow is at UE and the signal setting is optimal.

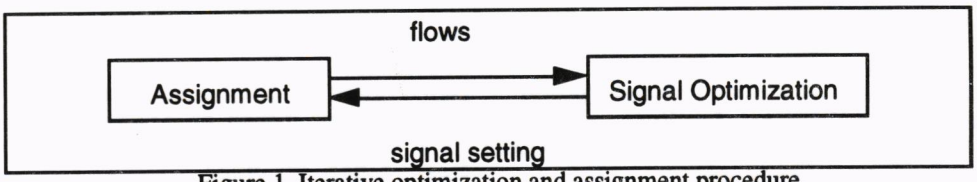

Figure 1 Iterative optimization and assignment procedure.

Allsop's conceptual algorithm was extended by Allsop and Charlesworth (1977) on a six-intersection network. Quite distinct mutually consistent solutions, i.e., different flow and green time patterns, were found but indicated similar total network travel times. Tan et al. (1979) expressed UE flow pattern as a set of constraints and suggested an Augmented Lagrangian Method for solutions. However the method is inappropriate for large networks because of path enumeration.

Dickson (1981) and Smith (1979) noted that the above iterative method is not guaranteed to converge even to a local optimum. Sheffi and Powell (1983) suggested a local search methodology and solved a small network problem with a simplified link performance function. Smith (1979) proposed a new signal control policy P_O with a capacity maximizing property, which is different from conventional delay minimization or Webster's equisaturation policy (1958). Smith and Van Vuren (1993) analyzed the convergence and uniqueness of solutions by the iterative approach for a large collection of control policies. They showed that the link performance functions and the control policies affect the convergence and uniqueness of mutually consistent solutions. Van Vuren and Van Vliet (1992) performed a comprehensive experimental study.

Cantarella et al. (1991) described the behavior of the iterative procedure graphically. Cantarella and Sforza (1995) included an offset optimizer in the iterative procedure. Gartner and Al-Malik (1996) introduced a simultaneous approach for both route choice behavior and optimal signal setting by representing signal control variables as equivalent flow variables at two phase operating intersections. Yang and Yagar (1995) developed a gradient descent algorithm to utilize the sensitivity analysis procedure. In this paper, the following notation is used,

CL	cycle length;
λ_a	green time ratio for intersection link or movement a;
s_a	saturation flow rate for intersection link or movement a;
t_a	travel time on link or movement a;
t_o	free flow travel time of nonintersection link;
x_a	flow on link or movement a.

2. OBJECTIVES, ASSUMPTIONS AND FORMULATION

2.1 Objectives

The main objectives of this study are to compare the iterative search with the local search, simulated annealing and a genetic algorithm for the equilibrium network traffic signal setting problem at different demand levels and in different size networks regarding solution quality and convergence pattern. Sheffi and Powell's local search was improved to be able to solve complex signal schemes in large networks and to adopt comprehensive link performance curves as well as their modified BPR type curve. Two stochastic global searches, simulated

annealing and a genetic algorithm were adopted to accommodate the nonconvexity of the problem and they seek to find (near) global optimal solutions.

2.2 Assumptions

To keep the problem manageable, the following assumptions are made:

- The control policy for signal settings is total travel time minimization:
- Traffic assignments are steady-state with fixed OD. Driver route choice rule is minimum time path selection so that drivers follow deterministic UE. The Frank-Wolfe algorithm (1956) is used to solve the UE flow problem:
- Webster's two term delay function is selected for intersection delay modeling and link cruise time is assumed constant. Webster's curve is incompatible with the Frank-Wolfe algorithm when flow exceeds capacity since the algorithm contains a series of all-or-nothing assignments, which may cause flows on some links to be more than their capacity during the iterations. Thus, link costs must be defined throughout the whole flow region. The linear adaptation to combine deterministic queueing and Webster's curve at some flow level where the two curves show the same slope is chosen to resolve this problem;
- Only networks with isolated intersections are investigated since including the offset effect *analytically* into the theoretical relationships between flow and control parameters has not yet been properly resolved. Green time is the design signal parameter and cycle length is assumed fixed. All phases are protected and phase sequences are exogenous.

2.3 Formulation

The policy to minimize the total travel time induces the following equilibrium network traffic signal setting problem, P1.

$$\text{P1:} \qquad z = \sum_a t_a(x_a, \lambda_a) \cdot x_a \tag{1a}$$

subject to:

$$\lambda^{min} \le \lambda \le \lambda^{max} \tag{1b}$$

$$x \sim UE \tag{1c}$$

$$x \ge 0 \text{ and } \lambda \ge 0 \tag{1d}$$

Since there is a unique feasible equilibrium flow vector x^* for any feasible λ, i.e., x^* is uniquely decided by λ, P1 can be transformed to P2.

$$\text{P2:} \qquad z = \sum_a t_a\left(x_a^*(\lambda), \lambda_a\right) \cdot x_a^*(\lambda) \tag{2a}$$

subject to:

$$\lambda^{min} \le \lambda \le \lambda^{max} \tag{2b}$$

$$x \ge 0 \text{ and } \lambda \ge 0 \tag{2c}$$

For simplicity, x^* will be denoted by x. Two difficulties in solving P1 or P2 have been frequently mentioned (Smith, 1985). First, due to the problem nonconvexity, z may have many local minima. Thus, any gradient based search will find only a local minimum. Second, z requires knowledge of OD pattern, which is not easily developed for large realistic networks. The iterative approach has been the most practical alternative strategy.

3. ALGORITHMS

3.1 Two Global Searches

3.1.1 Simulated Annealing

Kirkpatrick et al. (1983) proposed an algorithm, based on a strong analogy of the annealing process, to some NP-complete combinatorial optimization problems. This analogy is called simulated annealing (SA). Vanderbilt and Louie (1984) extended SA to continuous optimization problems. Energy level, E, in a thermal process is a surrogate for an objective function value in optimization. Possible configurations or states in annealing are comparable to feasible solutions in optimization. If an annealing process properly continues with cooling, a low energy configuration is realized, which is comparable to a desirable optimum in optimization. If annealing is fast such as quenching, the solid cannot reach the low energy configuration. Instead, it may form a locally defected meta-stable configuration, which is comparable to a local optimum.

SA consists of two major elements--the Metropolis algorithm (Metropolis et. al., 1953) and cooling. First, for a given control parameter (temperature), SA repeats the search by generating new candidate solutions and updating the best solution until accepted solutions properly realize the Gibbs distribution. Second, the control parameter is decreased and the search continues with updating the best configuration. Reducing the parameter is called cooling, which plays an important role in SA. The best current solution may not be updated at every step because SA accepts worse candidates stochastically as well as better ones. This stochastic accepting uphill steps provides a chance to escape from a local optimum and makes the best solution eventually close to a global optimum. For global optimization, SA has proven to be a powerful numerical tool and is considered an elegant example of a physical concept imported to other science fields.

Hajek (1988) derived necessary and sufficient conditions for the asymptotic global convergence of SA defined on discrete space using the Markovian property of SA. Vanderbilt and Louie (1984) developed the first SA study to examine continuous optimization problems. Belisle (1992) derived a condition under which SA, defined on a continuous domain, converges in arbitrarily small neighborhoods of global optima regardless of the cooling rate. Because Vanderbilt and Louie's algorithm restricts the search domain during the procedure, Belisle's convergence theorem is not necessarily satisfied. Hence it will find solutions close to a global optimum but not necessarily find the optimum. However, relaxing the restriction may require much more computational effort. Vanderbilt and Louie's test, on the average, detected global optima in 80% of their trials and always found at least local optima. The detailed mechanism of Vanderbilt and Louie's SA to solve P2 is described next.

The current split λ^n and the random step $\Delta\lambda^n$ decide λ^{n+1} as follows:

$$\lambda^{n+1} = \lambda^n + \Delta\lambda^n \tag{3}$$

$$\Delta\lambda^n = \mathbf{Q}\cdot\mathbf{u} \tag{4}$$

$$\mathbf{u} = (u_1, u_2, \ldots, u_I) \tag{5}$$

where, each uniform distribution u_i is independently and identically distributed on the interval $\left[-\sqrt{3},\sqrt{3}\right]$ (i.e., with zero mean and unit variance), I is the number of independent decision variables (here, the total number of phases - total number of junctions), and the matrix Q scales u. The resulting u holds a probability density h(u) which is constant inside a hypercube of volume $\left(2\sqrt{3}\right)^I$ and zero outside. Q and the covariance matrix s of $\Delta\lambda^n$ are related by (6):

$$s = \mathbf{Q}\cdot\mathbf{Q}^T. \tag{6}$$

Q can be obtained from s via an inverting procedure such as the Choleski decomposition. Q and u decide $\Delta\lambda^n$ using (4) and $\Delta\lambda^n$ decides λ^{n+1} using (3). Therefore, the decision of s should be completed before starting the next search set. Vanderbilt and Louie proposed a methodology for deciding the next s by the first and second moments of the walk segment, v and w as

$$s^{(n+1)} = \frac{\chi}{\mu M} w^{(n)} \tag{7}$$

The growth factor χ is chosen >1 so that a free random walk on the (n+1)-th set would cover, on average, $\sqrt{\chi}$ times as much space in each direction as on the n-th set iterations. The whole procedure is shown in Figure 2.

Step 0: Set m=0, n=1, and assume $s^{(n)}$, $\lambda^{(m;n)}$ and temperature $c^{(n)}$
Set μ, χ and a cooling schedule.

Step 1: Inverting procedure to get $Q^{(n)}$ by (6)
(Inner loop begins)

Step 2: Set m=m+1 and generate $u^{(m;n)}$

Step 3: Calculate $\Delta\lambda^{(m;n)} = Q^{(n)} \cdot u^{(m;n)}$

Step 4: Calculate the new point $\lambda^{new} = \lambda^{(m-1;n)} + \Delta\lambda^{(m;n)}$

Step 5: Solve UE with λ^{new}. Calculate $F\left(x^{new}, \lambda^{new}\right)$
And evaluate the new point by the Metropolis criterion
If the new point is accepted, $\lambda^{(m;n)} = \lambda^{new}$
Otherwise, $\lambda^{(m;n)} = \lambda^{(m-1;n)}$

Step 6: Return to Step 2 until m=M
(Inner loop ends)

Step 7: Calculate $v^{(n)}$, $w^{(n)}$ and $s^{(n+1)}$ unless converged

Step 8: Set the new $c^{(n+1)}$ ($c^{(n+1)} \le c^{(n)}$)
Set n=n+1, m=0 and return to Step 1

Figure 2 Adaptation of Vanderbilt and Louie's continuous SA

3.1.2 Genetic algorithm

A genetic algorithm (GA) is a stochastic algorithm based on the principle of evolution and survival of the fittest. Recently GA has received considerable attention regarding its potential as an optimization technique for complex problems. A basic form of GA was described by Goldberg (1989) and Gen and Chen (1997). GA has the following user-specified parameters as shown in Table 1: crossover rate, mutation rate, population size and maximum generation. One more parameter is the length of the encoded decision variable strings, which depends on the required precision.

Table 1 GA Parameters

GA parameters
• Crossover rate = 0.3
• Mutation rate = 0.01
• Substring length for a green split = 7bits[a]
• Population size = 300
• Maximum generations = 100

[a]Two places precision after the decimal

The P2 can be easily evaluated by the above GA format as in Figure 3. Each chromosome represents a signal setting in the network and needs one UE assignment to find flow and the system travel time. When the fitness is calculated in Step 4, scaling is required because the travel time is a negative utility and inversely proportional to the fitness. In this paper, a simple linear scaling is utilized. The following two major advantages of GA can be pointed out regarding this application.

- GA searches the entire domain for the ergodicity of evolution operators, which makes GA very effective at performing global searches;
- With minimal mathematical requirements, GA can handle various kinds of objective functions and constraints defined on discrete, continuous, or mixed search domains.

Step 1:	Generate the initial population of splits in binary bit $b^{(1)}$ in random, where $\left\|b^{(1)}\right\|$ is the population size N_p and b
Step 2:	Decode all elements of $b^{(n)}$ to get $\lambda_i^{(n)}$
Step 3:	Perform UE with decoded green splits $\lambda_i^{(n)}$ in Step 2 Calculate the total travel time of each flow pattern
Step 4:	Calculate the fitness of each flow pattern
Step 5:	Select signal settings for the next generation with the fitness
Step 6:	Perform the crossover operation with p_c
Step 7:	Perform the mutation operation with p_m
Step 8:	If $n=N_{max}$ or convergence obtained, stop; Otherwise, Set counter n = n+1 and go to Step 2

Figure 3 Genetic algorithm

3.2 Local Search

When performing a local search for P2, computing the gradient of z must be considered first. Typical terms of the gradient of z are given by

$$\frac{\partial z(\lambda)}{\partial \lambda_k} = \sum_a \left\{ \left[t_a\left(x_a(\lambda),\lambda_a\right) + x_a(\lambda)\frac{\partial t_a\left(x_a,\lambda_a\right)}{\partial x_a} \right] \frac{\partial x_a(\lambda)}{\partial \lambda_k} \right\} + x_k(\lambda)\cdot\frac{\partial t_k\left(x_k,\lambda_k\right)}{\partial \lambda_k} \quad (8)$$

The term $\frac{\partial x_a}{\partial \lambda_k}$ is the partial derivative of the equilibrium flow on a link with respect to the green split on another. Because it is not possible to derive this term analytically, Equation 4 is very difficult to use directly. Sheffi and Powell suggested an numerical approximate method. In (8), one more noticeable term is $\frac{\partial t_k}{\partial \lambda_k}$. It could affect $\frac{\partial z}{\partial \lambda_k}$ severely and give very different values depending on the λ location due to the linear adaptation of Webster's curve. Here two different methods can be devised, analytical and numerical.

The above approach contains both analytical and numerical terms. Thus it is also natural to use a full version of numerical gradient estimation to avoid bothersome differentiation of some complex delay functions as,

$$\frac{\partial z}{\partial \lambda_k} \cong \frac{z(\ldots,\lambda_k+\Delta,\ldots) - z(\ldots,\lambda_k,\ldots)}{\Delta} \quad (9)$$

Sheffi and Powell suggested, although not tested in their work, a simplified gradient approximation for a large network by assuming that the main gradient portion of stage k comes from stage k itself. Lee (1998) found that the numerical local searches are most stable among other local search variations. Thus this research chose the numerical local search as shown in Figure 4.

Step 0:	Initialization: obtain feasible splits, and set n=0.
Step 1:	Update: calculate travel time for given splits and perform UE assignment.
Step 2:	Gradient calculation: calculate $\frac{\partial z}{\partial \lambda_a}$ by (9)
Step 3:	Direction determination: decide descent direction and maximum step size.
Step 4:	Find an optimum step size and update splits.
Step 5:	Return to Step 1 until stopping criterion is met.

Figure 4 Local Search Algorithm

3.3 Iterative procedure

The iterative optimization and assignment procedure (IOA) is a heuristic approach finding mutually consistent points instead of a local or global optimal. The domain space, (x,λ), constitutes vectors of flow and splits. By fixing either one, the IOA decomposes the problem into two subspaces, (x,λ_{fix}) and (x_{fix},λ), and solves them alternatively. (x,λ_{fix}) relates to the traffic assignment to identify flow pattern assuming control variables fixed, and (x_{fix},λ) signal optimization to optimize control variables assuming flow pattern fixed.

Smith and Van Vuren (1993) introduced the concept "pressure" for the signal optimization. Pressure was defined such that if all pressure of stages in an intersection are balanced, the control objectives are achieved. Thus, depending on the control objectives, the real form of pressure can be different. For the total system time minimization policy, the link pressure of Webster's formula is explicitly derived as follows:

$$-x_a \frac{\partial t_a(x_a,\lambda_a)}{\partial \lambda_a} = \frac{x_a CL(1-\lambda_a)}{\left(1-\frac{x_a}{s_a}\right)} + \frac{x_a s_a}{2\cdot(\lambda_a s_a - x_a)^2} - \frac{x_a}{2\cdot\lambda_a^2 s_a} \qquad (10)$$

Notice that Equation (10) is not a stage pressure but a link pressure. The stage pressure is obtained by summing the pressure of the links which belong to that stage (i.e., receive the right-of-way). By swapping green from less pressured stages to more pressured stages, green splits are adjusted and moved to be balanced.

3.4 Chain Rule for Complex Multiphase Signal and Maintaining Feasibility

The gradient $\frac{\partial z}{\partial \lambda_a}$ is the rate of total travel time change with respect to the green time ratio change of movement a. In signal control, however, green time can be changed stage by stage, rather than movement by movement. Thus, when the signal control is complex, such that one movement receives green during more than one stage, a stage gradient is actually required. By the chain rule, the stage gradient can be calculated with (11).

$$d_t = \sum_{a \in S_t} \frac{\partial z}{\partial \lambda_a} \qquad (11)$$

where d_t is the gradient of stage t and S_t is a set of the movements that receive green during stage t. To save space, the detailed proof is omitted.

Finally, when an intersection has N stages, at least one stage green time ratio must be decided by the other N-1 stages. The former and the latter are called dependent and independent stage, respectively. To maintain the feasibility of SA and GA, independent stage green time will be decided first and then the dependent stage green time will be calculated by the independent stage green times. However, the newly generated green times are not guaranteed to maintain the feasibility. Thus, retrials are required when the feasibility violation is detected.

4. EXPERIMENT

4.1. Experimental Scheme

To compare the IOA, local and global searches, four example networks in Figure 5, 6 and 7 were chosen. Figure 5 contains two simple networks denoted by "VV", and "2x1." Figure 6 is a medium size network denoted by "MED", having 10 zone centroids and 11 signalized intersections under multiphase operation with overlapping movements, 40 total phases, 108 intersection links and 158 total links. Figure 7 is a network located in Austin, Texas, USA, denoted by "AST", consisting of 21 zone centroids and 27 signalized intersections under multiphase operation with overlapping movements, 83 total phases, 267 intersection links and 323 total links.

Network congestion can affect the relative performance of the algorithms. Five different OD demand levels from 1 to 5 are selected so that their network wide volume per capacity ratios show 0.1, 0.3, 0.5, 0.7 and 0.9, respectively. Initial control setting is an important factor of solution quality because of the problem nonconvexity. 32 different initial settings are selected. Therefore the experiment has 640 repetitions for each network and 2560 total cases. When any numerical approximation is involved for gradient calculation, Δ is set to 0.05, which means 3 seconds when the cycle length is 60 seconds. If Δ is too small, the derivative estimate is subject to roundoff noise; whereas if Δ is too large, it no longer measures the local gradient. According to the test, 3 seconds was an appropriate Δ value.

4.2 Experiment Results

Means of total travel times are summarized in Table 2 (Each table cell value is the mean of the 32 different initial settings). The best value of each row (i.e., across the algorithms) are shadowed. The results vary across the networks. The means by the IOA are relatively worse than those by the other algorithms for the VV and 2x1 networks. The phenomenon, however, becomes reversed as the network size increases. The IOA outperforms the other searches in the AST network, specially at high demand levels. Generally, the deviation grows as network size increases or as demand increases. The two global searches did not always find best solutions because of the finite iterations in real implementation.

This relative performance that the IOA is strong at high demand in a large network while local searches are strong at high demand in a small network is illustrated in Figure 8 and can be ascribed to the following:

> When the network is small, there may be very few distinct local solutions, which can be found by local searches. Although the mutually consistent solution is intrinsically suboptimal, for the small network, it is quite similar to the local solutions when demand level is low, and as demand grows, the difference grows. On the other hand, for the large network, there may be many local or quasi-local solutions. Thus, any local search can be easily trapped to worse solutions if the initial solution is not in a good domain neighborhood. Since the iterative approach includes a signal optimization procedure, it finds a good solution showing small total travel time, which may not be mutually consistent until convergence, whether the initial solution is in a good neighborhood or not. Then the search drifts to find a mutually consistent point. When the network is big and demand is high, there may be many mutually consistent points so that it is likely to find one around the signal optimized point.

Two global searches require much more computational times and iterations to stop than the iterative and local searches as shown in Table 3. The local search takes 3 to 5 iterations to converge and the IOA averages 4 to 15 iterations. Between SA and GA, SA is relatively faster than GA. The global searches, however, should be used for the limited cases because of the slow convergence. Although the IOA requires a bit more iterations than the local search, the UE iterations involved in the optimum step size decision of the local searches are

not negligible as shown in Table 3. To reduce this local search computational burden, many search algorithms utilize a streamlined method such as a predetermined step size or a limited maximum number of line searches. This study, however, does not include this. On the other hand, the IOA also involves a line search for optimal green swapping decisions in the control optimization procedure, which can be efficiently performed. Figure 9 shows the objective value improvement by the numerical local search marginally decreases while the IOA shows some hump. This hump is called a wrong way optimization, caused by no consideration of flow change during control optimization.

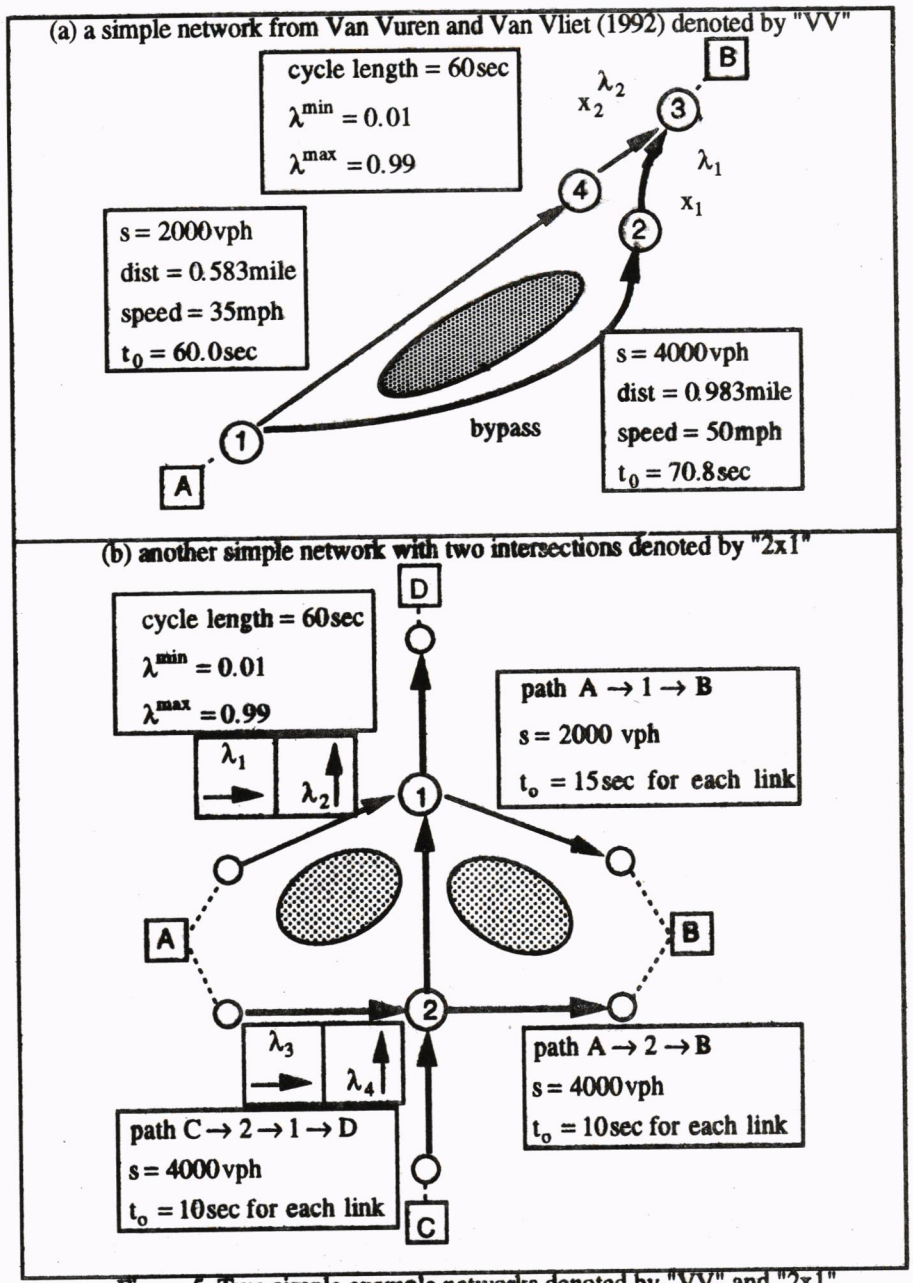

Figure 5 Two simple example networks denoted by "VV" and "2x1"

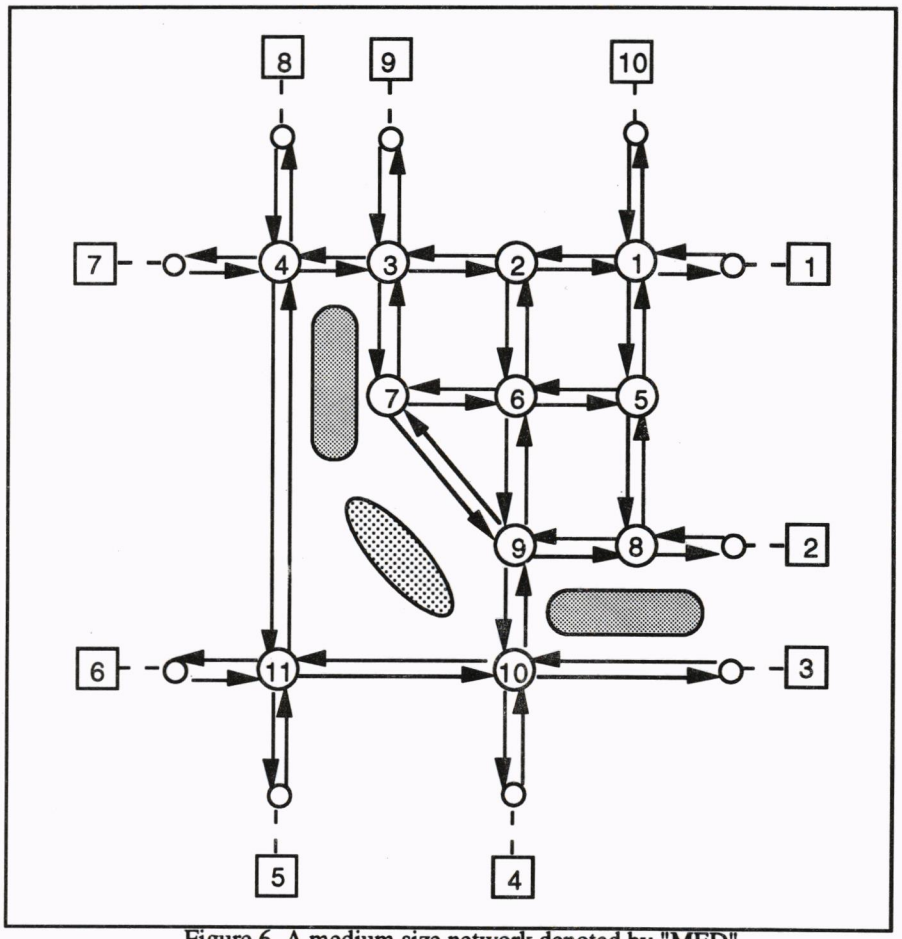

Figure 6 A medium size network denoted by "MED"

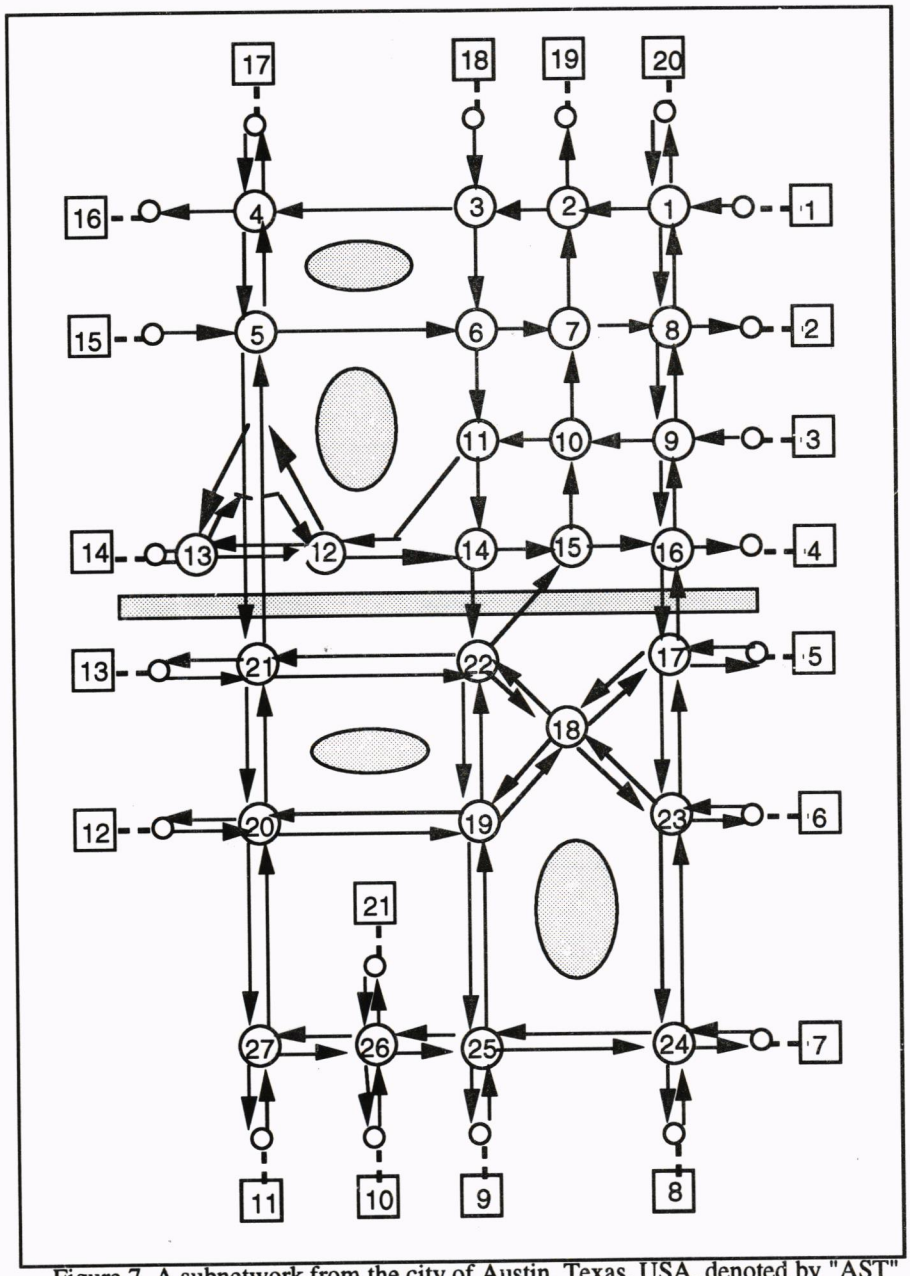

Figure 7 A subnetwork from the city of Austin, Texas, USA, denoted by "AST"

Table 2 Mean of total travel time by the four searches (hr)

N/W	OD	IOA	Local	SA	GA
W	1	6	5	5	5
	2	21	21	19	19
	3	46	39	39	39
	4	84	56	56	56
	5	172	80	75	75
2x1	1	4	4	4	4
	2	15	14	14	14
	3	25	25	25	25
	4	43	45	43	43
	5	83	80	78	78
MED	1	33	36	35	33
	2	131	144	139	132
	3	274	324	304	273
	4	657	788	738	624
	5	1398	1781	1637	1304
AST	1	147	155	152	148
	2	555	595	577	568
	3	1030	1250	1094	1060
	4	1838	2624	1970	1877
	5	3345	5487	3548	3648

best across algorithms

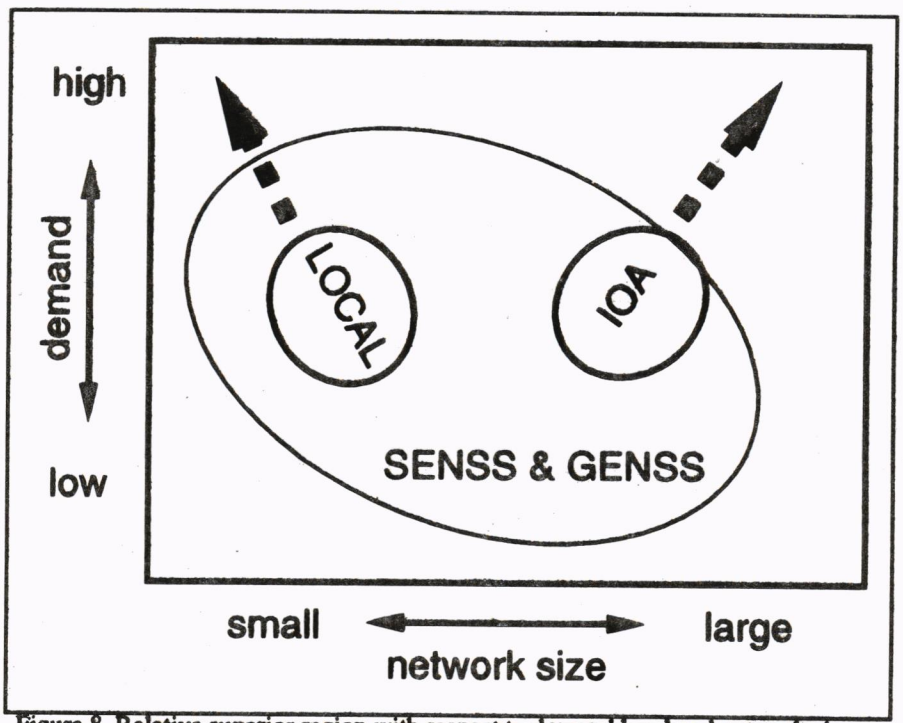

Figure 8 Relative superior region with respect to demand level and network size

Table 3 Computing times, average iterations until convergence and UE assignment repetitions of the local search[a]

Network	IOA[b]	Local[c]	SA[d]	GA[e]
VV	1.5sec, 3.7	1.2sec, 3.1, 93	4.1sec, 108	5.6sec, 147
2x1	1.8sec, 4.1	1.9sec, 5.1, 209	6.1sec, 492	7.5sec, 1932
MED	2.8sec, 8.3	2.2sec, 3.0, 90	35sec, 2044	2.2min, 37671
AST	5.1sec, 14.7	4.1sec, 3.0, 139	14min, 13952	31min, 77240

[a]Digital AlphaStation 600 5/26 with Alpha EV5 processor
[b](average CPU time, average iteration to converge)
[c](average CPU time, average iterations to converge, average UE repetitions)
[d](average CPU time, average iterations to converge)
[e](average CPU time, average iterations to converge)

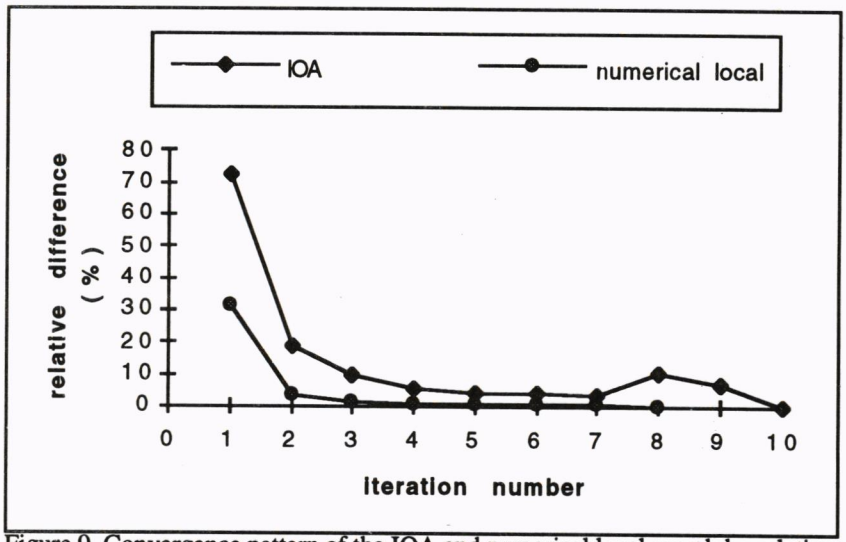

Figure 9 Convergence pattern of the IOA and numerical local search by relative objective value reduction between iterations for the 2x1 network at OD level 3

5. CONCLUSION

The combined control and assignment problem is examined focusing on three different approaches, stochastic global, gradient based and iterative approaches, under a planning (off-line) perspective. The iterative approach was also coded to compare the solution quality.

According to the comprehensive experimental test, when the network was small, the iterative and local searches found good solutions simultaneously, but only at low demand levels. When demand was high, the iterative approach failed to produce good solutions. On the other hand, when the network was big, as demand level became higher, the iterative approach tended to find better solutions. Two stochastic global searches are slow but find stable solutions when demand is low.

REFERENCES

Allsop, R. E. (1974) Some possibilities for using traffic control to influence trip distribution and route choice. **Proceedings of the Sixth International Symposium on Transportation and Traffic Theory** (D. Buckley, ed.), Sydney, 345-373.

Allsop, R. E., and J. A. Charlesworth (1977) Traffic in a single-controlled road network: an example of different signal timings inducing different routings. **Traffic Engineering & Control 18**, 262-264.

Belisle, C. (1992) Convergence theorems for a class of simulated annealing algorithms on R^d. **Journal of Applied Probability 29**, 885-895.

Cantarella, G. E., G. Improta, and A. Sforza (1991) Iterative procedure for equilibrium network traffic signal setting. **Transportation Research 25A No. 5**, 241-249.

Cantarella, G. E., and A. Sforza (1995) Network design models and methods for urban traffic management. In (N. H. Gartner and G. Improta (eds.), **Urban Traffic Networks: Dynamic Flow Modeling and Control**. Springer-Verlag, New York, 123-153.

Dickson, T. J. (1981) A note on traffic assignment and signal timings in a signal -controlled road network. **Transportation Research 15B No. 4**, 267-271.

Frank, M., and P. Wolfe (1956) An algorithm for quadratic programming. **Naval Research Logistics Quarterly 3**, 95-110.

Gartner, N. H. (1976) Area traffic control and network equilibrium. In M. A. Florian (ed.), **Traffic Equilibrium Methods**, Springer-Verlag, 274-297.

Gartner, N. H., and M. Al-Malik (1996) Combined model for signal control and route choice in urban traffic networks. **Transportation Research Record 1554**, TRB, National Research Council, Washington, D.C., 27-34.

Gen, M. and Cheng, P. (1997) **Genetic Algorithms and Engineering Design**. John Wiley, New York.

Goldberg, D. (1989) **Genetic Algorithms in Search, Optimization and Machine Learning**. Addison-Wesley, Reading, MA.

Hajek, B. (1988) Cooling schedules for optimal annealing. **Mathematics of Operations Research 13 No. 2**, 311-329.

Kirkpatrick, S., Gelatt, C. D., Jr. and Vecchi, M. P. (1983) Optimization by simulated annealing. **Science 220 No. 13**, 671-680.

Lee, C (1998) **Combined Traffic Signal Control and Traffic Assignment: Algorithms, Implementation and Numerical Results**. Dissertation, The University of Texas at Austin.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N. and Teller, A. H. (1953) Equation of state calculations by fast computing machines. **The Journal of Chemical Physics 21 No. 6**, 1087-1092.

Sheffi, Y., and W. B. Powell (1983) Optimal signal settings over transportation networks. **Journal of Transportation Engineering 109 No. 6**, 824-839.

Smith, M. J. (1979) Traffic control and route-choice; a simple example. **Transportation Research 13B**, 289-294.

Smith, M. J. (1979) The existence, uniqueness and stability of traffic equilibria. **Transportation Research 13B**, 295-304.

Smith, M. J. (1985) Traffic signals in assignment. **Transportation Research 19B No. 2**, 155-160.

Smith, M. J., and T. Van Vuren (1993) Traffic equilibrium with responsive traffic control. **Transportation Science 27 No. 2**, 118-132.

Tan, H.-N., S. B. Gershwin, and M. Athans (1979). Hybrid Optimization in Urban Traffic Networks, Laboratory for Information and Decision Systems Technical Report DOT-TSC-RSPA-79-7, Massachusetts Institute of Technology, Cambridge.

Van Vuren, T., and D. Van Vliet (1992) **Route Choice and Signal Control: The potential for Integrated Route Guidance**. Athenaeum Press Ltd, Newcastle, Great Britain.

Vanderbilt, D. and Louie, S. G. (1984). A Monte Carlo simulated annealing approach to optimization over continuous variables. **Journal of Computational Physics 56**, 259-271.

Wardrop, J. G. (1952) Some theoretical aspects of road traffic research. **Proceedings Institution of Civil Engineers II (1)**, 325-378.

Webster, F. V. Traffic signal settings (1958) Road Research Technical Paper 39. Her Majesty's Stationary Office, London, England.

Yang, H., and S. Yagar (1995) Traffic assignment and signal control in saturated road networks. **Transportation Research 29A No. 2**, 125-139.