PROCEEDINGS

OF THE 39TH ANNUAL MEETING OF THE

TRANSPORTATION RESEARCH FORUM

Montreal, PQ, Canada
October 16 - 18, 1997

Volume 2

Global, Local and Iterative Searches to the Equilibrium Network Traffic Signal Setting Problem.

Lee, C. and Machemehl, R.

The University of Texas at Austin

Austin, Texas, the USA.

ABSTRACT

A Simulated annealing (SA) algorithm is applied to the equilibrium network traffic signal setting problem seeking globally optimal, in some sense, signal settings and an equilibrium traffic flow pattern simultaneously. The SA method aims at overcoming the nonconvexity of the problem, which can undermine the quality of local search solutions.

A local search algorithm based on Sheffi and Powell (1984) is used to find local solutions. The iterative approach performs assignment and signal optimization sequentially until it converges to mutually consistent points where the flow is at user equilibrium and the signal setting is optimal.

Link performance is described, in two ways, by the BPR formula and Webster's delay functions. The origin-destination matrix is assumed fixed and optimal green time per cycle ratios are decision variables. The above three approaches are tested on a simple network and applied to a real network with 27 signalized intersections in a real southwestern USA city. The solution quality and convergence pattern are discussed.

Introduction

Signal setting parameters consist of several decision variables--green time per cycle ratios (green split), cycle length, phase sequence and offset. Most models and codes optimizes a subset of these decision variables assuming the others fixed. Changing signal settings may stimulate drivers to adjust their route choices. However changing flow may suggest re-setting signals. It is normally assumed drivers follow Wardrop's first principle (Wardrop, 1958), i.e., user equilibrium (UE) flow pattern. The problem of combining signal controls and an equilibrium assignment is called the equilibrium network traffic signal setting problem.

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. Allsop suggested an iterative procedure to solve such a problem and 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 Assignment Control 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.

Allsop's conceptual algorithm was extended by Allsop and Charlesworth (1977), in which the signal optimization subproblem is solved by TRANSYT, the old version of TRANSYT-7F. Allsop and Charlesworth carried out the procedure on a small six-intersection network. Quite distinct mutually consistent solutions, i.e., different flow and green time patterns, were found but indicated virtually equal total travel times in the test network.

Dickson (1981) and Smith (1979a) noted that the above iterative method is not guaranteed to converge even to a local optimum. Sheffi and Powell (1983) suggested a

local search but because of problem complexity, the iterative approach has been mainly studied. Smith (1979b) proposed a new signal control policy P_O(1979b), which is different from conventional delay minimization or Webster's equisaturation policy (Webster, 1958) and Smith and Van Vuren (1993) analyzed the convergence and uniqueness of solutions in the iterative approach. They showed that the link performance functions and the signal optimization policies affect the convergence and uniqueness. Van Vuren and Van Vliet (1992) performed numerical examples.

The equilibrium network traffic signal setting problem is know to be not necessarily convex, given a user equilibrium constraint. Therefore it may have lots of different local solutions. Thus, there is a possibility that local and mutually consistent points are severely different from global solutions. This paper attempts to find global, local and mutually consistent points using a global, a local and an iterative approach, respectively. Simulated annealing is adopted for the global search and other solutions are compared to determine whether local and iterative approaches severely impair solution quality.

In the next section, assumptions and the formulation of the problem are given. The explanation of three algorithms and application to a small network and a larger real network follow.

Assumptions and Formulation

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

1. The flow pattern follows user equilibrium (UE) and the origin-destination matrix (OD) is fixed and known. Thus, assignments are steady state.
2. The measure of network performance is total network travel time. Therefore minimizing total system time is the signal optimization policy.
3. Green split is the only explicitly considered signal decision variable. The main reason for this is that it enables analytical set up of the theoretical relationship between flows

and signals. Cycle length could be included in this way and other parameters can be included if simulation is used for the link performance function.

4. Link travel time consists of link cruise time and intersection delay. Link cruise time is assumed constant, and delay is modeled by the BPR function, extended to allow signal control as in Sheffi and Powell (1983) and Smith and Van Vuren (1993), and Webster's two term formula. For link a, travel time is

$$\text{BPR: } t_a(x_a, g_a) = t_0\left(1+\alpha\left(\frac{x_a}{g_a s_a}\right)^{\beta}\right) \tag{1}$$

$$\text{Webster's: } t_a(x_a, g_a) = t_0 + \frac{1}{2}\cdot\frac{C(1-g_a)^2}{\left(1-\frac{x_a}{s_a}\right)} + \frac{1}{2}\cdot\frac{x_a}{g_a s_a (g_a s_a - x_a)} \tag{2}$$

where t_a is travel time for link a, t_0 is free flow travel time (or cruise time), x_a is flow, s_a is saturation flow, g_a is green split, α and β are parameters (typically, 0.15 and 4.0 respectively), and C is cycle length. The BPR formula was not designed for a signal-setting context but it is used for its simplicity. Webster's formula has more theoretical and experimental background supporting use in an intersection delay context.

5. An intersection is confined with exactly two approaches (links) and two phases. Either approach has the right of way over only one phase. This assumption makes all formulae in this paper simple but does not apply in the Austin sub-network of Network Examples section, where multiphases and multimovements are used. The following relation and chain rule between green splits of phases and links are used to resolve this. For a link green split, g_i, and phase splits, λ_i's, of an intersection,

$$g_i = \sum_{\text{link i belongs to phase j}} \lambda_j \quad . \tag{3}$$

Then,

$$\frac{\partial g_i}{\partial \lambda_j} = \begin{cases} 1, & \text{if link i belongs to phase j} \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

Then for a function, $f(g_1, g_2, g_3, \cdots)$,

$$\frac{\partial f}{\partial \lambda_j} = \frac{\partial f}{\partial g_1} \cdot \frac{\partial g_1}{\partial \lambda_j} + \frac{\partial f}{\partial g_2} \cdot \frac{\partial g_2}{\partial \lambda_j} + \frac{\partial f}{\partial g_3} \cdot \frac{\partial g_3}{\partial \lambda_j} + \cdots$$

By (4)

$$= \sum_{\text{link i belongs to phase j}} \frac{\partial f}{\partial g_i} \tag{5}$$

Let $\{J,L\}$ denote a graph representation of a transportation network where J is a set of J signalized junctions and L is a set of L directed links. L_j is a subset of L, consisting of all links of which green splits are controlled by junction j. L_j's are mutually exclusive and collectively exhaustive subsets. The policy to minimize total travel time endues the flowing equilibrium network traffic signal setting problem, P1.

$$\text{P1:} \qquad \min_{x,g} F(x,g) = \sum_a x_a \cdot t_a(x_a, g_a) \tag{6}$$

$$\text{subject to} \qquad x \sim UE \tag{7}$$

$$g_{min} \le g_a \le g_{max} \quad \text{for each link a} \tag{8}$$

$$\sum_{\text{all phases of j}} g_i = 1 \sum_{\text{all phases of j}} \lambda_i \quad \text{for each } j \in J \tag{9}$$

where g_i is the same as λ_i (green split of phase i) by the assumption 5. Refer to Sheffi (1985) for the explicit representation of UE and the Frank-Wolfe algorithm to solve UE. (8) and (9) describe signal control relationships, however, if (8) and (9) are relaxed, $g_a^* = \infty$ would be an optimal solution which reduces t_a but is unrealistic. P1 is generally nonconvex because (7) may not necessarily be convex. Thus, P1 may have different local solutions, and there is no known equivalent mathematical optimization program that finds the equilibrium flow pattern and optimal splits of P1 simultaneously.

Simulated Annealing

Kirkpatrick, Gelett and Vecchi (1983) proposed an algorithm, based on a strong analogy of the annealing process, to some NP-complete combinatorial optimization problems, all of which are widely considered unsolvable by polynomial algorithms--for example, the traveling salesman problem (TSP). This analogy is called simulated annealing (SA). Independently Cerny (1985) applied the similar concept to TSP and 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 temperature lowering, a low energy configuration is realized, which is comparable to a desirable optimum in optimization. If annealing is fast such as quenching, the solid configuration cannot reach the low energy configuration. Instead, it may form a locally defected meta-stable configuration, which is comparable to a local optimum. Temperature is regarded as a control parameter in SA.

SA consists of two major folds--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 because it corresponds to temperature lowering in annealing. How to reduce the parameter is a key factor of SA and is called the cooling schedule.

The best current solution may not be updated at every step because the Metropolis algorithm accepts worse candidates by the Metropolis criterion. Accepting uphill steps as well as downhill steps provides a chance to escape from a local energy configuration, i.e., a local optimum. This escape allows the algorithm to search for different domain

neighborhoods. Eventually the best solution becomes close to the lowest energy state, i.e., a global optimum. For certain optimization problems, particularly 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 fields of science.

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) was the first SA study to examine continuous optimization problems. The method was adopted by Frietz et al. (1992) to solve a transportation network design problem, which is known to be nonconvex. 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. Vanderbilt and Louie's algorithm restricts the search domain during the procedure based on the previous search and Belisle's theorem is not necessarily satisfied. Hence it will find solutions close to the 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 P1 is described next.

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

$$\mathbf{g}^{n+1} = \mathbf{g}^n + \Delta\mathbf{g}^n \tag{10}$$

$$\Delta\mathbf{g}^n = \mathbf{Q} \bullet \mathbf{u} \tag{11}$$

$$\mathbf{u} = (u_1, u_2, \ldots, u_L) \tag{12}$$

where, each uniform distribution u_i is independently and identically distributed (iid) on the interval $\left[-\sqrt{3}, \sqrt{3}\right]$ (i.e., with zero mean and unit variance), L is the number of independent decision variables (Here, the total number of phases - total number of junctions), and the matrix $\mathbf{Q}$ scales u. The resulting u holds a probability density h(u)

which is constant inside a hypercube of volume $\left(2\sqrt{3}\right)^L$ and zero outside. $\mathbf{Q}$ and the covariance matrix $\mathbf{s} = \left(s_{ij}\right)$ of $\Delta\mathbf{g}^n$ are related as follows:

$$\begin{aligned} s_{ij} &= \int d^L u\, h(\mathbf{u})\, \Delta x_i\, \Delta x_j \qquad \text{(by definition)} \\ &= \frac{1}{\left(2\sqrt{3}\right)^L} \int_{-\sqrt{3}}^{\sqrt{3}} du_1 \int_{-\sqrt{3}}^{\sqrt{3}} du_2 \cdots \int_{-\sqrt{3}}^{\sqrt{3}} du_L \left(\sum_k Q_{ik} u_k\right) \cdot \left(\sum_m Q_{jm} u_m\right) \\ &= \sum_k Q_{ik} Q_{jk} \qquad \text{(by } u_i\text{'s are iid)} \end{aligned} \tag{13}$$

Thus,

$$\mathbf{s} = \mathbf{Q} \cdot \mathbf{Q}^T. \tag{14}$$

Using $\mathbf{Q}$, $\mathbf{s}$ can be calculated in (14). On the other hand, $\mathbf{Q}$ can be obtained from $\mathbf{s}$ via inverting procedure such as the Choleski decomposition for (14). $\mathbf{Q}$ and $\mathbf{u}$ decide $\Delta\mathbf{g}^n$ with (11). Finally this $\Delta\mathbf{g}^n$ decides $\mathbf{g}^n$ with (10). The method continues a set of searches at a control parameter c with $\mathbf{s}$ and then reduces temperature for the next set of searches. Therefore, the decision of $\mathbf{s}$ should be completed before starting the next set of searches, which means $\mathbf{s}$ cannot be directly calculated from (13). Vanderbilt and Louie proposed a methodology for deciding the next $\mathbf{s}$.

First assume that SA carries out M steps (inner iterations, and each of these steps requires one UE assignment) at each c until a certain stopping rule is met. This is similar to the random walk in stochastic processes. After the n-th set of M steps, the outer iteration proceeds. The first and second moments of the walk segment, $\mathbf{v} = \left(v_1, v_2, \ldots, v_L\right)$ and $\mathbf{w} = \left(w_{ij}\right)$ are calculated:

$$v_i^{(n)} = \frac{1}{M} \sum_{m=1}^{M} g_i^{(m;n)} \tag{15}$$

$$w_{ij}^{(n)} = \frac{1}{M} \sum_{m=1}^{M} \left[g_i^{(m;n)} - v_i^{(n)}\right] \cdot \left[g_i^{(m;n)} - v_j^{(n)}\right] \tag{16}$$

where $g_i^{(m;n)}$ is the value of g_i on the m-th step of the n-th set. At sufficiently high c, no steps will be rejected (i.e., free random walk) and, on average, the following would be expected:

$$\left\langle \mathbf{w}_{free}^{(n)} \right\rangle = \mu M \mathbf{s}^{(n)} \tag{17}$$

where the brackets indicate an average over the LM random variables $g_i^{(m;n)}$ for the walks. Vanderbilt and Louie gave $\mu = \frac{1}{6}$ for the arithmetic average. They suggested the following next iteration prescription for s:

$$\mathbf{s}^{(n+1)} = \frac{\chi}{\mu M} \mathbf{w}^{(n)} \tag{18}$$

so that

$$\left\langle \mathbf{w}_{free}^{(n+1)} \right\rangle = \chi \mathbf{w}^{(n)} \tag{19}$$

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. Now $\mathbf{s}^{(n+1)}$ is invertible by (11) and $\mathbf{Q}^{(n+1)}$ is ready to use for the (n+1)-th set of searches. Vanderbilt and Louie used a μ based on a geometric average since the walk size after many free sets is really a product rather than a sum of the growth factors. Through one dimensional numerical calculation 0.11 is suggested for μ. The whole procedure is shown in Figure 2.

Vanderbilt and Louie suggested the stopping criterion:

$$\frac{\bar{F}^{(n)} - F_{min}^{(n)}}{\bar{F}^{(n)}} < \varepsilon \tag{20}$$

where $\bar{F}^{(n)}$ is the average of F for the M steps of the n-th set, and $F_{min}^{(n)}$ is the best solution during the M steps of the n-th set. For typical applications, $\varepsilon = 10^{-3}$ is suggested. M should be large enough to guarantee reasonable statistics. At its minimum, M>L is

required to avoid a singular w matrix and therefore a singular s, which will cause failure of the inverting procedure in (14). Vanderbilt and Louie suggested $M \approx 15L$ for $L \leq 8$ and larger values of M for large-dimensional systems.

Local Search

The objective of P1, called F, may possess noncontinuous (piecewise continuous with finite jumps) derivatives at a finite number of points because small changes in the green time split may cause a change in the number of paths used between some OD pair. Therefore the gradient required for the local search must be interpreted as a finite difference over the point of derivative discontinuity (Sheffi and Powell, 1983). Typical terms of the gradient of F with respect to g_a can be given by

$$\frac{\partial F}{\partial g_a} = x_a \frac{\partial t_a(x_a, g_a)}{\partial g_a} + \sum_b \frac{\partial x_b}{\partial g_a}\left[t_b(x_b, g_b) + x_b \frac{\partial t_b(x_b, g_b)}{\partial g_b}\right]. \qquad (21)$$

The term $\frac{\partial x_b}{\partial g_a}$ is the partial derivative of the equilibrium flow on a link with respect to the green split on another. It is impossible to derive the term analytically. Thus, (21) is very difficult to use directly. This paper introduced an approximation for the gradient as follows,

$$\frac{\partial F}{\partial g_i} \approx \frac{F(\ldots, g_i + \Delta g_i, \ldots) - F(\ldots, g_i, \ldots)}{\Delta g_i}. \qquad (22)$$

(22) is combined with the gradient projection method by Sheffi and Powell (1983) to keep feasibility constraints (8) and (9), which can be accomplished by computing the average of the gradient terms associated with L_j and subtracting this average form all these gradient terms. Figure 3 describes the local search procedure. For each iteration, the number of UE assignments required equals the number of links (L) and the additional iterations in the step size decision of STEP 5 .

Iterative Approach

As explained in the Introduction, the iterative approach is a heuristic approach finding mutually consistent point instead of a local or global optimal. Sheffi and Powell (1983) called this a naive approach. The domain space of P1, $(\mathbf{x},\mathbf{g})$, constitutes vectors of flow and splits. By fixing either one, the iterative approach decomposes the problem into two subspaces, $(\mathbf{x},\mathbf{g}_{fix})$ and $(\mathbf{x}_{fix},\mathbf{g})$, and solves them alternatively. The first domain relates to the traffic assignment and the second signal optimization.

Smith and co-workers have studied this approach (Smith, 1979a; Smith, 1979b; Smith, 1981a; Smith, 1981b; Smith, 1981c; Smith, 1982; Smith et. al., 1987; Smith and Van Vuren, 1993). In order to find the mutually consistent points, a new concept, pressure, is introduced, which is perceived by the signal setter. For total system minimization policy, the pressure will be $-x_a \frac{\partial t_a(x_a,g_a)}{\partial g_a}$ (Smith and Van Vuren, 1993). In a UE context, users are assumed to seek a less costly route, which results in no flow on more costly routes for every OD pair . Similarly, if the policy is satisfied, a less pressurized phase receives no green time for every intersection.

For the total system time minimization policy, the link pressure of the BPR and Webster's are derived as (Smith and Van Vuren, 1993),

$$\text{BPR: } -x_a \frac{\partial t_a(x_a,g_a)}{\partial g_a} = -t_a \alpha\beta \left(\frac{x_a}{g_a s_a}\right)^{\beta} \frac{x_a}{g_a} \tag{23}$$

$$\text{Webster's: } -x_a \frac{\partial t_a(x_a,g_a)}{\partial g_a} = \frac{x_a C(1-g_a)}{\left(1-\frac{x_a}{s_a}\right)} + \frac{1}{2}\cdot\frac{x_a s_a}{(g_a s_a - x_a)^2} - \frac{1}{2}\cdot\frac{x_a}{g_a^2 s_a} \tag{24}$$

The phase pressure is obtained by summing the pressure of the links which belong to the phase (i.e., receive the right of way). By swapping green from less pressured phases to more pressured phases, green splits are adjusted. Since $(x_a,g_a)=(0,0)$ has no

useful meaning, it is excluded. According to the discussion by Smith and Van Vuren (1993), for the BPR function, the solution may be unique, but for Webster's formula the solution may not be unique.

For the next experiment, a streamlined form of the iterative approach proposed by Smith and Van Vuren (1993) is applied. It performs a single Frank-Wolfe iteration between assignments with new green splits so that the assignment is expedited. The proof of convergence is given in Smith and Van Vuren (1993).

Network Examples

The three algorithms were tested first on a small network. This network is shown in Figure 4 and included four zones each of which is an origin and a destination. OD pairs, 1 to 2 and 4 to 3 have two alternative routes while OD pairs, 2 to 1 and 3 to 4 have only one route. There are nine intersections, each operated with two signal phases. Another experiment using a larger real network from Austin, Texas can be seen in Figure 5. It includes 27 signalized intersections featuring multiphase operation. For each network, the two link cost functions, BPR and Webster's, were applied with three different initial signal settings.

Table 1 contains the parameter information for SA. Initial temperature is set very high and lowered continually based on iterations at the previous temperature. If the best solution at a temperature is significantly improved, then the next temperature should be lowered but close to the current one in order to intensively search the current neighborhood. On the other hand, if the best solution is not much improved, then the next temperature should be lowered enough not to waste time in the current neighborhood. This concept is reflected in an exponential type cooling schedule.

The following assumptions are further made:

- Cycle time=100 sec;
- No lost time;

• Minimum green time=1 sec;

• Maximum green time=99 sec.

Table 2 and 3 contain the experimental results. Case 1 starts with equal green signal time over all intersections (e.g., if an intersection is under three phase operation, then green fraction for each phase would be one third.). This starting solution might be a local optimum of the simple network because of the network symmetry. Case 2 begins with random splits for each intersection generated by a random number generator, so that intersections may initially have very different green splits. Case 3 initially has unequal splits within intersections but the same splits across intersections. In the case of the small network example, the three algorithms found very similar solutions for the BPR link cost as in Table 2 and Figure 6 and the SA did not always produce the best solutions but the differences are minimal. However Webster's formula and the iterative approach with Case 2 initial settings produced a relatively worse solution. In terms of iterations, the local and iterative algorithms are much faster than SA, which is natural because SA tries to intensively search the whole domain by probabilistically accepting worse points (climbing-up).

Table 3 and Figure 7 show the results of the larger Austin sub-network. For the BPR function, all results range from 800 to 1000 system time hours. For Webster's function and the Case 2 starting condition, both the local search and the iterative approach show poor solutions.

In general, the local and iterative approaches find good solutions much faster than SA, however, they tend to get stalled producing worse solutions than SA in some cases. The convergence patterns depicted in Figure 8 and Figure 9 are interesting. Local and iterative algorithms show significant objective function reductions in the initial few iterations regardless of the number of iterations required to converge. SA, which is intensive in computational effort, fluctuates at initial high temperature iterations and becomes stable as temperature decreases.

More comprehensive experiments with different networks and OD levels are required to obtain general results. Particularly, OD levels can affect solutions because if traffic demand is very low or high, green signal proportions may not easily change the minimum path to a specific OD pair. On the other hand if traffic demand is medium, then the minimum path can be sensitive to green signal proportions. In this experiments, traffic demands are moderate, producing volume to capacity ratios, weighted by link flow, between 0.35 to 0.55.

Conclusions

Simulated annealing, local search, and iterative algorithms have been investigated and applied to solve the equilibrium network signal setting problem. With limited experiments, SA is found to minimize the risk of stalling the solution process at poor solution domain at the cost of extensive computation. The convergence patterns of the local and iterative approaches indicate limited iterations may be enough to obtain a good solution, however both exhibit tendencies to produce poor solutions under certain cases. More comprehensive experiments, however, are required.

Reference

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

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

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

Cerny, V. (1985). Thermodynamical approach to the traveling salesman problem: an efficient simulation algorithm. *Journal of Optimization Theory and Applications* , Vol. 45, pp. 41-51.

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

Friesz, T. L., Cho, H-J., Mehta, N. J., Tobin, R. L. and Anandalingam, G. (1992). A simulated annealing approach to the network design problem with variational inequality constraints. *Transportation Science,* Vol. 26, No. 1, pp. 18-26.

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

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

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,* Vol. 21, No. 6, pp. 1087-1092.

Sheffi, Y. (1985). *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*. Prentice-Hall, Englewood Cliffs, New Jersey.

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

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

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

Smith, M. J. (1980). A local traffic control policy which automatically maximizes the overall travel capacity of an urban road network. *Traffic Engineering & Control*, Vol. 21, pp. 298-302.

Smith, M. J. (1981a). The existence of an equilibrium solution to the traffic assignment problem when there are junction interaction. *Transportation Research*, Vol. 15B, No. 6, pp. 443-451.

Smith, M. J. (1981b). Properties of a traffic control policy which ensure the existence of a traffic equilibrium consistent with the policy. *Transportation Research*, Vol. 15B, No. 6, pp. 453-462.

Smith, M. J. (1981c). A theoretical study of traffic assignment and traffic control. V. F. Hurdle, E. Hauer and G. N. Steuart (ed.), *Proceedings of the Eighth International Symposium on Transportation and Traffic Theory*, pp. 490-526..

Smith, M. J. (1982). Junction interactions and monotonicity in traffic assignment. *Transportation Research*, Vol. 16B, No. 1, pp. 1-3.

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

Smith, M. J., Van Vuren, T., Heydecker, B. G. and Van Vliet, D. (1987). The interactions between signal control policies and route choice. N. H. Gartner and N. H. M. Wilson (ed.), *Proceedings of the Tenth International Symposium on Transportation and Traffic Theory*, pp. 319-338.

Van Vuren, T. and Van Vliet, D. (1992). *Route Choice and Signal Control: The Potential for Integrated Route Guidance*. Ashgate Publishing Company, Vermont.

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

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

Webster, F. V. (1958). Traffic signal settings. *Department of Scientific and Industrial Research, Road Research Technical Paper No. 39*. London, 1958.

Table 1 Simulated Annealing Parameters

Parameter	Value
• Growth factor	• 2.0
• Initial temperature	• 9999900000
• Cooling schedule	• Exponential adaptive cooling based on the objective value reduction at the previous temperature.
• Iterations at each temperature (i.e., inner iterations)	• 300 (the simple network) • 3000 (Austin sub-network)

Table 2 Comparisons of the Three Algorithms, Giving Total System Travel Time and Iteration Numbers for the Simple Network.

Link Cost Function	Initial Split[a]	SA[b]	Local[c]	Iterative[d]
BPR	Case 1	113.6/31	113.6/2	113.6/1
	Case 2	114.0/32	113.6/14	113.6/9
	Case 3	114.5/34	113.6/5	113.6/2
Webster	Case 1	225.2/33	229.5/2	229.5/1
	Case 2	226.5/35	222.3/21	290.4/8
	Case 3	225.1/33	229.5/5	229.5/2

[a]Case1-equal green splits, Case2-random green splits and different across intersections, Case3-unequal green splits but same ones across intersections.
[b]Function values (in hr)/outer iterations, (inner iterations=300).
[c]Function values (in hr)/iterations.
[d]Function values (in hr)/iterations.

Table 3 Comparisons of the Three Algorithms, Giving Total System Travel Time and Iteration Numbers for the Austin Sub-Network.

Link Cost Function	Initial Split[a]	SA[b]	Local[c]	Iterative[d]
BPR	Case 1	908/36	846/16	839/253
	Case 2	861/36	881/8	839/242
	Case 3	973/36	846/40	842/246
Webster	Case 1	2538/30	2929/193	2475/4385
	Case 2	2560/26	37396/13	7253/143
	Case 3	3735/24	2732/30	2474/4359

[a]Same as in Table 2.
[b]Function values (in hr)/outer iterations, (inner iterations=3000).
[c]Function values (in hr)/iterations.
[d]Function values (in hr)/iterations.

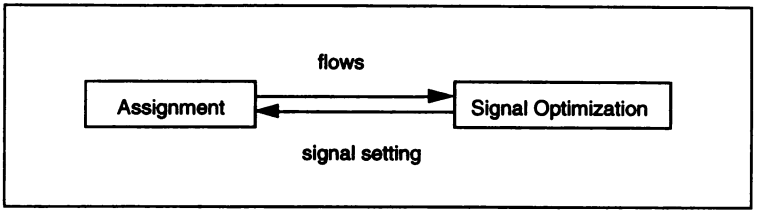

Figure 1 Iterative Assignment Control Procedure

STEP 0: Set m=0, n=1, and assume $\mathbf{s}^{(n)}$, $\mathbf{g}^{(m;n)}$ and the initial temperature $c^{(n)}$.
Set a constant μ, a growth factor χ, and a cooling schedule.
(Outer) STEP 1: Inverting procedure to get $\mathbf{Q}^{(n)}$ by (14).
↓ (Inner) STEP 2: Set m=m+1 and generate $\mathbf{u}^{(m;n)}$.
↓ STEP 3: Calculate $\Delta\mathbf{g}^{(m;n)} = \mathbf{Q}^{(n)} \bullet \mathbf{u}^{(m;n)}$.
STEP 4: Calculate the new point $\mathbf{g}^{(new)} = \mathbf{g}^{(m-1;n)} + \Delta\mathbf{g}^{(m;n)}$.
STEP 5: Solve UE with $\mathbf{g}^{(new)}$. Calculate $F\left(\mathbf{x}^{(new)}, \mathbf{t}\left(\mathbf{x}, \mathbf{g}^{(new)}\right)\right)$.
And evaluate the new point by the Metropolis criterion.
If the new point is accepted, $\mathbf{g}^{(m;n)} = \mathbf{g}^{(new)}$.
↑ Otherwise, $\mathbf{g}^{(m;n)} = \mathbf{g}^{(m-1;n)}$.
(Inner) STEP 6: Return to STEP 2 until m=M.
↑ STEP 7: Calculate $\mathbf{v}^{(n)}$, $\mathbf{w}^{(n)}$ and $\mathbf{s}^{(n+1)}$ unless converged.
(Outer) STEP 8: Set n=n+1, m=0, and reduce temperature $c^{(n)} \le c^{(n-1)}$.
And return to STEP 1.

Figure 2 Vanderbilt and Louie's Continuous Simulated Annealing

STEP 0: Initialization--obtain some feasible splits, $\{g_i^0\}$. Set iteration counter, n=0.

STEP 1: Update-- calculate travel time, $\{t_i^n\}$, given $\{g_i^n\}$ and perform assignment obtaining user equilibrium flow pattern, $\mathbf{x}$.

STEP 2: Gradient Calculation--For each phase, perform (1) to (3).

(1) change g_i^n by Δg_i^n and update t_i^n.

(2) find a new equilibrium flow, $\mathbf{y}$

(3) calculate $\frac{\partial F}{\partial g_i} \approx \frac{F(\mathbf{y}, g_i+\Delta g_i, \ldots) - F(\mathbf{x}, g_i, \ldots)}{\Delta g_i}$.

STEP 3: Decent Direction Determination.

(1) for each junction j, compute $A_j = \frac{\sum_{i \in L_j} \frac{\partial F}{\partial g_i}}{|L_j|}$, where L_j is the link set ending junction j and $|L_j|$ is the carnality of the set L_j.

(2) for each phase, $d_i^n = -\frac{\partial F}{\partial g_i} + A_j$

STEP 4: Determination of Maximum Step Size

(1) for each phase, if $d_i^n > 0$, set $\alpha_i^n = \frac{g_i^{max} - g_i^n}{d_i^n}$, and if $d_i^n < 0$, set $\alpha_i^n = \frac{g_i^n - g_i^{min}}{d_i^n}$.

(2) set $\alpha^{max} = \min_i \{\alpha_i^n\}$.

STEP 5: Step Size Determination and Split Updating.

(1) find α_n^* that solve $\min_\alpha F(\mathbf{g}^n + \alpha_n \cdot \mathbf{d}^n)$ subject to $0 \le \alpha \le \alpha^{max}$.

(2) set $g_i^{n+1} = g_i^n + \alpha_n^* \cdot d_i^n$.

STEP 6: Stopping Test--If $\max_i \{|g_i^{n+1} - g_i^n|\} \le \varepsilon$ then stop. Otherwise set n=n+1 and go to STEP 1.

Figure 3 Local Search

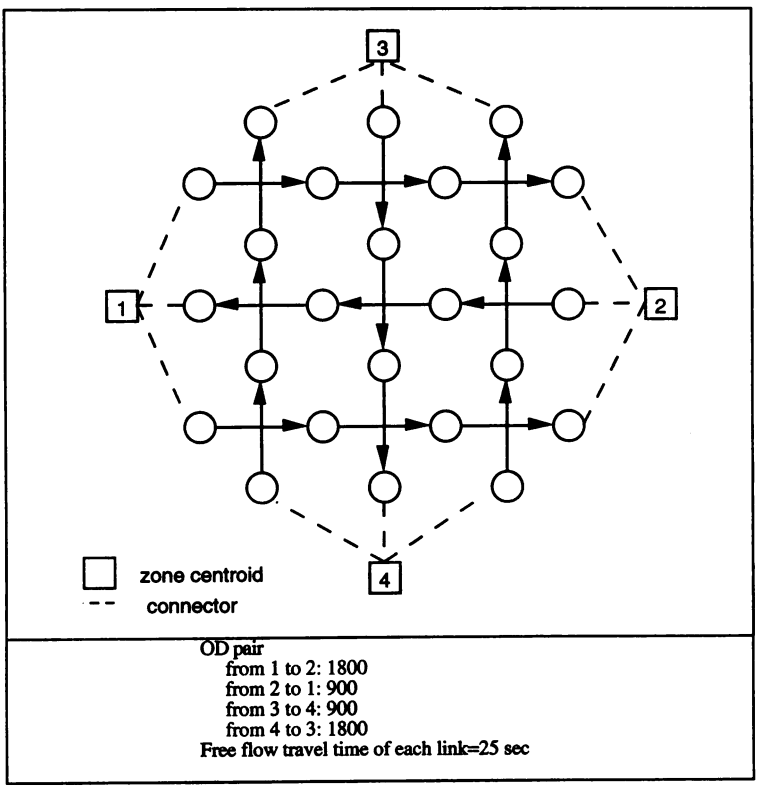

Figure 4 A Simple Network.

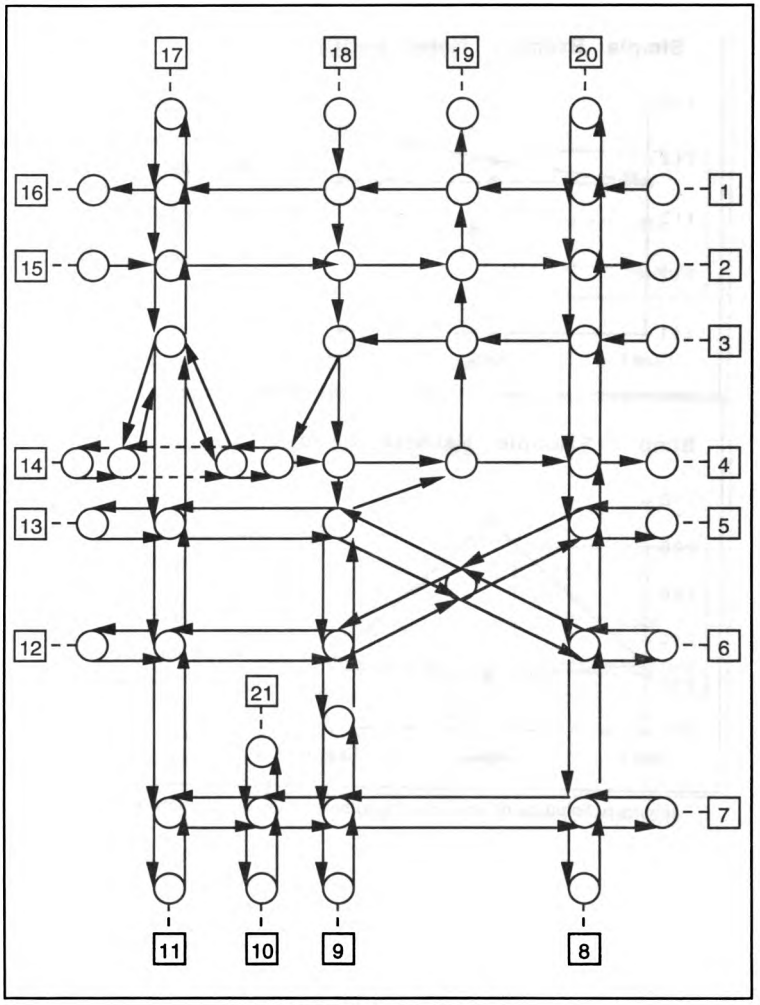

Figure 5 Austin Sub-Network.

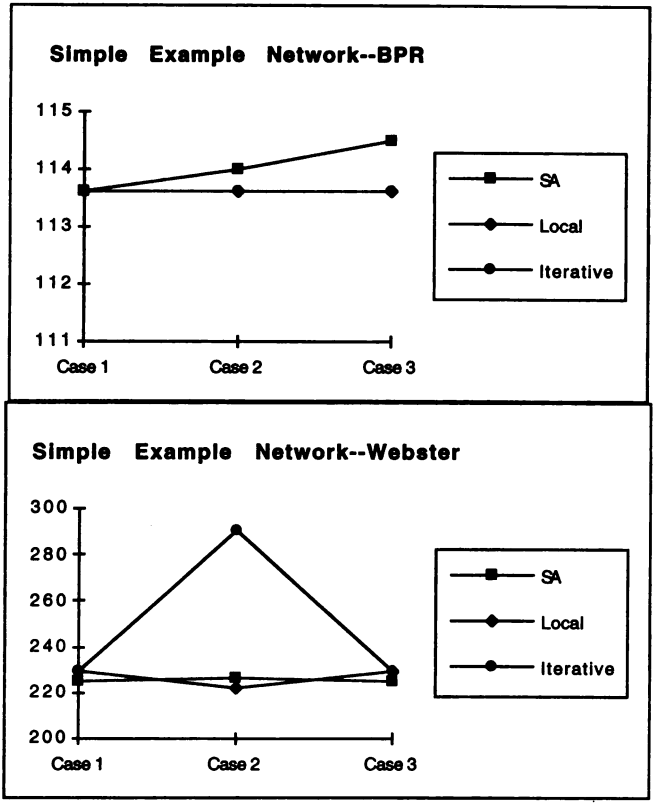

Figure 6 Solutions of the Three Algorithms for the Simple Network

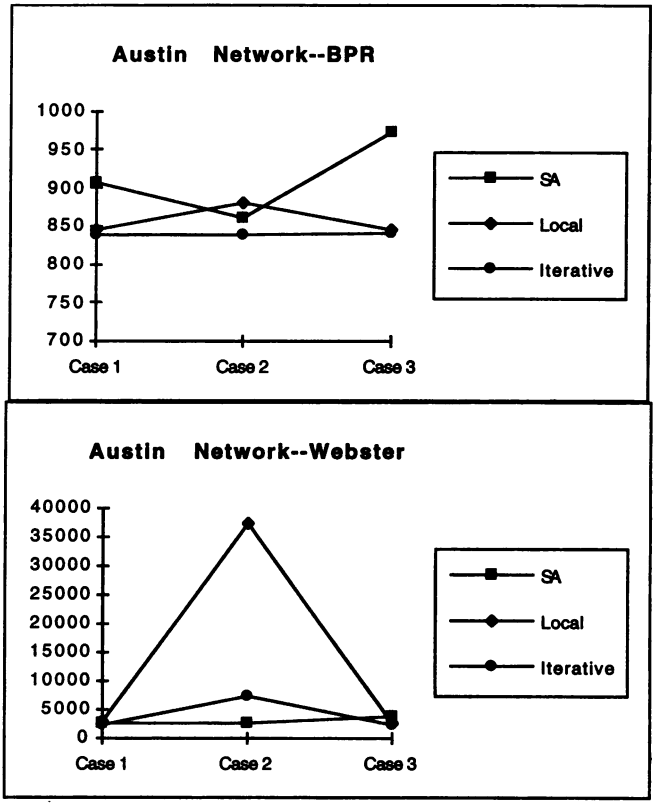

Figure 7 Solutions of the Three Algorithms for the Austin Sub-Network

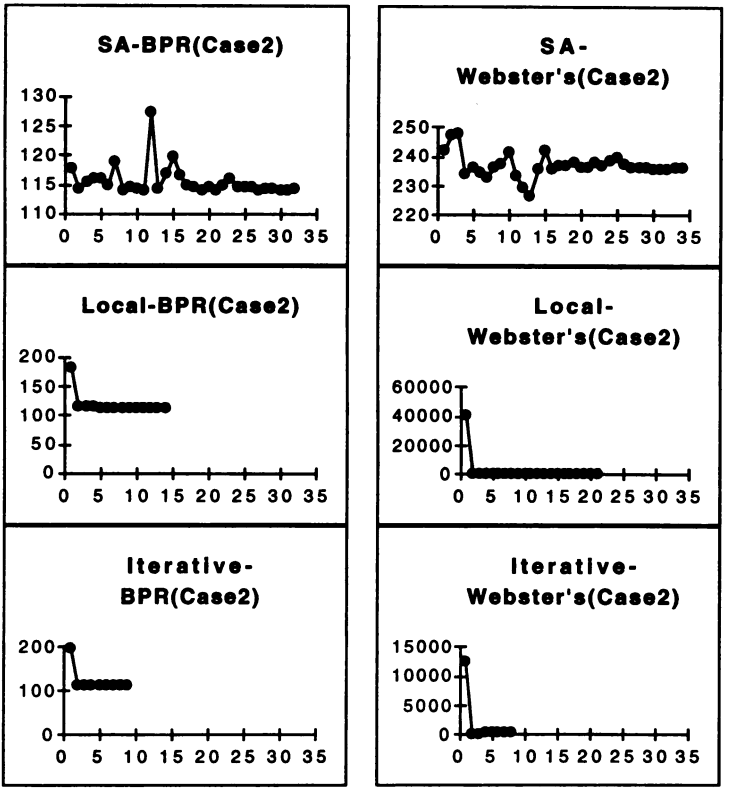

Figure 8 Convergence Patterns of the Three Algorithms for the Simple Example Network (SA points are the best solutions within each outer iteration, i.e., a set of inner iterations at a temperature. Local points are the solutions after each UE assignment. Iterative points are the solutions after each streamlined UE assignment.)

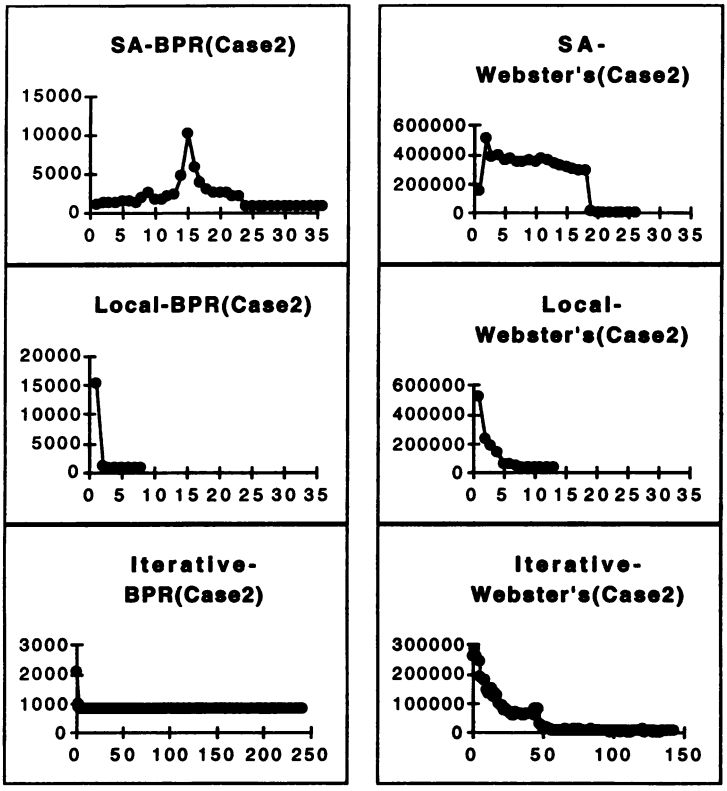

Figure 9 Convergence Patterns of the Three Algorithms for Austin Sub-Network (SA points are the best solutions within each outer iteration, i.e., a set of inner iterations at a temperature. Local points are the solutions after each UE assignment. Iterative points are the solutions after each streamlined UE assignment.)