A Direct Method for Security–Constrained Unit Commitment

John J. Shaw

ALPHATECH, Inc., Burlington, MA 01803

ABSTRACT

Secure operation is an enduring concern to electric utilities. Several factors oblige utilities to operate their systems—particularly transmission systems—at or close to their capacities. Few methods for scheduling electric generating units, however, take security into account. Utilities, therefore, may inadvertently risk overloading critical transmission lines if units are not rescheduled. Without advance planning a utility may incur significant costs if it must reschedule generation retroactively.

This paper presents a practical method to account for security constraints when scheduling electric generating units. Our results indicate that this method obtains lower–cost solutions faster than if the security constraints are considered retroactively.

1. INTRODUCTION

1.1 MOTIVATION

We present a new algorithm for scheduling electric generating units over a stipulated period of time to meet forecasted demand while satisfying system security and transmission line flow constraints. We refer to this problem as *security–constrained unit commitment* (SCUC). Electric utilities establish security requirements to insure the system will continue to operate following a significant disturbance, such as the loss of a large generating unit. The response of a utility following such a contingency is constrained by the generating capacity readily available to serve the electric load and by the capacity of the transmission system to link generators to loads.

Unit commitment is frequently solved without accounting for security and transmission constraints. If the resulting schedule ignores these requirements, the utility may inadvertently overload critical transmission lines. Without advance planning the utility may incur significant costs if it must redistribute generation.

A better approach in our view is to account for the security requirements directly in the scheduling algorithm. This strategy *plans* for a cost–effective way to distribute generation throughout the electric grid so as to meet anticipated security requirements.

1.2 OVERVIEW OF THE ALGORITHM

The SCUC problem has two broad types of constraints: i) *unit–level* constraints that affect units individually, and ii) *system–level* constraints that affect units as a group. Our SCUC algorithm replaces the system–level constraints with penalty functions that appear directly in the objective function. This transformation creates a *dual optimization* problem. The dual optimization problem adjusts these penalty functions to obtain unit schedules that better satisfy the system–level constraints.

The solution to the dual programming problem provides two values pieces of information. First, it provides a *lower bound* on the cost of the best possible unit schedule. This lower bound lets us assess whether a candidate feasible solution is "close" to optimal. Second, the dual programming solution indicates a "neighborhood" of scheduling solutions where we can expect to find a best feasible solution.

Our algorithm develops a feasible SCUC solution in two stages. The algorithm first solves a dual programming problem in the *Dual Optimization* stage. This solution, as we just noted, locates scheduling *options* that are likely to lead to an optimal feasible solution. The algorithm then selects a number of these options and evaluates the cost of each in the *Primal Solution* stage. The algorithm assesses the hour–by–hour cost of each option by solving a series of security–constrained economic dispatch (SCED) problems. Upon completion, the algorithm selects the option providing the lowest cost encountered. We summarize in Fig. 1 the main computation steps performed by our algorithm.

Our innovation is to account for the security constraints in *both* the Dual Optimization *and* the Primal Solution stages. We call this a *Direct* method to indicate that the security constraints are incorporated into the optimization algorithm at the outset. We contrast this innovation with other methods that omit the security constraints from the optimization stage, and consider them only to construct a feasible unit commitment schedule. We refer to these methods collectively as *Indirect*.

94 SM 591-8 PWRS A paper recommended and approved by the IEEE Power System Engineering Committee of the IEEE Power Engineering Society for presentation at the IEEE/PES 1994 Summer Meeting, San Francisco, CA, July 24 - 28, 1994. Manuscript submitted January 4, 1994; made available for printing May 3, 1994.

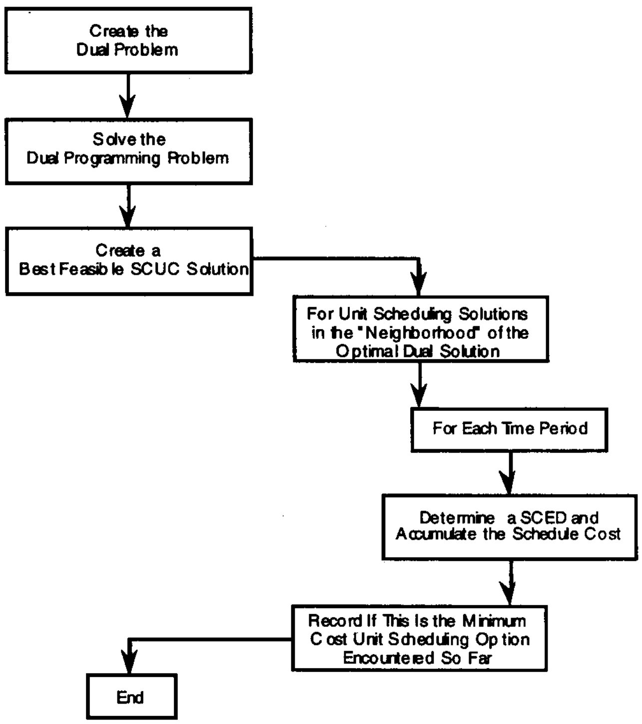

Figure 1. Overview of the SCUC Algorithm

1.3 RELATION TO PRIOR RESEARCH

Until recently few researchers included security requirements, Eq. 4, into the unit commitment problem. Those researchers who have included security into the commitment problem appear to have used *Indirect* methods. Indirect methods ignore the security requirements during the dual optimization, and introduce them retroactively when attempting to obtain feasible solutions. Indirect methods, however, can lead to severely suboptimal results. The problem arises when the electric utility is best off running marginally economic units to satisfy the security requirements. With Indirect methods, the dual solution may turn off several of these units unnecessarily: the utility will need to either turn those units on again or turn on other units that are less economic still.

Some researchers have reported on Direct methods. Dodu et al. [7] accounted for transmission capacity constraints in a greatly simplified version of the unit commitment problem that did not include startup/shutdown costs and minimum uptime/downtime constraints. Ruzic and Rajokovic [8] report on a Direct method that did include the startup/shutdown costs and minimum uptime/downtime constraints. Their computational results, however, were limited to two transmission lines. In addition, the computational time of their algorithm increased sharply (by ~40%) simply when those two lines were added. Zhuang [5] also reports on a Direct method where power flows on transmission lines can be routed to try to avoid overloads. Recently, Batut and Renaud [2] have presented an algorithm that uses an Augmented Lagrangian to decompose the scheduling and security–constrained dispatch problems into to subproblems. They address the case where transmission constraints are applied in one period, but do not present computational results.

1.4 ORGANIZATION OF THIS PAPER

The remainder of this paper spans four sections. We formulate the SCUC problem in the next section. In Section 3 we present our algorithm for solving the SCUC problem, and discuss our rationale for developing the algorithm along the lines proposed. In Section 4 we present illustrative computational results. In Section 5 we close this paper by discussing promising areas for future research.

2. THE SECURITY CONSTRAINED UNIT COMMITMENT PROBLEM

We formulate the security–constrained unit–commitment problem in this section. In this paper we omit consideration of hydroelectric units in this paper and direct our discussion toward the development of SCUC solutions for thermal electric units only. Our algorithm can be extended to combined hydrothermal scheduling using the approach described in [11].

2.1 THE UNIT COMMITMENT PROBLEM

The objective of unit commitment is to develop a schedule for starting up and shutting down thermal power plants that satisfies forecasted power and reserve obligations while minimizing operating costs.

We define the following terms:

T — the number of time periods (e.g., hours) considered when developing the unit commitment schedule;

I — the number of thermal units (with the index i denoting the i^{th} unit);

K — the number of loads (with the index k denoting the k^{th} load);

L — the number of transmission lines (with the index l denoting the l^{th} line);

p_i^t — the power produced by thermal unit i during time period t;

u_i^t — 1 if thermal unit i is turned on (or kept on) during time period t; 0 otherwise;

x_i^t — the number of hours thermal unit i has been on (> 0) or off (< 0) at the beginning of time period t;

$\overline{x_i}$ — the minimum number of hours unit i must remain on after it has been turned on;

$\underline{x}_i$ the minimum number of hours unit i must remain off after it has been turned off;

$C_i(P_i^t)$ the cost of producing power at thermal unit i;

$S_i(x_i^t, u_i^t)$ the startup/shutdown cost for thermal unit i;

D_k^t the forecasted electric demand at load k during period t;

D^t The forecasted system–wide demand for time period t;

R^t the spinning reserve requirement for time period t;

$R_i(P_i^t)$ the reserve available from unit i when it is producing P_i^t;

$\overline{F_l}$ the (real) power flow limit on transmission line l;

Γ_l the matrix relating generator output to power flow on transmission line l;

$\overline{P_i}$ the upper limit on generation available from unit i; and

$\underline{P}_i$ the lower limit on generation available from unit i.

We express the SCUC problem as an optimization problem. The objective is to minimize the sum of generation and unit startup/shutdown costs over the scheduling horizon:

$$J(\underline{u}^*, \underline{P}^*)^1 =$$

$$\min_{P_i^t, u_i^t, x_i^t} \sum_{t=1}^{T} \sum_{i=1}^{I} \left[C_i(P_i^t) + S_i(x_i^t, u_i^t) \right] \quad (1)$$

- System–level demand, reserve, and security constraints must be met:

  demand obligations must be met in every time period:

$$\sum_{i=1}^{I} P_i^t \geq D^t = \sum_{k=1}^{K} D_k^t \quad (2)$$

  sufficient spinning reserve must be available in every time period:

$$\sum_{i=1}^{I} R_i(P_i^t) \geq R^t \quad (3)$$

  and generation must be distributed sufficiently throughout the system to prevent transmission lines from being overloaded:

$$-\overline{F_l} \leq F_l^t = \sum_{i=1}^{I} \Gamma_{l,i} P_i^t - \sum_{k=1}^{K} \Gamma_{l,k} D_k^t \leq \overline{F_l} \quad (4)$$

We presume the loads are fixed at each time period and are not altered by the unit scheduling optimization. We can remove them from Eq. 4 and adjust the left-hand and right-hand limits accordingly, replacing them with the time–varying terms $\underline{F}_l^t$ and $\overline{F}_l^t$, respectively.

- Unit–level operating constraints must be satisfied:

  the power produced by unit i is bounded by lower and upper limits:

$$P_i^t = \begin{cases} 0 & \text{if } u_i^t = 0 \\ \left[\underline{P}_i, \overline{P}_i\right] & \text{if } u_i^t = 1 \end{cases} \quad (5)$$

  the unit may need to adhere to must–run/must–off constraints, and the durations of time the unit is on or off must satisfy minimum up–time and down–time limits.

We refer to the "state" of a thermal unit as the number of hours the unit has been on ($x_i^t \geq 1$) or off ($x_i^t \leq -1$) at the beginning of time period t. We express the evolution of the state of a unit as follows:

$$x_i^{t+1} = \begin{cases} \max\left[1, x_i^t + 1\right] & \text{if } u_i^t = 1 \\ \min\left[-1, x_i^t - 1\right] & \text{if } u_i^t = 0 \end{cases} \quad (6)$$

The permissible scheduling options for a thermal unit at each time period are constrained by the minimum up–time and down–time constraints. A unit that is on must stay on if it has not been on for its minimum up–time, and may stay on or shut off otherwise:

$$u_i^t = \begin{cases} 1 & \text{if } 1 \leq x_i^t < \overline{x}_i \\ 0 \text{ or } 1 & \text{if } x_i^t \geq \overline{x}_i \end{cases} \quad (7)$$

Similarly, a unit that is off must stay off if it has not been off for its minimum down–time, and may stay off or turn on otherwise:

$$u_i^t = \begin{cases} 0 & \text{if } -1 \geq x_i^t \geq -\underline{x}_i \\ 0 \text{ or } 1 & \text{if } x_i^t \leq -\underline{x}_i \end{cases} \quad (8)$$

[1] We use the underscore notation (e.g., $\underline{u}$) to denote the entire set of variables (e.g., all u_i^t); we use the asterisk to denote the optimal value of the variable indicated.

2.2 THE DUAL OPTIMIZATION PROBLEM: DIRECT METHODS

The optimization problem presented in Eqs. 1 thru 8 has a special decomposable structure. The objective function (Eq. 1) is separable in the units, while the thermal units are coupled only by the demand (Eq. 2), reserve (Eq. 3), and transmission line flow constraints (Eq. 4). We can relax these constraints and append them to the objective function using *non–negative* Lagrange multipliers to obtain the following problem formulation:

$$
\begin{aligned}
q(\lambda, \mu, \alpha, \zeta) = \min_{P_i^t, u_i^t, x_i^t} \; & \sum_{t=1}^{T} \sum_{i=1}^{I} \left[C_i(P_i^t) + S_i(x_i^t, u_i^t) \right] \\
& - \sum_{t=1}^{T} \lambda^t \left(\sum_{i=1}^{I} P_i^t - D^t \right) \\
& - \sum_{t=1}^{T} \mu^t \left(\sum_{i=1}^{I} R_i(P_i^t) - R^t \right) \\
& + \sum_{t=1}^{T} \sum_{l=1}^{L} \alpha_l^t \left(\Gamma_l P^t - \overline{F}_l^t \right) \\
& - \sum_{t=1}^{T} \sum_{l=1}^{L} \zeta_l^t \left(\Gamma_l P^t - \underline{F}_l^t \right)
\end{aligned}
\tag{9}
$$

subject to the unit–level constraints (Eqs. 5–8). The variable λ^t is the Lagrange multiplier for the demand constraint at time t, while μ^t is the Lagrange multiplier for the reserve constraint at time t. The variable α_l^t is the Lagrange multiplier for the line capacity in the positive flow direction for line l at time t; the variable ζ_l^t is the complementary Lagrange multiplier for the line capacity in the negative flow direction. In all, 2T·(1+L) Lagrange multipliers are added to the problem.

We can re–write Eq. 9 as:

$$
\begin{aligned}
q(\lambda, \mu, \alpha, \zeta) = & \sum_{i=1}^{I} q_i(\lambda, \mu, \alpha, \zeta) + \sum_{t=1}^{T} \lambda^t D^t \\
& + \sum_{t=1}^{T} \mu^t R^t - \sum_{t=1}^{T} \sum_{l=1}^{L} \alpha_l^t \overline{F}_l^t + \sum_{t=1}^{T} \sum_{l=1}^{L} \zeta_l^t \underline{F}_l^t
\end{aligned}
\tag{10}
$$

where

$$
\begin{aligned}
q_i(\lambda, \mu, \alpha, \zeta) = \min_{P_i^t, u_i^t, x_i^t} \; & \sum_{t=1}^{T} \Big[C_i(P_i^t) + S_i(x_i^t, u_i^t) - \lambda^t P_i^t - \mu^t R_i(P_i^t) \\
& + \sum_{l=1}^{L} (\alpha_l^t - \zeta_l^t) \Gamma_{l,i} P_i^t \Big]
\end{aligned}
\tag{11}
$$

subject to Eqs. 5–8. The dual objective function encompasses the sum of I unit–level optimization problems which can be solved separately using dynamic programming; the last four summation terms are simply constants when the values of λ, μ, α, and ζ are fixed.

Equation 11 lends itself to a useful observation if we interpret the Lagrange multipliers as 'prices' that the utility is willing to pay for generation. The Lagrange multiplier associated with an overloaded line will *discourage* generation from units that contribute to the excess flow, and will *encourage* generation from units that can relieve the overload. The Direct method, therefore, establishes a mechanism within the dual optimization problem for distributing generation throughout the electric grid so as to meet load while satisfying security requirements. This mechanism is important because the solution to the dual optimization problem establishes a "neighborhood" of scheduling options for developing a feasible schedule. If this mechanism is not present then the dual solution may misrepresent which units should be turned on and off throughout the electric grid to best comply with transmission constraints; excessive effort may be required to obtain a feasible solution from the dual solution.

2.3 THE DUAL OPTIMIZATION PROBLEM: INDIRECT METHODS

The Direct method has the potential drawback that it adds 2T·(1+L) Lagrange multipliers to the optimization problem. The addition of these constraints may hamper the convergence of the dual optimization problem. One simplifying approach is to omit the security constraints from the dual optimization problem, and re–introduce them when trying to develop feasible unit commitment schedules from the dual solution. This is the rationale for *Indirect* methods, which omit the Lagrange multipliers for the transmission line constraints (the α's and ζ's), but retain the multipliers for the demand and reserve constraints (the λ's and μ's). This simplification not only decreases the computational effort needed to evaluate Eq. 11, but may also decrease the number of iterations needed to obtain an optimal dual solution.

The Indirect method has a potentially serious drawback: it does not provide a mechanism within the dual optimization problem for distributing generation throughout a transmission grid. The dual solutions it produces, therefore, may lead to inferior starting points for constructing primal feasible solutions.

3. THE SECURITY–CONSTRAINED UNIT COMMITMENT ALGORITHM

3.1 SOLVING THE DUAL OPTIMIZATION PROBLEM

We solve the dual optimization problem using an iterative algorithm; we illustrate the major steps of this algorithm in Fig. 2.

3.1.1 Evaluating the Dual Objective

At each iteration the algorithm evaluates the dual objective function for the current set of Lagrange multipliers using Eq. 10. The I optimization problems appearing in this equation are solved using dynamic programming [3, 6, 11].

We extract additional information from the solution to the dynamic programming problem that we use later when we develop primal feasible solutions. Let $J^t[x_i^t]$ denote the cost–to–go from time period t if unit i is in state x_i^t at the beginning of time period t. We define an *opportunity value* for turning on (or keeping on) a unit at every time period; we denote the opportunity value by $\Psi[x_i^t]$. If a unit in state x_i^t were to stay on during period t then it would realize both a "return" for generating power during this period and a cost–to–go from being in state x_i^t+1 at time t+1. Let $\phi_i(P_i^t)$ represent the return for generating power during this period :

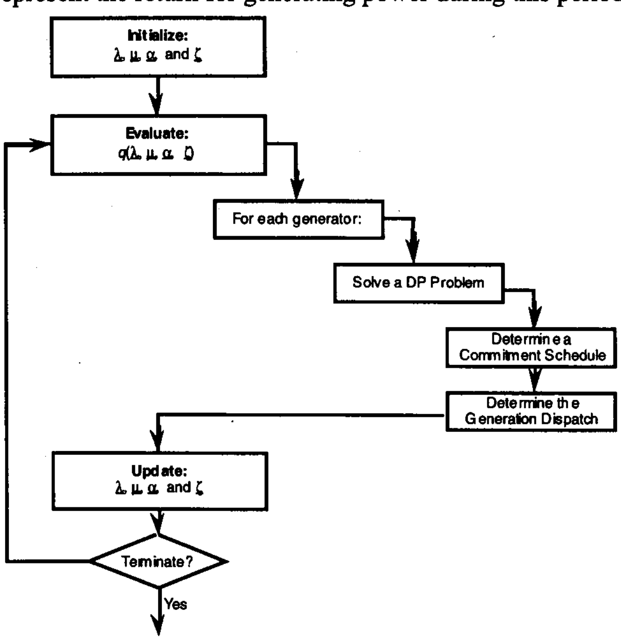

Figure 2. Overview of the Dual Optimization Algorithm

$$\phi_i(P_i^t) = \min_{P_i^t} \left[C_i(P_i^t) - \mu^t R_i(P_i^t) + \sum_{l=1}^{L} (\alpha_l^t - \zeta_l^t) \Gamma_{l,i} P_i^t \right] \quad (12)$$

If that unit were, instead, to turn off then it would realize both a shutdown cost and a cost–to–go from being in state –1 at time t+1. We set the opportunity value equal to the difference between the cost of turning *off* and the cost of staying *on*, or:

$$\Psi[x_i^t] = S_i(x_i^t, 0) + J^{t+1}[-1] - \phi_i(P_i^t) - J^{t+1}[x_i^t + 1] \quad (13)$$

if the unit is on for its minimum up time or longer, and ∞ otherwise.

Similarly, for a unit that is off we compute $\Psi[x_i^t]$ as:

$$\Psi[x_i^t] = J^{t+1}[x_i^t - 1] - \phi_i(P_i^t) - S_i(x_i^t, 1) - J^{t+1}[1] \quad (14)$$

if the unit is off for its minimum down time or longer, and $-\infty$ otherwise.

It should be clear that the best decision from state x_i^t is to stay on (or turn on) if $\Psi[x_i^t]$ is positive, and to turn off (or stay off) otherwise.

3.1.2 Updating the Lagrange Multipliers

We update the Lagrange multipliers using a subgradient method. Subgradient methods, however, are notorious for having inconsistent convergence, and they often have a tendency to "oscillate" around inferior solutions as the optimal solution is approached. This can present a potential setback for using a Direct method since such a method hampers the solution to the dual optimization problem by adding additional Lagrange multipliers. Space dilation techniques [9] have proven to be successful in this regard, but they can be difficult to implement. We use an adaptive step–size technique that detects oscillations and adjusts subsequent step sizes. We report these techniques and their computational properties in [10].

3.2 OBTAINING A FEASIBLE UNIT SCHEDULE

The unit schedule that emerges from the optimal dual solution will not normally satisfy the system–level constraints. It is necessary, therefore, to have an algorithm to construct feasible commitment schedules from the dual solution. Most dual optimization–based approaches for solving the unit–commitment problem use similar steps to solve the dual, but differ greatly in how they obtain feasible schedules. We summarize our algorithms in four major steps:

1. Construct a commitment schedule for the entire scheduling horizon $\mathbf{U} = \{u_i^t \forall i, \forall t\}$ by turning on units in every hour if their corresponding opportunity costs (Eqs. 13 and 14) exceed a specified threshold.
2. Determine for this fixed schedule the startup costs, the shutdown costs, and the productions costs in every hour. Determine the hourly production costs by solving an optimal security–constrained economic dispatch (SCED) subproblem in every time period.
3. If the cumulative cost of U is the best observed so far, record; else, discard.
4. Terminate if termination criteria have been met; else, reduce the threshold used to commit units in Step 2 and return to Step 2.

For a fixed schedule U, the SCED problem in every time period t is:

Minimize generation costs:

$$J(\underline{P}^{t*}) = \min_{P_i^t} \sum_{i=1}^{I} C_i\left(P_i^t\right) \qquad (15)$$

while meeting system–level demand (Eq. 2), reserve (Eq. 3), and transmission line constraints (Eq. 4).

We use a dual method. to solve this subproblem The critical observation here is that we can use the Lagrange multipliers from the Dual Optimization stage to *start* the dual method. This approach, as we report in the next section, can reduce appreciably the computational effort needed to obtain a feasible generator dispatch solution.

4. COMPUTATIONAL RESULTS

4.1 THE TEST SYSTEM

Our test system is a hypothetical 31–bus system comprising 43 lines, 16 generating units, and 11 load centers. Figure 3 depicts the one–line diagram for this system, showing the peak demands imposed at the 11 load centers. The most efficient generators tend to be some distance from the major load centers. These base–load units are located at the top of this illustration, at busses 3001, 1600, 101, and 102. Peak loads can be covered by the cycling units and/or by gas turbines. The gas turbines are expensive to operate, but can be brought on-line and off–line much more quickly than can the cycling units. If transmission constraints were not a factor, economy of operation for this system favors using the base–load units to carry most of the load, and using some cycling units, primarily to help cover the peak demands.

We summarize the key parameters for this system in Tables 1 thru 3. For convenience we presume the individual loads are a constant fraction of the total load; we list the contribution of each load to the total load in Table 1. Table

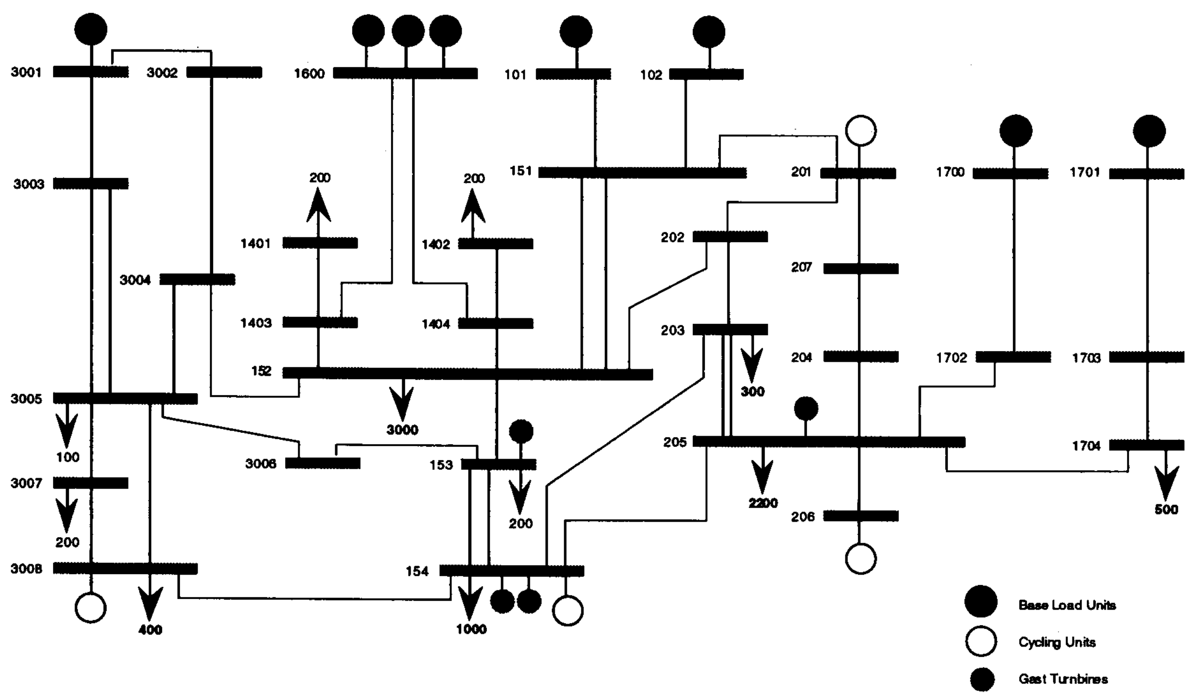

Figure 3. The One–Line Diagram for the 31–Bus Test System

2 lists the system loads by hour; hour 1 corresponds to 1 AM. Our algorithm accepts any arbitrary load profile, however. Table 3 lists several key transmission line parameters, including capacity and per unit reactance; the latter is needed to compute the distribution factor matrix Γ.

TABLE 1. PERCENT OF SYSTEM LOAD DRAWN BY EACH LOAD BUS

Bus	Percent
152	36.1
153	2.4
154	12.0
203	3.6
205	26.5
1401	2.4
1402	2.4
1704	6.2
3005	1.2
3007	2.4
3008	4.8

TABLE 2. SYSTEM LOAD BY HOUR

Hour	Load	Proportion of Peak
1	2502	30%
2	2441	29%
3	2197	26%
4	2075	25%
5	2502	30%
6	3418	41%
7	4809	58%
8	5859	71%
9	6957	84%
10	7690	93%
11	8056	97%
12	8300	100%
13	7995	96%
14	7201	87%
15	6591	79%
16	6225	75%
17	6652	80%
18	7812	94%
19	8056	97%
20	7079	85%
21	5188	63%
22	4028	49%
23	3174	38%
24	2807	34%

TABLE 3. TRANSMISSION LINE PARAMETERS

Line	Capacity (MW)	Reactance (p.u.)
151_to_101	1500	0.01
151_to_102	1500	0.01
151_to_152	1200	0.046
151_to_152	1200	0.046
151_to_201	1200	0.015
152_to_153	2500	0.005
152_to_202	1200	0.01
152_to_1403	1780	0.00815
152_to_1404	1780	0.00815
152_to_3004	1000	0.03
153_to_154	250	0.045
153_to_154	250	0.054
153_to_3006	1000	0.012
154_to_203	200	0.04
154_to_205	600	0.0333
154_to_3008	700	0.022
201_to_202	1200	0.025
201_to_207	1200	0.015
202_to_203	800	0.01
203_to_205	250	0.045
203_to_205	250	0.045
204_to_205	900	0.01

Line	Capacity (MW)	Reactance (p.u.)
204_to_207	1200	0.015
205_to_206	800	0.01
205_to_1702	1250	0.01
205_to_1704	1200	0.025
1401_to_1403	1000	0.0035
1402_to_1404	1000	0.0035
1703_to_1704	1250	0.01
3001_to_3002	1000	0.025
3001_to_3003	1000	0.008
3002_to_3004	1000	0.054
3003_to_3005	1000	0.054
3003_to_3005	1000	0.054
3004_to_3005	1000	0.01
3005_to_3006	1000	0.03
3005_to_3007	1000	0.025
3005_to_3008	1000	0.05
3007_to_3008	1000	0.025
1600_to_1403	2000	0.01
1600_to_1404	2000	0.01
1700_to_1702	1000	0.01
1701_to_1703	1000	0.01

Finally, Table 4 lists key generator parameters. For this test system the generation costs are quadratic in the power generated:

$$C_i(P_i) = a_{i1}P_i + a_{i2}P_i^2 \tag{16}$$

We presume that unit startup costs have a fixed component plus a component that varies with the length of time the unit has been off. We presume that the boilers serving unit i cool at an exponential rate inversely proportional to a cooling constant τ_i. Startup costs are computed, therefore, as:

$$S_i\left(x_i^t, u_i^t\right) = b_{i1}\left(1-e^{\frac{-x_i^t}{\tau_i}}\right) + b_{i2} \tag{17}$$

4.2 SCHEDULE COST

Were security not an issue, the system could be operated with the generation profile depicted in Fig. 4. Our algorithm evaluated its cost at \$1,107,882 for this 24–hour period. This schedule does not use the gas turbines at all, but draws on the base load units and select cycling units. For comparison, the algorithm computed a lower bound of \$1,107,583; the feasible cost is within 0.03% of the lower bound. This schedule will, however, overload numerous transmission lines during the peak demand hours, notably the transmission lines serving busses 154 and 205. Furthermore, it is noteworthy that this schedule does not use the cycling unit at bus 206 or the gas turbine at bus 205: operation of these units—together with slight redistribution of generation elsewhere—could remove these line overloads.

Our algorithm develops the generation profile depicted in Fig. 5 when generation is scheduled to avoid overloading the transmission lines. This schedule makes extensive use of the cycling unit at bus 206 and the base–load unit at bus 3001 to work around the line overload. This schedule also redistributes some generation away from the base–load units at busses 101 and 102 and toward the three base–load units at bus 1600. Our algorithm evaluated the cost of this solution at \$1,111,049, and the cost of the lower bound at \$1,110,319; the feasible cost is within 0.07% of the lower bound.

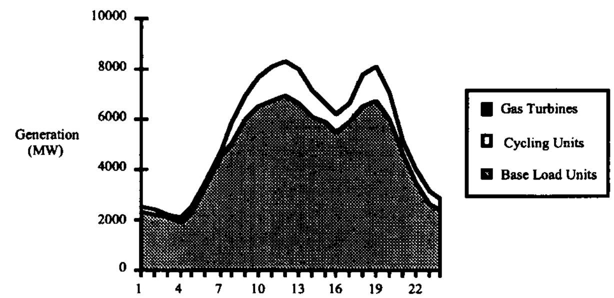

Figure 4. Generation Profile Absent Security Constraints

TABLE 4. GENERATOR PARAMETERS

Unit Name	$\overline{R_i}$	$\underline{P_i}$	$\overline{x_i}$	$\underline{x_i}$	x_i^0	a_{i1}	a_{i2}	b_{i1}	b_{i2}	t_i
Bus 3001	1000	300	5	6	5	8.752	0.0015	2050	825	6
Bus 1600_1	1200	360	5	6	5	7.654	0.0016	1460	650	6
Bus 1600_2	1200	360	5	6	5	7.654	0.0016	2100	950	6
Bus 1600_3	1200	360	5	6	5	7.654	0.0016	2200	960	6
Bus 101	1500	300	5	6	5	6.052	0.0013	2100	900	6
Bus 102	1500	300	5	6	5	6.052	0.0013	2100	750	6
Bus 201	800	240	5	6	5	9.072	0.0015	2300	950	6
Bus 1700	500	150	3	4	3	8.752	0.0015	1370	550	4
Bus 1701	500	100	3	4	3	8.752	0.0015	1180	550	4
Bus 206	600	120	3	4	-4	9.543	0.0017	1180	625	4
Bus 205	300	90	1	1	-1	11.62	0.0018	25	12	1
Bus 153	150	45	1	1	-1	12.54	0.0019	10	30	1
Bus 154_1	175	52	1	1	-1	13.00	0.0019	25	30	1
Bus 154_2	200	60	1	1	-1	14.62	0.0018	20	12	1
Bus 154_3	600	120	3	4	-4	9.543	0.0017	1200	440	4
Bus 3008	750	150	3	4	-4	8.352	0.0015	1700	550	4

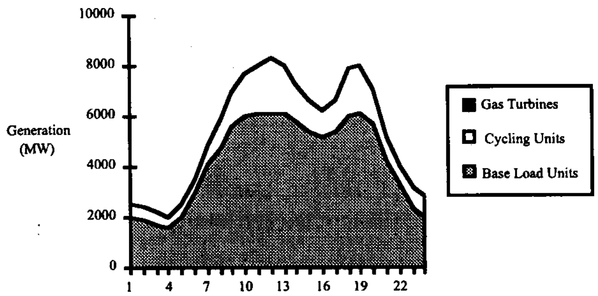

Figure 5. SCUC Generation Profile When Security Constraints Are Handled Directly

Our hypothesis is that Direct solution methods should provide better unit commitment schedules than Indirect methods; computational experience bears this out. We provided an option in the implementation of our Direct method to ignore the security constraints during the Dual Optimization stage, thereby creating an Indirect method. At the risk of belaboring the obvious, this implementation of an Indirect method differs from the implementation of our Direct method only in this one respect. Our implementation of an Indirect method produced the schedule profile illustrated in Fig. 6. This schedule turns off the unit at bus 206 during the first hour and, like the schedule depicted in Fig 4, attempts to meet the load with the base–load units and select cycling units. This strategy, as we noted earlier, cannot meet the peak load and avoid overloading transmission lines. The schedule, therefore, attempts to meet the peak load by turning on the gas turbine at bus 205. This measure is insufficient, and the schedule ultimately turns on the cycling unit at bus 206. This schedule costs $1,124,149, roughly $13,000, or 1.2%, greater than that from our Direct method.

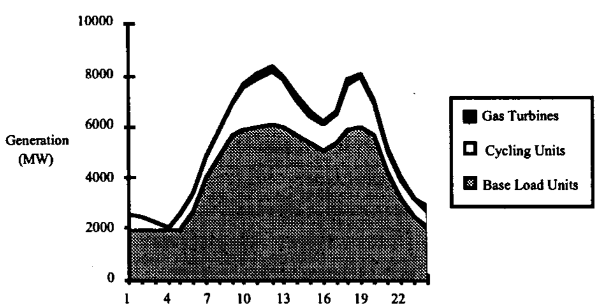

Figure 6. SCUC Generation Profile When Security Constraints Are Handled Indirectly

4.3 COMPUTATIONAL REQUIREMENTS

The Indirect method is appreciably less efficient computationally than the Direct method. The Direct method obtained its solutions within 4.4 CPU–seconds on a Macintosh Quadra 950 (with floating–point co–processor). The Indirect method required 6.2 CPU–seconds. The breakout of computation effort by solution activity for each method is listed in Table 5. We speculated in Section 2 that the Direct method would require more effort than the Indirect method to obtain an optimal dual solution; our computational results bear this out. This additional effort is spent evaluating the dual cost element $\phi_i\left(P_i^t\right)$ defined in Eq. 12.

TABLE 5. COMPUTATION TIME BY SOLUTION ACTIVITY

Activity	Direct	Indirect
Obtain an optimal dual solution	3.5	3.0
Obtain a feasible solution:	0.4	2.8
Miscellaneous	0.5	0.4
Total	4.4	6.2

The Indirect method, however, required considerably more computation time to develop a primal feasible solution. Two factors contribute to this. First, the Indirect method does not use the Dual Optimization stage to pare down the number of candidate scheduling solutions that are examined during the Primal Solution stage. The Indirect method needed to reduce the opportunity value threshold eight times to locate a first feasible solution; the Direct method needed to reduce this threshold only once. Second, the Indirect method does not provide starting solutions to the SCED algorithm—initial values for the Lagrange multipliers—that are as good as those provided by the Direct method. The Indirect method, therefore, causes the SCED algorithm to iterate longer to converge to a feasible economic dispatch.

4.4 SUMMARY

The Direct method obtains better solutions faster than the Indirect method. The Indirect method is less effective because it fails to value properly the generation that can be provided by units located close to major load centers. Instead, it attempts first to schedule units strictly on the basis of cost, and then fix that schedule to satisfy security constraints. The fix applied retroactively, however, is no match in cost for a schedule developed by the Direct method that adheres to security constraints at the outset. The effort needed to develop a fix is also greater than the effort needed to develop the correct schedule via the Direct method. Table 6 summarizes these findings.

5. CLOSING REMARKS

In this paper we have demonstrated a practical algorithm for solving the security–constrained unit–commitment problem. This algorithm uses Lagrangian relaxation to identify a "neighborhood" of solutions where we expect to find a best feasible solution. The algorithm examines candidate solutions within this neighborhood to identify a good feasible solution. Our computational results suggest that this approach obtains lower–cost solutions faster than if the security constraints are considered retroactively. Our results suggest that the retroactive consideration of the security constraints leads the algorithm to explore a larger number of infeasible candidate solutions.

Our results suggest that Lagrangian relaxation methods for unit–commitment can handle a variety of system–wide operating constraints that must be met in every time period. The security constraints, Eq. 4, do not have a particular structure and could just as easily have represented other types of constraints acting on units as a group.

5.1 POLLUTION EMISSION RESTRICTIONS

A growing number of electric utilities require the capability to schedule generating units so as to meet *pollution emission restrictions*. We feel that our algorithm can handle such restrictions. The key observation is to reinterpret the meaning of Eq. 4, replacing *transmission lines* with *'emission control areas'* and redefining the distribution factor to mean the proportion of a plant's emissions that enter each emission control area. In practice this proportion might be all or nothing, but Eq. 4 permits intermediate proportions as well. In this fashion we can express the system–wide restriction on emissions of pollutant p entering control area a as:

$$\Gamma_{\mathrm{p,a}}\,\underline{P}^t \le \overline{M}^{\,t}_{\mathrm{p,a}} \tag{18}$$

where $\overline{\mathrm{M}}^{\,t}_{p,a}$ denotes the maximum allowable mass of pollutant p that can enter control area a during time period t.

5.2 MULTI–AREA INTERCHANGE RESTRICTIONS

Utilities frequently operate with constraints on the amount of power generated outside a specified area that can serve loads in that area and/or that can be transferred through that area. We refer to these constraints collectively as *area–interchange restrictions*. These restrictions can appear in several guises; we summarize three:

1. the total power generated outside a specified area to serve loads in that area cannot exceed some threshold (possibly expressed as a fraction of the load in that area);
2. the amount of power transferred between two specified areas cannot exceed some threshold; and
3. the amount of power transferred through a specified area cannot exceed some threshold.

We add that all three restrictions can be present at once, and all can vary with time. We direct the interested reader to [5, 12, 13] for additional treatments of this problem.

The first variation listed above is really nothing more than a special restriction on the normal constraint that generation meet load (Eq. 2). For some area a, the additional constraint could stipulate that internal generation be at least as great as some proportion of the area load, or:

$$\sum_{i \in U_{\mathrm{a}}} P_i^{\,t} \ge \chi\, D_{\mathrm{a}}^{\,t} \tag{19}$$

where:

- U_a denotes the set of generating units within area a;
- $\mathrm{D}_a^{\,t}$ denotes the forecasted demand within area a during time period t; and
- χ denotes the proportion of area load that must be met by internal generation.

The second variation on the area–interchange constraint can be expressed as a limit on the total flow that can be carried on transmission lines directly entering/leaving one area that *communicate* with another area. We say that an area communicates with a transmission line if one or more busses in the area have nonzero line flow distribution factors for that line.

The third variation on the area interchange constraint can be formulated as an elaboration on the second variation. Specifically, such a constraint can be expressed as a limit on the total flow that can be carried on transmission lines directly entering/leaving a specified area, area of origin/destination notwithstanding.

TABLE 6 SUMMARY OF KEY COMPUTATIONAL RESULTS

Solution Method	Lower Bound Cost	Feasible Schedule Cost	Computation Time (CPU–seconds)†
Unconstrained‡	$1,107,583	$1,107,882	3.2
Direct	$1,110,319	$1,111,049	4.4
Indirect	$1,107,583	$1,124,149	6.2

† Macintosh Quadra 950

‡ Security constraints ignored altogether

REFERENCES

1. Muckstadt, J.A., and S.A. Koenig, "An Application of Lagrangian Relaxation to Scheduling Power–Generation Systems," Operations Research, Vol. 25, No. 3, 1977, pp. 387–403.

2. Batut, J., and A. Renaud, "Daily Generation Scheduling Optimization with Transmission Constraints: A New Class of Algorithms," IEEE Trans. on Power Systems, Vol. 7, No. 3, August 1992, pp 982–989.

3. Bard, J. F., "Short–Term Scheduling of Thermal–Electric Generators Using Lagrangian Relaxation," Operations Research, Vol. 36, No. 5, Sept.–Oct., 1988, pp 756–766.

4. Merlin, A., and P. Sandrin, "A New Method for Unit Commitment at Electricite de France," IEEE Trans. on PAS, Vol. PAS–102, No. 5, May, 1983, pp 1218–1225.

5. Zhuang, F., "Optimal Generation Unit Commitment in Thermal Electric Power Systems," Ph.D. Thesis, McGill University, December, 1988.

6. Lauer, G.S., D.P. Bertsekas, N.R. Sandell, Jr., and T.A. Posbergh, "Solution of Large–Scale Optimal Unit Commitment Problems," IEEE Trans. on PAS, Vol. PAS–101, No. 1, January, 1982 pp. 79–86.

7. Dodu, J.C., P. Martin, A. Merlin, and J. Pouget, "An Optimal Formulation and Solution of Short–Range Operating Problems for a Power System with Flow Constraints," Proceedings of the IEEE, Vol. 60, No. 1, January, 1972, pp 54–63.

8. Ruzic, S., and N. Rajokovic, "A New Approach for Solving Extended Unit Commitment Problem," IEEE Trans. on Power Systems, Vol. 6, No. 1, February, 1991, pp 269–277.

9. Poljak, B.T., "Subgradient Methods: A Survey of Soviet Research," in Nonsmooth Minimization, C. Lemarechal and R. Mifflin, Eds., Pergammon Press, London, 1978.

10. Shaw, J., "Lagrangian Relaxation Methods for Electric Power Systems," 1991 PICA Tutorial on Advanced Optimization Techniques, Baltimore, 1991.

11. Shaw, J.J., and D.P. Bertsekas, "Optimal Scheduling of Large Hydrothermal Power Systems," IEEE Trans. on PAS, Vol. PAS–104, No. 2, February, 1985, pp. 286–294.

12. Hsu, Y.Y., C.C. Su, C.C. Liang, C.J. Lin, and C.T. Huang, "Dynamic Security Constrained Multi–Area Unit Commitment," IEEE Trans. on Power Systems, Vol. 6, No. 3, August, 1991, pp 1049–1055.

13. Lee, F.N., and Q. Feng, "Multi–Area Unit Commitment," Presented at the IEEE/PES 1991 Winter Meeting, Paper 91 WM 180–0 PWS, 1991.

Discussion

Xiaohong Guan (Consultant with PG&E), **Peter B. Luh** (University of Connecticut): The discussers would like to congratulate the author for his achievement on solving the problems of transmission-constrained unit commitment and power system scheduling. As the emergence of the competitive electric energy market, transmission system is coming to be seen not just as electric wires to transport power, but as a resource to be more aggressively marketed and efficiently managed to generate more revenue. Incorporating transmission system constraints in unit commitment becomes increasingly important. The current paper will no doubt make contribution to resolving the new issues challenging power system operators. We would appreciate the authors' comments on the following problems:

1. In the paper, as the author mentioned, the direct method needs to add 2T(1+L) Lagrange multipliers to relax the transmission line flow constraints. For a small system tested in the paper with only 43 lines, more than 14,000 multipliers will be needed for an hourly commitment problem of one week. This may result in convergence problem at the high level even though adaptive step-size technique can be used to update the multipliers. Could the author comment how to resolve this problem when the direct method is applied to schedule systems of practical size, say for a system with 1000 lines.

2. The author has presented an interesting method to modify the dual solution into a feasible solution, and an impressive duality gap has been obtained. We would like to know that in Step 1, pp. 6, which opportunity value is used to determine to turn on the unit since there are many states at a particular hour. Would the state with the largest opportunity value be selected? How is the minimum up and down time satisfied in this procedure? Also if a unit has coupling constraints across hours, e.g., energy or fuel constraints, how are these constraints guaranteed to be satisfied when the commitment is modified based on cost comparison. It seems that the algorithm should return to Step 1 instead of Step 2, if it does not terminate in Step 4.

3. What kind of dual method is used for the SCED problem, which is similar to an optimal power flow problem? The computational time comparison of direct method with indirect method for solving an SCED would depend on the method used. Again, if there are hydro or fuel constrained thermal units in the system, how can the coupling constraints across hours be satisfied.

4. The CPU time seems to be very impressive for the problem tested in the paper. How many high level iterations are generally needed and how does CPU time approximately increase with problem size? Would it increase linearly?

Manuscript received August 22, 1994.

Alva J. Svoboda, Chao-an Li and **Raymond B. Johnson**, (Pacific Gas and Electric, San Francisco, CA): The author presents a direct algorithm incorporating transmission line constraints as a straightforward extension of existing Lagrangian relaxation scheduling algorithms such as the one used in PG&E's Hydro-Thermal Optimization program [A]. As the author points out, the price interpretation of the Lagrange multipliers is preserved, and the dual optimum may be closer to a primal schedule feasible to the transmission constraints. We would like the author to address the following comments and questions in this closure.

1) The line flows are based on a distribution factor matrix derived from line reactances which is equivalent to using a DC power flow. In a heavily loaded system, line flows calculated using an AC power flow algorithm may differ significantly from those calculated using a DC power flow. Has the author considered constructing an algorithm using an AC power flow or even a linearized AC power flow? What sort of hierarchical scheme would be required and how would such an algorithm perform?

2) The direct method of the author's paper adds 2T*(1+L) Lagrangian multipliers to the Unit Commitment optimization problem to incorporate the transmission security constraints. As the author himself has indicated the addition of this large a number of constraints may hamper the convergence of the dual optimization. However, no indication is given on how to handle this problem. Has the author tested the direct method with the adaptive step-size technique for practical power systems of realistic size?

3) The nature of the adaptive multiplier updates will be critical to the success of this approach. We have experienced difficulty in either fixing the scaling factors used for different sets of constraints (e.g., system constraints versus generation area constraints) or modifying the scaling factors adaptively. More details of the scaling techniques used in this algorithm may provide useful indications about how various types of coupling constraints can be dealt with simultaneously.

4) We would also like to point out that emission constraints, like fuel limits, are usually specified over multiple time periods rather than for individual time periods. How would the author extend his method to handle such constraints?

[A] L.A.F.M. Ferreira, T. Andersson, C.F. Imparato, T.E. Miller, C.K. Pang, A. Svoboda and A.F. Vojdani, "Short-term Resource Scheduling in Multi-Area Hydro-thermal Power Systems", Electrical Power and Energy Systems, Vol. 11, No. 3, July 1989.

Manuscript received August 22, 1994.

John J. Shaw: I would like to thank the discussors for their observations and comments. I first address the questions posed by Drs. Guan and Luh, and then respond to the questions posed by Drs. Svoboda, Li, and Johnson.

Response to Drs. Guan and Luh

(1) I have not yet applied the method described herein to power systems comprising several thousand transmission lines, and my comments here are, therefore, speculative rather than definitive. Our computational experience with Lagrangian Relaxation indicates that the values of the dual variables from one day to the next tend to be coupled only weakly. (The values of the dual variables for Friday thru Monday tend to be more closely coupled, but usually not to the same degree of coupling *within* a day.) I anticipate, therefore, that convergence problems will be more a function of the number of 'coupling' constraints (e.g., demand, reserve, and transmission line limits) that are actually **binding** during any day. In my view, the two issues that arise when handling several thousand transmission lines are: a) how many line constraints will actually be binding in every hour?, and b) how to avoid needless computations for lines that are never binding? Unfortunately, I have no answer for the former issue. The latter issue can be dealt with using active set methods, and my colleagues and I are currently developing one such implementation; I look forward to reporting results at a later date.

(2) We compute two opportunity costs (Eqs. 13 and 14 in the paper) for every unit for every time period: one for the state where the unit has been on for its minimum up time and one for the state where the unit has been off for its minimum down time. As Drs. Guan and Luh observe, there are possibly additional down states beyond the minimum down time that should be accounted for when startup costs depend on the number of hours a unit has been off. We do not compute opportunity costs for those down states per se. Instead, for any of those states we simply add the difference between the startup cost for that state and the startup cost for the minimum down time state to the opportunity cost for the minimum down time state. Minimum up time and down time constraints are preserved throughout. We have not yet dealt with units that have coupling constraints across hours (e.g., fuel or energy constraints). Those constraints, however, should be introduced directly into each generating unit's dual optimization problem (Eq. 12) to better account for the opportunity costs. Drs. Guan and Luh are correct in their observation that the algorithm should return to Step 1 if it does not terminate in Step 4: my paper has a typographical error on page 6.

(3) To date, we have experimented with two dual methods for solving the subproblem defined by Eq. 15: coordinate ascent and conjugate gradient. The coordinate ascent method does not require a line search to determine the best step size, but can converge slowly if two or more transmission lines have mutually competing line flow constraints (i.e., when actions taken to relieve overflow on one line create overflows on other lines). The conjugate gradient method tends to outperform the coordinate ascent method when a time period has many active coupling constraints. My colleagues and I are currently working on an algorithm that combines both methods and switches adaptively from one to the other.

(4) The computation times were for 40 subgradient iterations and 10 iterations of the primal algorithm described on page 6. The computation time per subgradient iteration is linear in the number of transmission lines, in the number of generators, and in the number of time periods, and grows with the *product* of these three parameters. The computational time per primal iteration is linear in the number of time periods, tends to be linear in the number of generators, and quadratic in the number of transmission lines.

Response to Drs. Svoboda, Li, and Johnson

(1) We have considered introducing AC flow into our algorithm. The key difficulty is that AC optimal power flow problems are demanding computationally. Our current thinking, therefore, is to introduce AC power flow selectively during our algorithm's primal stage and not during its dual stage. The hierarchy we propose currently is as follows:

a) Solve the Dual optimization problem (Eq. 9) using the DC power flow approximation.

b) In the Primal stage (described on page 6) introduce a new step between the current steps 3 & 4 to check the feasibility and cost of the incumbent solution using AC power flows.

(2) See my response to Drs. Guan & Luh in (1) above.

(3) We have experimented with a number of adaptive step size rules for this problem, and describe them in detail, together with computational results, in Reference 10. Briefly, these rules reduce the step size if the current subgradient iteration does not improve the dual solution, and take progressively larger step sizes (up to some threshold) otherwise.

(4) Our current thinking on handling constraints that span several hours (e.g., fuel and emission limits) is as follows. We relax these constraints (just as we relax the system demand, reserve, and transmission constraints) and introduce them into the dual objective. We compute the fuel use (or emission) obtained by the dual solution in every hour and from that establish for the primal phase a *target* fuel use (or emission) *allowance* for every hour, where the sum of the hourly targets satisfies the applicable constraint for the entire schedule. We then impose those targets in every hour when computing a primal solution. We are working on an implementation of this approach, and hope to report results soon.

Manuscript received October 27, 1994.