Generation Scheduling With Thermal Stress Constraints

Zuyi Li, *Member, IEEE*, and Mohammad Shahidehpour, *Fellow, IEEE*

***Abstract*—This paper describes a scheduling method for representing the thermal stress of turbine shafts as ramp rate constraints in the thermal commitment and dispatch of generating units. The paper uses Lagrangian relaxation for optimal generation scheduling. In applying the unit commitment, thermal stress over the elastic limit is used for calculating the ramping cost. The thermal stress contribution to generation cost requires the calculation of a set that includes thermal stress at the end of each time step; this requirement presents a complicated problem which cannot be solved by an ordinary optimization method such as dynamic programming. The paper uses an improved simulated annealing method to determine the optimal trajectory of each generating unit. Furthermore, the paper uses linear programming for economic dispatch in which thermal stress limits are incorporated in place of fixed ramp rate limits. The paper illustrates the economics of frequently ramping up/down of low cost generating units versus the cost of replacement of their turbine rotors with a shorter life span. The experimental results for a practical system demonstrate the effectiveness of the proposed method in optimizing the power system generation scheduling.**

***Index Terms*—Lagrangian relaxation, ramp rates, short term generation scheduling, simulated annealing, thermal stress.**

I. Introduction

ELECTRICAL power systems are designed and operated to meet the continuous variation of consumer demands. In order to ensure an optimal economic operation, power systems scheduling is based on two important tasks (i.e., unit commitment and economic dispatch). The generation scheduling objective is to minimize the system generation cost over a period of 24 to 168 h, while satisfying multiple physical and operational constraints. From the mathematical point of view, this problem represents a large and complex mixed integer programming. Research and development endeavors on this subject have been directed toward more efficient and near-optimal generation scheduling solutions in power systems with increasingly more diverse constraints.

Traditionally, unit commitment [1]–[4] has adopted a fixed ramp-rate limit for representing physical limits of generating units, such as the rotor fatigue effect, on unit ramping characteristics. This scheme presumes that as long as the ramp-rate is within preselected limits, the ramping process will not shorten the life of the rotor. Even though the use of properly preselected ramp-rate limits guarantees a more reasonable generation schedule, it fails to provide the flexibility for allowing system operators to select proper ramping rates for various levels of power.

In [5], we use a fatigue index, instead of fixed ramp rates, to balance the benefits of ramping the units with low operating cost against the cost of shortening the service life of the turbine rotor. We use two parameters: one is the elastic range which represents the limited change in generating power for a given ramping time that would not shorten the service life of the rotor, and the other one is the ramping penalty, which is based on the cost of the future rotor replacement allocated to each megawatt of ramping beyond the elastic range. These two parameters are used as limits in our economic dispatch or as additional cost in unit commitment, which result in a more exact solution in operation of units.

In [6], we present another model for optimal generation scheduling, which includes unit ramping cost as part of the overall system operation cost. In that model, the effect of fatigue on the life of a unit is converted into factors of ramping cost, which is minimized by considering various ramping policies. The ramping cost in power systems operation includes the fuel cost for generating the required electrical energy and starting up decommitted units as well as the rotor depreciation during ramping processes such as starting up, shutting down, loading, and unloading. In that model, the unit minimum up/down times constraints are represented by ramping costs, which conceptually and computationally improve the optimal generation scheduling.

Reference [7] presents a method for steam electric generator dispatch which replaces fixed ramp rate constraints by thermal stress computation in the dynamic dispatch of generating units. At power stations, changes in turbine rotor temperature can be monitored by temperature sensors, and a process computer uses this information to control the rate at which a generator is loaded and unloaded. The approach incorporates into modified dispatch equations a calculation of thermal stress of the rotor. The calculation parameters are taken from turbine manufacturers' fatigue index curves.

This study proposes a generation scheduling method, which considers thermal stress as a limit for generation scheduling and dispatch. Thermal stress cost denotes the effect of strain due to peak tensile or compressive thermal stress in the nonelastic region of generating units. Therefore, instead of fixing ramp-rate limits as one set of unit operating constraints, the proposed approach permits variable ramping rates and determines a suitable generation schedule by minimizing the overall system operation cost. The cost, however, includes a component that depends on the number of stress/strain cycles and their peak values.

Manuscript received January 14, 2003.

The authors are with the Electric Power and Power Electronics Center, Illinois Institute of Technology, Chicago, IL 60616 USA (e-mail: lizuyi@iit.edu; ms@iit.edu).

Digital Object Identifier 10.1109/TPWRS.2003.818698

Following a mathematical formulation of the problem, the proposed approach will be described. Then, the experimental results for a practical system demonstrate the effectiveness of the proposed approach in optimizing the power system generation scheduling.

II. Mathematical Models

The objective of the generation scheduling problem studied in this paper is to minimize the system operation cost, which includes the cost of fuel consumed for electrical energy generation and starting up processes, as well as the thermal stress contributions to operating cost, without violating any system operation constraints.

The list of symbols used in this paper is as follows

$\mathcal{F}$	total operating cost of the entire period;
w	width of time step; in our study, w is half an hour;
N_t	total study time span in half hours, while N_t is the set index;
N	total number of units;
$I_i(t)$	commitment state of unit i at end of time step t (1 or 0); no power is generated if $I_i(t) = 0$ at either the beginning or end of a time step;
$P_i(t)$	power output of unit i at end of time step t (MW);
$\tilde{P}_i(t)$	average power of unit i during time step t (MW);
$E_D(t)$	system energy demand for interval ending at time t (MWh);
$\bar{P}_i$	rated upper power limit of unit i (MW);
$\underline{P}_i$	rated lower power limit of unit i (MW);
$P_R(t)$	system spinning reserve requirement at time step t;
$\lambda(t)$	Lagrangian multiplier for system power balance constraint;
$\mu(t)$	Lagrangian multiplier for system spinning reserve constraint;
L_i	thermal stress limit for turbine rotor of unit i, °F;
$S_i(t)$	thermal stress on unit i at end of time step t, °F;
k_i	slope of turbine metal temperature versus power output, $^\circ\mathrm{F/MW}$;
T_i^f	first stage temperature at full power of unit i, °F;
T_i^v	first stage temperature at transition to partial arc steam admission, °F;
P_i^v	turbine loading at transition to partial arc steam admission, MW;
τ_i	turbine rotor thermal time constant for unit i, h;
$C_{si}(t)$	start up cost, \$ of unit i at time step t;
α_i, β_i	coefficients of the start-up cost function of thermal unit i;
ρ_i	time constant in the start up cost function of thermal unit i;
$C_{hi}(t)$	cost of turbine thermal stress, \$ of unit i at time step t;
$Q_i(t)$	magnitude of peak tensile or compressive thermal stress in the nonelastic region for unit i, $^\circ\mathrm{F}$ at time step t;
TE_i	temperature corresponding to elastic limit of unit i, °F;
d_i	cost coefficient for elastic strain $\$/(^\circ\mathrm{F})^2$;
T_0	initial control temperature;
$T_{\min}$	lower bound on the control temperature;
$F_i(\tilde{P}_i(t))$	fuel cost rate of unit i, U.S.\$/h
$X_i^{on/off}(t)$	time duration for which unit i has been on/off at hour t (Hr);
$T_i^{on/off}(t)$	minimum up/down time of unit i.

Mathematically, the optimization problem can be described as follows. The objective of the problem is to minimize

$$\mathcal{F} = \sum_{i=1}^{N}\sum_{t=1}^{N_t}\left[I_i(t)wF_i(\tilde{P}_i(t)) + C_{si}(t) + C_{hi}(t)\right] \quad (1)$$

where

$$\begin{aligned}\tilde{P}_i(t) =& \frac{1}{2}\left(P_i(t) + P_i(t-1)\right)\\ C_{si}(X_i^{off}(t)) =& \alpha_i + \beta_i\left(1 - e^{X_i^{off}(t)/\rho_i}\right)\\ C_{hi}(t) =& d_i \cdot (Q_i(t) - TE_i)^2.\end{aligned}$$

The objective is to minimize the system operation cost which consists of three elements. One is the power generation cost for supplying the system load demand and reserves, the second is the start-up cost, and the last one is the cost attributed to the turbine thermal stress.

The constraints for the problem are

1) System energy balance

$$\sum_{i=1}^{N} I_i(t)\tilde{P}_i(t) = \frac{E_D(t)}{w}, \quad t = 1, \ldots, N_t. \quad (2)$$

2) System spinning reserve requirements

$$\sum_{i=1}^{N} \bar{P}_i(t)I_i(t) \geq \frac{E_D(t)}{w} + P_R(t), \quad t = 1, \ldots, N_t. \quad (3)$$

3) Unit generation limits

$$\underline{P}_i \leq P_i(t) \leq \bar{P}_i(t), \quad i = 1, \ldots, N; t = 1, \ldots, N_t. \quad (4)$$

4) Thermal stress constraints

$$-L_i \leq S_i(t) \leq L_i, \quad i = 1, \ldots, N; t = 1, \ldots, N_t \quad (5)$$

where

$$\begin{aligned}S_i(t) =& S_i(t-1) \cdot e^{-w/\tau_i} + a_i \cdot [P_i(t) - P_i(t-1)]\\ a_i =& k_i \frac{1 - e^{-w/\tau_i}}{\frac{w}{\tau_i}}\\ k_i =& \frac{T_i^f - T_i^v}{\bar{P}_i - P_i^v}.\end{aligned} \quad (6)$$

5) Thermal unit minimum start up/shut down times

$$\begin{aligned}(X_i^{on}(t-1) - T_i^{on}) * (I_i(t-1) - I_i(t)) \geq& 0\\ \left(X_i^{off}(t-1) - T_i^{off}\right) * (I_i(t) - I_i(t-1)) \geq& 0.\end{aligned} \quad (7)$$

Equation (6) represents the dynamic equation of the thermal stress of a unit at a specific time step, which is a linear function of the unit's thermal stress at the previous time step and the unit's power output at the current and the previous time steps. Due to the dynamic process of the thermal stress change, the peak value $Q_i(t)$ may be different than $S_i(t)$. See [7, Fig. 1] for more details.

The difference between this study and our earlier study [8]–[12] is that the thermal stress is used as a constraint to unit operation and its excess over the elastic limit will cause a penalty to the operation cost of the unit $C_{hi}(t) = d_i \cdot (Q_i(t) - TE_i)^2$. From (6), we know that the thermal stress of a unit at time t is determined by a dynamic equation, which is a linear function of its thermal stress at time $t-1$ and the unit power outputs at times t and $t-1$. The unit power output is already determined by the economic dispatch subproblem; thus as long as the unit states at time t are given, the thermal stress at time t will be calculated step by step using (6). If the unit power output jumps between two power values, the thermal stress may exceed its elastic limit which would contribute to the operation cost. However, if the unit power output change is moderate (e.g., a small change between two power values), the thermal stress will be in its elastic range and there will be no penalty attributed to the operation cost.

In some cases, the thermal stress contribution to operation cost $C_{hi}(t)$, along with the inclusion of start-up cost $C_{si}(t)$, allows us to use rotor fatigue costs to represent minimum up-time constraints. Minimum down-time, however, may be based on the maximum speed in which station operators may be able to start a unit.

III. Proposed Solution Strategy

The Lagrangian relaxation method (LRM) is used in this study for adjoining the system constraints into cost factors via Lagrange multipliers, thus reducing the number of variables in the optimization process by decomposing the original problem into several subproblems. Since each subproblem in LRM considers the operation strategy of each unit over the entire time span, LRM is proposed for handling ramp-rate constraints. In our study, thermal stress below the elastic limit will not lead to fatigue; otherwise, it will cause rotor depreciation during ramping processes. The excess amount of thermal stress over the elastic limit is used to calculate the fatigue contribution to operation cost. Since the calculation of ramping cost needs the complete set of thermal stress values at each time step, we present an improved simulated annealing method to determine the optimal trajectory of each generating unit. Once the status of units is determined, we apply linear programming (LP) to economic dispatch. In our method, thermal stress is used in place of ramping limits to mimic the process that takes place in power stations; this is more accurate than the use of a fixed ramp rate limit.

In the following, the proposed approach for the optimal generation scheduling is given. An iterative procedure is used to search for a suitable solution for $\mathcal{L}(P_i(t), I_i(t), \lambda(t), \mu(t))$ with relaxed system energy balance constraints. The procedure for finding a suitable unit commitment schedule is as follows:

Step 1) Assign an initial value to $\lambda(t)$ and $\mu(t)$, $t = 1, \ldots, N_t$, denoted by $\hat{\lambda}(t)$ and $\hat{\mu}(t)$.

Step 2) Find the suitable unit commitment schedule under current system operating status by simulated annealing method. That is

$$\min \mathcal{L}\left(P_i(t), I_i(t), \hat{\lambda}(t), \hat{\mu}(t)\right)$$

where the solution is denoted by $\hat{P}_i(t)$ and $\hat{I}_i(t)$.

Step 3) If $\sum_{i=1}^{N} \bar{P}_i(t)\hat{I}_i(t) \geq (E_D(t)/w) + P_R(t)$ for $t = 1, \ldots, N_t$, we have obtained a satisfactory unit commitment schedule; proceed with the economic dispatch algorithm; otherwise, go to Step 4).

Step 4) Corresponding to the unit states $\hat{P}_i(t)$ and $\hat{I}_i(t)$, use the subgradient method to modify $\lambda(t)$ and $\mu(t)$. Once the updated $\hat{\lambda}(t)$ and $\hat{\mu}(t)$ are obtained, go back to Step 2).

A. Lagrangian Function for Unit Commitment

LRM decomposes the original scheduling problem into subproblems. Each subproblem corresponds to the most economical operation strategy in operating a generating unit during the entire study period and under certain system conditions. The system conditions are described by Lagrangian multipliers which physically represent the system operating condition, including the incremental cost of the system energy, the incremental cost of the system generating capacity, and so on. Each subproblem considers a dynamic process which determines the unit state (either on or off) in each time interval for minimizing the operation cost of the unit over the entire study period.

The Lagrangian function for unit commitment is

$$\begin{aligned}
&\mathcal{L}\left(P_i(t), I_i(t), \lambda(t), \mu(t)\right)\\
&\quad= \sum_{i=1}^{N}\sum_{t=1}^{N_t}\left[wF_i\left(\tilde{P}_i(t)\right)I_i(t) + C_{si}(t) + C_{hi}(t)\right]\\
&\qquad - \sum_{t=1}^{N_t}\lambda(t)\left(\sum_{i=1}^{N} I_i(t)\tilde{P}_i(t) - \frac{E_D(t)}{w}\right)\\
&\qquad - \sum_{t=1}^{N_t}\mu(t)\left(\sum_{i=1}^{N}\bar{P}_i(t)I_i(t) - \frac{E_D(t)}{w} - P_R(t)\right) \quad (8)
\end{aligned}$$

with constraints expressed by (4)–(6).

Therefore, the search for the optimal commitment schedule under certain system condition, that is for a given $\lambda(t)$ and $\mu(t)$, $t = 1, \ldots, N_t$ which are denoted by $\hat{\lambda}(t)$ and $\hat{\mu}(t)$, becomes a minimization process for variables $P_i(t)$ and $I_i(t)$.

Rearranging (8) to categorize all variables and constants, we have

$$\begin{aligned}
&\min \mathcal{L}\left(P_i(t), I_i(t), \hat{\lambda}(t), \hat{\mu}(t)\right)\\
&\quad= \sum_{i=1}^{N}\left\{\sum_{t=1}^{N_t}\left[wF_i\left(\tilde{P}_i(t)\right)I_i(t) + C_{si}(t) + C_{hi}(t)\right]\right.\\
&\qquad\left. - \sum_{t=1}^{N_t}(\hat{\lambda}(t)I_i(t)\tilde{P}_i(t) - \sum_{t=1}^{N_t}\hat{\mu}(t)\bar{P}_i(t)I_i(t)\right\}\\
&\qquad + \sum_{t=1}^{N_t}\left[\hat{\lambda}(t)\frac{E_D(t)}{w} + \hat{\mu}(t)\left(\frac{E_D(t)}{w} + P_R(t)\right)\right]. \quad (9)
\end{aligned}$$

Since the last two items of (9) are constant, they do not affect the final solution of (9), and thus, can be ignored in searching the optimal commitment schedule. In this regard, the optimization problem presented in (9) can be decomposed into N subproblems, each corresponding to the optimal unit commitment of individual units over the entire study period, which is expressed as

$$\min \mathcal{L}\left(P_i(t), I_i(t), \hat{\lambda}(t), \hat{\mu}(t)\right) = \sum_{t=1}^{N_t}\left[wF_i\left(\tilde{P}_i(t)\right) - \hat{\lambda}(t)\tilde{P}_i(t) - \hat{\mu}(t)\bar{P}_i(t)\right] I_i(t) + \sum_{t=1}^{N_t}\left[C_{si}(t) + C_{hi}(t)\right] \quad (10)$$

subject to (4)–(6).

Since the stress contribution to cost $C_{hi}(t)$ requires the complete set of values for thermal stress at the end of each time step, it will be difficult to use traditional optimization methods such as dynamic programming to obtain the optimal solution. So we adopt an improved simulated annealing method to find the suitable values of $I_i(t)$, which minimize (10) and satisfy (4) and (6).

B. Simulated Annealing Algorithm for Unit Commitment

In this study, the operation cost consists of three terms, one is the system generation cost which represents the fuel and variable operation and maintenance (O&M) expense for supplying the system load demand, another is the start-up cost $C_{si}(t)$ which can be determined by the unit operation trajectory and the third one is the thermal stress contribution $C_{si}(t)$ to the unit operation cost.

1) Peak Stress: The total study time N_t can be divided into a sequence of intervals in which the thermal stress is i) tensile beyond the elastic limit, ii) compressive beyond the elastic limit, or iii) within the elastic limit. Changing over to and from the elastic region separates the time intervals. For each set of continuous time steps of type i) or ii), there is a cost $C_{hi}(t)$ associated with the peak value of the stress $Q_i(t)$ that occurs within the time interval. Since simulated annealing can identify the complete set of stresses $S_i(t), t = 1, \ldots, N_t$, its application to our approach can spot the peak stress in each type i) or type ii) continuous interval, and thus, account for the cost $C_{hi}(t)$ contributed during each of these intervals.

Mathematically, each subproblem represents a 0–1 integer programming (IP). For a small size problem, the enumeration method is feasible since the number of configurations is small. However, in this study, the number of configurations becomes excessively large ($\mathrm{NC} = 2^{N_t}$), and the applicability of the enumeration method will be infeasible. If we apply DP, which is used extensively in LRM, with a single state variable to determine an optimal trajectory with minimum operation cost, we will have difficulty in obtaining the optimal solution. One condition for the application of dynamic programming (DP) is the separability of the objective function, which means the objective function must be separable for all k, or the effect of final k stages on the objective function of an n-stage process must depend only on states $I_i(n-k)$ and upon the final k decisions $x_{n-k+1}, x_{n-k+2}, \ldots, x_n$. However, for our commitment subproblem, the calculation of stress contribution to cost does require the previous values of the thermal stress at times $1, 2, \ldots, k$; thus, the separability condition of the objective function is not satisfied. So, we cannot resort to the one-dimensional (1-D) DP for the solution of this problem. One possible approach would be to add another state variable to describe dynamic stress equations, thus a two-dimensional (2-D) problem is required to formulate and obtain a DP solution; however, this effort will require a much greater numerical computation and storage which preclude its application to our problem.

We envisioned that the calculation of stress contribution to cost is relatively simple. Once the unit states $I_i(t)$ at each time t are given, the complete set of values for thermal stress at each time can be determined by (6), while the operation cost of unit will be calculated by (9). Hence, simulated annealing [13]–[15] can be applied readily to obtain the optimal trajectory; in this study, we present an improved simulated annealing method to determine the optimal trajectory of each generating unit more efficiently.

2) Simulated Annealing: Assume $S = (s_1, s_2, \ldots, s_{NC})$ is the set of all combinatorial configurations, and $C: S \longrightarrow R$ is a cost function given by (9), which assigns a real number to each configuration. Thus, the combinatorial problem is to find a configuration $s^* \in S$ for which

$$C(s^*) = \min C(s_i), \quad \forall\, s_i \in S.$$

For our unit commitment subproblem, this configuration is equivalent to finding an optimal trajectory for each unit i.

The fundamental concept of simulated annealing is based on a strong analogy between the physical annealing process of solid materials and the problem of solving large combinatorial optimization problem. In that case, the configurations assume the role of states of a solid material, while the cost function C and the control parameter T represent energy and temperature, respectively. Let T decrease slowly from a high value; for each T, use the Metropolis algorithm to simulate the evolution of a solid to thermal equilibrium, that is, given a configuration s_i, generate another configuration s_j by choosing an element at random from the neighborhood of s_i. If $\Delta C_{ij} = C(s_j) - C(s_i)$, then accept s_j as a new state with a probability $\exp(-\Delta C_{ij}/kT)$. This process is continued until the probability distribution of configurations satisfies the Boltzmann distribution as follows:

$$f = Z(T)e^{-C(s_i)/kT}, \quad Z(T) = \frac{1}{\sum_i e^{-C(s_i)/kT}}.$$

If T decreases very slowly and $T \longrightarrow 0$, then, based on the above equation, the final state is the one with minimum cost.

The above concepts can be implemented by two separate algorithms, one is the annealing process (AP), and the other is the Metropolis Sampling (MS) algorithm.

Annealing Process (AP) algorithm:

1 Select an initial state s_0 at random; let $s(0) = s_0$, the initial temperature be T_0 and $i = 0$

2 Let $T = T_i$; call Metropolis algorithm using T and $s(i)$, and let the returned state s of MS be the present state $s(i) = s$

3 Decrease T according to some perceived rules (e.g., $T = T_{i+1}$) where $T_{i+1} < T_i$ and $i = i + 1$

4 Check the stopping criterion, if satisfied, go to 5; otherwise, go to 2
5 Represent $s(i)$ as the optimal solution, stop.

Metropolis Sampling (MS) algorithm: According to the input s and T from AP,
1 Let $k = 0$ and $s(0) = s$
2 Generate a trial state $s(k)$ from the neighbor set of s. For our commitment subproblem, this new state can be obtained by choosing a random number t from $(1, 2, \ldots, N_t)$, then changing the status of a unit at time t; that is, if the unit is on at t, let it off and vice versa. We denote the new state as s'. Calculate the cost of s' by **(9)** and obtain the change of cost as

$$\Delta C' = C(s') - C(s(k)).$$

3 If $\Delta c' < 0$, accept s' as the present state; if $\Delta C' \geq 0$, then accept s' with a probability of $\exp(-\Delta C'/T)$. If s' is accepted, let $s(k+1) = s'$; otherwise, $s(k+1) = s(k)$
4 $k = k + 1$, check the convergence criteria, if satisfied, go to 5; otherwise, go to 2
5 Return AP with $s(k)$ as the present state.

From Steps 2 and 3, we observe that, for a given temperature T, the time series $s(0), s(1), \ldots, s(k)$ reflects not only the updating procedure of the solution but also the searching trace in the configuration set. Since the algorithm accepts the bad solution with probability $\exp(-\Delta C'/T)$, the time series does not decrease monotonically. In simulated annealing, the present solution is a time updating series which is cascaded by series at different decreasing values of T. This is why the simulated annealing has the ability to exit from a local minima; due to this reason, the present solution may be worse than some of the solutions searched. From the above analysis, we construct the monotonically decreasing series $ss(k)$ as follows:

$$ss(k) = \begin{cases} s(0), & \text{if } k = 0 \\ s(k), & \text{if } C(s(k)) < C(ss(k-1)) \\ ss(k-1), & \text{otherwise} \end{cases} \tag{11}$$

with the following advantages.

1) Since the searching strategy is not changed, the ability to exit from local minima is maintained; the final solution must be the best one of all the states visited, which enhances the ability to obtain the optimal solution without increasing the computation requirements. So, the solution obtained by this improvement is always superior to that of the original simulated annealing method.
2) For a given temperature T, $ss(i) = ss(i+1) = \cdots = ss(i+q)$ indicates that after q consecutive searches, the solution is not better than $ss(i)$. If q is large enough, there is no sense to continue the search. So, we should provide a threshold value q_0, and if $q > q_0$, we stop the MS process for a given temperature T.
3) For a given temperature T, denote its final solution $ss(k)$ as $ss(T_i)$. If for each i, $ss(T_i) = ss(T_{i+1}) = \cdots = ss(T_{i+p})$, then after p continuous drop in the temperature, the optimal solution is not improved. Thus, there is no sense to decrease the temperature any further. So, we provide a threshold value p_0 and stop the annealing process for $p > p_0$.

Therefore, the improved simulated annealing algorithm is described as follows:

Improved Annealing Process (IAP) algorithm
1 Select an initial optimal solution s^* at random with an initial state s_0 and let $s(0) = s_0$; also let the initial temperature be T_0, for $i = 0$ and $p = 0$.
2 Let $T = T_i$, and call the Improved Metropolis algorithm using T, s^*, and $s(i)$. Let the returned optimal solution be $(s')^*$ and the returned present state be $s(i) = s'(k)$.
3 If $C(s^*) < C((s')^*)$, $p = p + 1$; otherwise, let $s^* = (s')^*$ and $p = 0$.
4 Decrease the temperature according to some perceived rules (i.e., $T = T_{i+1}$), where $T_{i+1} < T_i$; let $i = i + 1$.
5 If $p > p_0$, go to 6; otherwise, go to 2.
6 Print s^* as the optimal solution, stop.

Improved Metropolis Sampling (IMS) algorithm: According to $s^*, s(i)$ and T obtained from IAP,
1 Let $k = 0$ and the present state be $s'(0) = s(i)$. The initial optimal solution is $(s')^* = s^*$ and $p = 0$.
2 Generate a trial state $s(k)$ from the neighboring set of $s = s'(k)$ by using the same random unit commitment generator as in the MS algorithm. Calculate the change of cost as

$$\Delta C' = C(s') - C(s(k)).$$

3 If $\Delta C' < 0$, then s' is the next present state. If $C((s')^*) > C(s')$, then $(s')^* = s'$ and $q = 0$; otherwise, $q = q + 1$; if $\Delta C' \geq 0$, then accept s' with probability $\exp(-\Delta C'/T)$. If s' is accepted, let $s'(k+1) = s'$ and $q = q + 1$; otherwise, $s'(k+1) = s'(k)$.
4 $k = k + 1$, if $q > q_0$ (convergence criteria) go to 5; otherwise, go to 2.
5 Return to IAP with $(s')^*$ as the optimal solution and $s'(k)$ as the present state.

C. Constrained Economic Dispatch

Once the status of each unit is determined, the second part of the problem concerning thermal stress constraints will be to solve the economic dispatch problem. In this study, thermal stress of each unit is used as a limit, instead of fixed ramp rates, for the operation of the unit which permits a dispatch at higher ramp rates without increasing turbine rotor fatigue. Linear programming (LP) is used for the optimal economic dispatch in which a thermal model of the rotor is incorporated to limit the effect of fatigue due to thermal stress. The objective of the economic dispatch is as follows:

$$\min \sum_{i \in U_k^{on}} \sum_{t \in T} F_i(\tilde{P}_i(t)) \tag{12}$$

where U_k^{on} is the set of all committed units.

We apply the upper bounded technique to the simplex method. However, as we use the piecewise linear fuel cost function, the dimension of the problem will exceed the capability of personal computers. Since the piecewise fuel cost of a unit is continuously differentiable except on vertices, it would be possible to find the optimal solution as long as the vertices can be checked. In this regard, the objective function is (12), subject to (4) and (5). The details for solving this problem are the same as that in [8]–[10].

D. Subgradient Method for Modifying $\lambda(t)$ and $\mu(t)$

The method is based on the approach given in [8]–[12], and is implemented as follows.

1) The incremental cost of the system power generation $\lambda(t)$.
    a) In Step 4 of the solution procedure, the power balance constraint is relaxed, so

$$\sum_{i=1}^{N} \hat{I}_i(t)\tilde{P}_i(t) \neq \frac{E_D(t)}{w}.$$

Using the subgradient method, $\lambda(t)$ can be updated as

$$\lambda(t) = \max\left\{\left[\lambda(t) + \xi\left(\frac{E_D(t)}{w} - \sum_{i=1}^{N}\hat{I}_i(t)\tilde{P}_i(t)\right)\right], 0\right\} \tag{13}$$

where ξ is the step size as a function of the number of iterations.

    b) System energy balance constraints will be satisfied by economic dispatch. If the convergence criterion is not met, $\lambda(t)$ will be modified. Since we use a piecewise linear fuel cost function, the highest incremental cost of all committed units, denoted by η, is used in order to compensate for the system $\lambda(t)$. Therefore

$$\lambda(t) = \sigma\lambda(t) + (1 - \sigma)\eta \tag{14}$$

where σ is a relaxation coefficient.

2) The incremental cost of system generating capacity $\mu(t)$. The following procedure is applied to the iterative process.
    a) Set $\varepsilon(t) = 0$.
    b) If $(E_D(t)/w) + P_R(t) \leq \sum_{i=1}^{N} \bar{P}_i(t)\hat{I}_i(t) \leq (E_D(t)/w) + P_R(t) + \varepsilon(t)$, then neither $\mu(t)$ nor $\varepsilon(t)$ will be modified; otherwise, go to C.
    c) According to the subgradient method, $\mu(t)$ is updated as

$$\mu(t) = \max\left\{\left[\mu(t) + \xi'\left(\frac{E_D(t)}{w} + P_R(t) + \varepsilon(t) - \sum_{i=1}^{N}\hat{I}_i(t)\bar{P}_i(t)\right)\right], 0\right\} \tag{15}$$

where ξ' is the step size as a function of the number of iterations. As for $\varepsilon(t)$, if $\sum_{i=1}^{N} \bar{P}_i\hat{I}_i(t) \geq (E_D(t)/w) + P_R(t) + \varepsilon(t)$, then $\varepsilon(t)$ will not be adjusted; otherwise, increase $\varepsilon(t)$ by a constant value equal to the smallest capacity of all units (i.e., $\varepsilon(t) = \varepsilon(t) + Constant$).

TABLE I
EQUIPMENT OUTAGE TABLE

Station	Available Capacity (MW)	Reason	Off Date	Due Date
Oil 4	0.0	Periodic inspection	01/13/2002	02/03/2002
Oil 5	0.0	Repair pump	01/13/2002	02/03/2002
Coal 1	197	Reduction: replace seals on No. 1-B high pressure boiler feed pump	01/29/2002	01/30/2002

TABLE II
CHARACTERISTICS OF THE AVAILABLE UNITS

		Coal 1	Coal 2	Coal 3	Oil 1-2	Oil 3
fixed O&M cost * ($/kW-YR)		22.4	24.9	21.6	21.6	15.3
variable O&M cost ($/MWh)		12.4	12.4	13.7	13.7	0.0
fuel cost ($/MBtu)		163.2	162.7	143.9	204.6	204.6
unit startup cost coefficients	α_i ($)	46930	38416	6085	3841.6	3941.6
	β_i ($)	0.0	0.0	0.0	0.0	7910
	ρ_i (h)	1.0	1.0	1.0	1.0	8.8
ramp up rate (MW/h)		140.0	120.0	80.0	130.0	170.0
ramp down rate (MW/h)		130.0	110.0	70.0	160.0	200.0
min up time (h)		10	9	9.5	8	7.5
min down time (h)		10	8.5	10	6	5.5
unit capacity block (MW)		90.0 192.0 261.0 277.0	112.0 184.0 294.0	9.0 64.0 127.0 144.0	30.0 265.0 361.0 381.0	29.0 70.0 141.0 187.0 216.0
block heat rate (MBtu/MW)		13.385 10.421 9.746 9.721	13.371 11.358 10.212	36.856 11.440 9.770 9.745	18.126 11.014 10.946 10.931	24.470 18.517 16.494 16.138 16.054
thermal stress up limit (Fahrenheit)		200.0	200.0	90.0	120.0	150.0
thermal stress down limit (Fahrenheit)		-200.0	-200.0	-90.0	-120.0	-150.0

O&M cost: operation and maintenance cost

IV. EXPERIMENTAL RESULTS

The data for this project were provided by a generating company in the U.S., which owns 11 generating units including three nuclear, three coal-fired, and five oil-fired units. Since nuclear units are almost always in service, they are excluded from unit commitment. An operating policy exists regarding the two peaking units, Oil 1–2, indicating that these two units should never be operated at the same time. Since these two units have identical characteristics, they can be treated as one unit in the generation scheduling process. As to which unit should be operated, the operating station has the right to decide. The most recent operating record provided by the generating company is for the week of January 26 to February 1. During that week, three units were under scheduled outage for repair or inspection as shown in Table I. The capacity of Coal 1 is reduced by 80 MW for two days (01/29 to 01/30). The remaining five units supply the system load demand and their characteristics are listed in Table II.

In order to study the efficiency of our algorithm, we set the parameters for the "end-test" of large and small iterations in simulated annealing as $p_0 = \infty$, $q_0 = \infty$, which correspond to a usual simulated annealing method. Let $T_0 = 1500$, $T_{i+1} = \gamma T_i$, $\gamma = 0.95$, $T_{\min} = 50$, and set the initial unit states to be off for the whole period. After 11 iterations, LRM converges to a near optimal solution of U.S.$ 803 846.71. If we set $T_{\min} = 10$ and maintain all other parameters unchanged, LRM converges in seven iterations with a near optimal solution of U.S.$ 624 029.06; obviously, this result is much better than the earlier

TABLE III
GENERATION AND LOAD (MW) FOR JANUARY 28, 2002

h	Coal 1 Cal.1	Coal 1 Cal.2	Coal 1 Act.	Coal 2 Cal.1	Coal 2 Cal.2	Coal 2 Act.	Coal 3 Cal.1	Coal 3 Cal.2
1	261.0	261.5	204.2	190.0	215.0	272.8	127.0	101.5
2	261.0	251.0	106.6	112.0	112.0	269.3	84.0	94.0
3	261.0	251.0	105.8	112.0	112.0	269.6	83.0	93.0
4	261.0	251.0	106.5	112.0	112.0	269.0	74.0	84.0
5	261.0	268.0	116.5	112.0	112.0	268.8	105.0	98.0
6	261.0	277.0	251.1	232.0	262.0	267.7	144.0	127.0
7	277.0	277.0	260.6	294.0	284.0	268.8	144.0	144.0
8	277.0	277.0	259.5	294.0	289.0	268.8	144.0	144.0
9	277.0	277.0	257.2	294.0	294.0	270.4	144.0	144.0
10	277.0	277.0	259.7	294.0	294.0	269.4	144.0	144.0
11	277.0	277.0	259.5	294.0	294.0	269.1	144.0	144.0
12	277.0	277.0	254.3	294.0	294.0	272.3	144.0	144.0
13	277.0	277.0	258.3	294.0	294.0	269.6	144.0	144.0
14	277.0	277.0	257.6	294.0	294.0	268.8	144.0	144.0
15	277.0	277.0	265.8	294.0	294.0	269.4	144.0	144.0
16	277.0	277.0	268.7	294.0	294.0	268.8	144.0	144.0
17	240.7	277.0	267.7	258.6	294.0	266.2	121.1	144.0
18	225.1	259.0	265.9	243.5	272.0	269.4	111.3	144.0
19	277.0	267.0	267.9	294.0	274.0	270.6	144.0	144.0
20	267.9	277.0	267.7	285.1	294.0	269.8	138.3	144.0
21	249.4	274.0	267.7	267.2	294.0	268.5	126.6	127.0
22	228.7	266.0	266.8	247.0	294.0	267.3	113.5	127.0
23	211.6	264.0	267.3	230.4	294.0	267.1	102.8	127.0
24	207.1	261.4	267.0	226.0	281.6	265.4	99.9	127.0

h	Coal 3 Act.	Oil 1-2 Cal.1	Oil 1-2 Cal.2	Oil 1-2 Act.	Oil 3 Cal.1	Oil 3 Cal.2	Oil 3 Act.	System Load
1	103.0	0.0	0.0	0.0	0.0	0.0	0.0	578.0
2	82.8	0.0	0.0	0.0	0.0	0.0	0.0	457.0
3	82.2	0.0	0.0	0.0	0.0	0.0	0.0	456.0
4	74.0	0.0	0.0	0.0	0.0	0.0	0.0	447.0
5	94.7	0.0	0.0	0.0	0.0	0.0	0.0	478.0
6	108.1	0.0	0.0	14.5	29.0	0.0	24.8	666.0
7	109.2	0.0	60.0	64.6	113.0	63.0	124.7	828.0
8	110.4	30.0	110.3	151.2	216.0	141.7	171.0	961.0
9	110.0	65.0	140.0	183.4	216.0	141.0	173.8	996.0
10	110.3	129.0	204.0	245.7	216.0	141.0	174.9	1060.0
11	110.0	169.0	244.0	279.2	216.0	141.0	182.3	1100.0
12	110.0	219.0	279.0	319.3	216.0	155.0	194.1	1150.0
13	110.6	268.0	305.2	361.3	216.0	178.8	199.1	1199.0
14	109.3	280.0	309.0	376.3	216.0	187.0	198.5	1211.0
15	110.5	268.0	297.0	376.5	216.0	187.0	176.3	1199.0
16	111.8	268.0	297.0	377.2	216.0	187.0	172.1	1199.0
17	111.8	262.0	206.0	264.3	179.7	141.0	152.0	1062.0
18	113.2	62.0	70.0	75.6	164.1	61.0	81.8	806.0
19	114.7	30.0	99.0	117.8	200.0	161.0	173.7	945.0
20	114.7	0.0	71.0	44.4	191.7	97.0	186.7	883.0
21	114.9	0.0	0.0	0.0	174.8	123.0	167.9	818.0
22	115.4	0.0	0.0	0.0	155.8	58.0	96.3	745.0
23	116.0	0.0	0.0	0.0	140.2	0.0	35.7	685.0
24	105.1	0.0	0.0	0.0	136.1	0.0	32.3	669.0

one. If we set $T_{\min} = 1$, LRM converges to a near optimal solution of U.S.$ 628 354.07; however, the CPU time increases to 16 s due to additional iterations of simulated annealing in each subproblem. These results imply that $T_{\min}$ is a crucial factor for the implementation of the algorithm; it influences not only the CPU time but also the quality of the solution. In some cases, simulated annealing cannot exit from local minima even though $T_{\min}$ is small.

Now, we use the improved simulated annealing to obtain the near optimal solution; after several experiments, we set $p_0 = 300$, $q_0 = 20$ for the given system. The values of p_0 and q_0 are determined empirically for a near optimal solution within a reasonable computation time. Let $T_0 = 1500$, $\gamma = 0.95$, $T_{\min} = 5$, and set the initial unit states to off for the whole period; we obtain the near optimal solution of U.S.$ 624 029.06 in seven iterations. If we set $T_{\min}$ to be 10 or 1, we obtain the same solution with a slightly higher CPU time. This experiment indicates that the values of $T_{\min}$, p_0, and q_0 play a major role in calculating the near optimal solution. Furthermore, the improved simulated annealing is able to provide more efficient results within a shorter period of time.

Table III presents the results for Monday, January 28, 2002 based on the proposed method (denoted as Cal.2). In addition, results based on [5] which considers fixed ramp-rate limits in unit commitment and economic dispatch as it incorporates the rotor fatigue effect (denoted as Cal.1), as well as the actual operating costs (denoted as Act.) are listed in Table III. The corresponding operating costs for actual, Cal.1, and Cal.2 cases are U.S.$ 629 124.69, U.S.$ 626 495.00, and U.S.$ 624 190.00, respectively. Although both calculation methods satisfy thermal stress limits, Cal.2 results in a lower operating cost because Cal.2 is not as constrained as Cal.1. In comparison, the results for Sunday (day before) and Tuesday (day after) are given as follows. On Sunday, hourly loads are smaller and operating costs for actual, Cal.1, and Cal.2 are U.S.$ 365 637.00, U.S.$ 364 583.00, and U.S.$ 363 031.03, respectively. Likewise, for Tuesday January 29, operating costs are higher since one of the base units (Coal 1) is on scheduled partial outage as listed in Table I. The Tuesday operating costs for actual, Cal.1, and Cal.2 cases are U.S.$ 528 920.63, U.S.$ 527 412.05, and U.S.$ 526 799.00, respectively.

The saving in operating cost, based on the application of the proposed method, will depend on the daily load profile and the amount of unit ramping required to follow the load variations. All experiments performed on the test system demonstrate that the proposed method (Cal.2) provides better and more realistic generation schedules.

V. Conclusion

In this paper, a new method for incorporating the thermal stress during the loading and unloading of steam-electric units is discussed. The method permits a dispatch at higher rates than fixed ramp rates without increasing the turbine rotor fatigue. The calculation of thermal stress is made from a thermal model of the rotor that does not require temperature sensors. The equations are validated against fatigue index curves provided by turbine manufacturers. The objective of the optimization process is to minimize the system operation cost, which includes those of energy production, start up, and thermal stress contribution to operating cost. The power dispatch based on thermal stress represents stress as a function of power output, which allows dispatchers to mimic the process that takes place in power stations, and essentially does not increase the dispatch equation complexity. LRM is used to decompose the original problem into several subproblems, each corresponding to a search for the best commitment under certain system operating conditions. This problem is solved by an improved simulated annealing method, which yields near-optimal and faster solutions than DP. In economic dispatch, thermal stress limits along with power generation limits are set as constraints, which is solved by LP. The experimental results demonstrate the effectiveness of the proposed method in providing reasonable constrained operation schedules for practical studies.

References

[1] J. Delson and M. Shahidehpour, "Linear programming applications to power system economics, operation and planning," *IEEE Trans. Power Syst.*, vol. 7, pp. 1155–1162, Aug. 1992.

[2] J. Arroyo and A. Conejo, "Optimal response of a thermal unit to an electricity spot market," *IEEE Trans. Power Syst.*, vol. 15, pp. 1098–1104, Aug. 2000.

[3] N. Sisworahardjo and A. El-Keib, "Unit commitment using the ant colony search algorithm," in *Proc. Large Eng. Syst. Conf. Power Eng.*, July 2002, pp. 2–6.

[4] C. Wang and M. Shahidehpour, "Effects of ramp-rate limits on unit commitment and economic dispatch," *IEEE Trans. Power Syst.*, vol. 8, pp. 1341–1350, Aug. 1993.

[5] ——, "Ramp-rate limits in unit commitment and economic dispatch incorporating rotor fatigue effect," *IEEE Trans. Power Syst.*, vol. 9, pp. 1539–1545, Aug. 1994.

[6] ——, "Optimal generation scheduling with ramping costs," *IEEE Trans. Power Syst.*, vol. 10, pp. 60–67, Feb. 1995.

[7] J. Delson, "Thermal stress computation for steam-electric generator dispatch," *IEEE Trans. Power Syst.*, vol. 9, pp. 120–127, Feb. 1994.

[8] M. Shahidehpour and M. Marwali, *Maintenance Scheduling in Restructured Power Systems*. Norwell, MA: Kluwer, 2000.

[9] M. Shahidehpour, H. Yamin, and Z. Li, *Market Operations in Electric Power Systems*. New York: Wiley, 2002.

[10] H. Ma and M. Shahidehpour, "Unit commitment with transmission security and voltage constraints," *IEEE Trans. Power Syst.*, vol. 14, pp. 757–764, May 1999.

[11] X. Guan, P. Luh, H. Yan, and J. Amalfi, "Short-term scheduling of thermal power systems," in *Proc. IEEE Power Ind. Comput. Applicat. Conf.*, 1991, pp. 120–126.

[12] A. I. Cohen, "Modeling unit ramp limitations in unit commitment," in *Proc. 10th Power Syst. Comput. Conf.*, Graz, Austria, Aug. 1990, pp. 1107–1114.

[13] F. Zhuang and F. Galiana, "Unit commitment by simulated annealing," *IEEE Trans. Power Syst.*, vol. 5, pp. 311–318, Feb. 1990.

[14] S. Kirkpatrick, G. D. Gelatt Jr, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, pp. 671–680, May 1983.

[15] P. J. M. van Laarhoven and E. H. L. Aarts, *Simulated Annealing: Theory and Application*. Norwell, MA: Kluwer, 1987.

Zuyi Li (M'01) received the B.S. degree in electrical engineering from the Shanghai Jiaotong University, Shanghai, China, in 1995, the M.S. degree electrical engineering from the Tsinghua University, Beijing, China, in 1998, and the Ph.D. degree electrical engineering from the Illinois Institute of Technology, Chicago, in 2002.

His research interests are focused on market operation of power systems.

Mohammad Shahidehpour (F'01) is a Professor in the Electrical and Computer Engineering Department and Director of the Center for Electric Power and Power Electronics at the Illinois Institute of Technology, Chicago.

He has written extensively on issues related to power systems restructuring and optimization.