A Dual Method for the Unit Commitment Problem

A. Möller and W. Römisch
Humboldt-Universität Berlin, Institut für Mathematik,
D-10099 Berlin, Germany

Abstract. In this paper a method for solving a mid-term unit commitment problem in a large-scale thermal power system is presented. This method is based on Lagrangian relaxation and uses a bundle method for solving the nonsmooth dual problem. Computational results are presented for systems up to 168 time periods (hours) and 100 units.

1 Introduction and Model

The unit commitment problem consists in determining a start-up/shut-down schedule and the corresponding production levels for each unit of a power system over a planning period so that the resulting total system costs are minimized. The schedules and the production levels have to satisfy demand and capacity constraints, and single unit constraints such as low and high generation limits, minimal up and down times as well as must-run and must-down time periods.

The mathematical model for our unit commitment problem is given by

$$\min_{(u,p)} f(u, p) := \sum_{t=0}^{T-1} \sum_{i=1}^{I} \left[ u_{t+1}^i FC_i(p_t^i) + K HC_i(u_{t+2}^i) + SUC_i(x_t^i, u_t^i) + HUC_i(x_t^i, u_t^i) \right]$$

subject to (for \(i = 1, \ldots, I\) and \(t = 1, \ldots, T\), respectively)

$$\sum_{i=1}^{I} u_{t+1}^i p_t^i = D_t \quad \text{(demand constraint)}$$

*This research is supported by a grant of the German Federal Ministry of Research and Technology (BMFT).*
\[
\sum_{i=1}^{I} u^t_i p^\max_i \geq D^t + R^t \quad \text{(capacity constraint)} \tag{1.3}
\]
\[
u^t_i p^\min_i \leq p^t_i \leq u^t_i p^\max_i \quad \text{ (unit capacity limits)} \tag{1.4}
\]
\[
u^t_i = (u^t_1, u^t_2) \in \{(0,0), (0,1), (1,0)\} \quad \text{ (decision set)} \tag{1.5}
\]
minimal up and down times \tag{1.6}
must on/must off constraints \tag{1.7}

where \( I \) is the number of units, \( T \) the number of time periods (hours), \( D^t \) and \( R^t \) the demand and reserve in time period \( t \), respectively, \( p^t_i \) the power production level of unit \( i \) in time period \( t \), \( u^t_i \) the decision variable describing whether the unit \( i \) in time period \( t \) has to be online \( (u^t_i = (1,0)) \), to be hot \( (u^t_i = (0,1)) \) or cool \( (u^t_i = (0,0)) \), \( x^t_i \) the state of unit \( i \) at time \( t \) indicating the status of the unit and how long the status has not been changed. Further, \( FC_i \), \( KH C_i \), \( SUC_i \) and \( SUC_i \) denote the fuel costs (as a quadratic function of \( p^t_i \)), the keeping hot costs, the start up costs and heat up costs, respectively.

\((1.1)\) - \((1.7)\) is a large-scale mixed-variable mathematical programming problem with nonlinear objective.

During the last decades numerous approaches for solving the unit commitment problem have been proposed. They are based on several optimization techniques such as heuristics and priority lists, dynamic programming, branch-and-bound methods, Benders decomposition, Lagrangian relaxation and combinations of them (cf. [1,6,8,9,11,15]). For a survey we refer to [14], where the authors come to the following conclusion: “A clear consensus is presently tending toward the Lagrangian relaxation approach over other methodologies.” Further applications of optimization techniques in power dispatch are considered in [4].

The idea of Lagrangian relaxation for unit commitment consists in including the demand and capacity constraints together with corresponding Lagrange multipliers into the objective function, so that the original problem decomposes into \( I \) independent single unit subproblems of lower dimension. By maximizing the dual function, the optimal Lagrange multipliers and a near-optimal solution of \((1.1)\) - \((1.7)\) are obtained.

In our algorithm outlined in Section 2 we use a bundle method to solve the dual problem. Compared to subgradient methods, bundle methods are based on a cutting-plane approximation for the objective associated with a bundle of information of function values and subgradients ([5,7,13]). The
computational results presented in Section 3 illustrate the performance of the proposed algorithm.

2 Lagrangian Relaxation and Solution Method

The Lagrangian of (1.1) with respect to the demand and capacity constraints is defined by

\[ L(u, p, \lambda, \mu) := f(u, p) + \sum_{t=0}^{T-1} \left[ \lambda^t \left( D^t - \sum_{i=1}^I u^T_i p^T_i \right) \right] + \sum_{t=0}^{T-1} \left[ \mu^t \left( D^t + R^t - \sum_{i=1}^I u^T_i p^\text{max}_i \right) \right], \]

where \( \lambda^t \) and \( \mu^t \) are certain Lagrange multipliers. The corresponding dual problem of (1.1) is

\[ \max_{(\lambda, \mu) \in \mathbb{R}^T \times \mathbb{R}^T_+} d(\lambda, \mu), \]

(2.1)

where

\[ d(\lambda, \mu) := \min_{(u, p)} \{ L(u, p, \lambda, \mu) \mid (u, p) \text{ satisfies (1.4) - (1.7)} \}. \]

(2.2)

For any given multipliers \( (\lambda, \mu) \in \mathbb{R}^T \times \mathbb{R}^T_+ \), the minimization problem on the right-hand side of equation (2.2) is called a Lagrangian relaxation of (1.1) - (1.7).

The function \( d : \mathbb{R}^T \times \mathbb{R}^T_+ \rightarrow \mathbb{R} \) is concave and a subgradient of \( d \) in \( (\lambda, \mu) \) is given by

\[
\begin{align*}
(D^0 - \sum_{i=1}^I u^0_i p^0_i, \ldots, D^{T-1} - \sum_{i=1}^I u^{T-1}_i p^{T-1}_i), \\
D^0 + R^0 - \sum_{i=1}^I u^0_i p^\text{max}_i, \ldots, D^{T-1} + R^{T-1} - \sum_{i=1}^I u^{T-1}_i p^{\text{max}}_i,
\end{align*}
\]

where \((u, p)\) minimizes \( L(u, p, \lambda, \mu) \) subject to (1.4) - (1.7) (see e.g. [11]). It holds

\[ \max_{(\lambda, \mu) \in \mathbb{R}^T \times \mathbb{R}^T_+} d(\lambda, \mu) \leq \min_{(u, p)} \{ f(u, p) \mid (u, p) \text{ satisfies (1.2) - (1.7)} \} \]

(weak duality theorem), but in general the equality is not satisfied (see e.g. [11]). An estimate of the relative duality gap is given by

\[ \frac{f^* - d^*}{d^*} \leq \text{const.} \frac{T + 1}{d^*}, \]
where \( f^* \) and \( d^* \) denote the optimal value of the original problem and its dual, respectively \([1,2]\). Since \( d^* \rightarrow \infty \) as \( I \rightarrow \infty \), the relative duality gap tends to zero if the number of units increases. Computational experiments indicate that the relative duality gap becomes also small for large \( T \). In general the primal variables obtained by solving the dual problem do not satisfy all demand and capacity constraints. Thus, a primal feasible solution has to be determined after the dual problem has been solved. The function \( d \) has the separability structure

\[
d(\lambda, \mu) = \sum_{i=1}^{l} d_i(\lambda, \mu) + \sum_{t=0}^{T-1} [\lambda^t D^t + \mu^t (D^t + R^t)],
\]

where \( d_i(\lambda, \mu) \) represents a single unit subproblem

\[
d_i(\lambda, \mu) = \min_{u_i} \left\{ \sum_{t=0}^{T-1} \min_{p^i_t \in [p^i_{\text{min}}, p^i_{\text{max}}]} \left\{ FC_i(p^i_t) - \lambda^t p^i_t \right\} \right.
+ K HC_i(u^i_2) + SUC_i(x^i_t, u^i_t)
+ H UC_i(x^i_t, u^i_t) - \mu^t u^i_{i1} p^i_{\text{max}} \right\}.
\]

In formula (2.3) the minimization with respect to \( p^i_t \) can be carried out explicitly and the minimization with respect to \( u_i \) subject to (1.5) - (1.7) is done by dynamic programming.

The solution strategy now consists in solving the dual maximization problem by nonsmooth optimization methods, which is followed by determining a primal feasible solution.

Our algorithm follows the general concept of \([15]\). A simplified flow chart is shown in figure 1 (see \([10]\)).

![Flow chart](chart.png)

Fig. 1

To initialize the Lagrange multiplier we generate a priority list based on the average fuel costs at the maximum power level. The units are considered in priority order and set online hourly if the capacity constraint has not yet
been satisfied or if it is required by other constraints of type (1.6) and (1.7). During this procedure we try to satisfy the inequality \( \sum_{t=1}^{T} u_t^i p_{t}^{m_i} \leq D_t \), too. After that we run the economic dispatch algorithm [3] and initialize the multipliers \( \lambda_t^i \) with the values of the Lagrange multipliers \( \lambda_{ED}^i \) obtained by the economic dispatch. The multipliers \( \mu_t^i \) are initialized by \( \mu_t^i = 0 \).

For solving the dual problem we use the bundle–trust (BT–)algorithm BTNCBC [12]. To illustrate the method we consider the problem

\[
\min_{x \in \mathbb{R}^m \times \mathbb{R}^m_n} f(x), \quad \text{where } f : \mathbb{R}^{m+n} \rightarrow \mathbb{R} \text{ is a convex function.}
\]

At the k-th iteration we have a sequence of iteration points \( x_k \), a set \( J_k \subseteq \{1, \ldots, k\} \) and a collection of auxiliary points \( y_i \). For each auxiliary point we also have a subgradient \( g_i \in \partial f(y_i) \) and the linearization error \( \alpha_k^i := f(x_k) - [f(y_i) + g_k^T(x_k - y_i)] \). The next auxiliary and iteration points are calculated as follows:

1. \( (v_k, d_k) := \arg\min_{(v,d) \in \mathbb{R}^m \times \mathbb{R}^{m+n}} \{ v + \frac{1}{2t_k} \| d \|^2 \mid v \geq g_k^T d - \alpha_k^i \forall i \in J_k, x_k + d \in \mathbb{R}^m \times \mathbb{R}^n_+ \} \)
2. If \( \frac{1}{t_k} \| d_k \| < \varepsilon \) and \( -\frac{1}{t_k} \| d_k \|^2 - v_k < \varepsilon \), then stop.

(In this case \( f(x_k) \leq f(x) + \varepsilon \| x - x_k \| + \varepsilon \forall x \in \mathbb{R}^m \times \mathbb{R}^n_+ \).)
3. If \( f(x_k + d_k) \) is “sufficiently smaller” than \( f(x_k) \), then either (a) enlarge \( t_k \) and goto step 1, or (b) \( x_{k+1} := y_{k+1} := x_k + d_k \) (serious step). If \( f(x_k + d_k) \) is “not sufficiently smaller” than \( f(x_k) \), then either (a) reduce \( t_k \) and goto step 1, or (b) \( x_{k+1} := x_k \) and \( y_{k+1} := x_k + d_k \) (null step).

A detailed description of the algorithm as well as convergence results are presented in [12,13].

As mentioned above the single unit problems are solved by dynamic programming.

The procedure to search for a reserve feasible solution (RFS) is essentially the same as in [15]. The idea consists in finding the time interval \( t \) for which the capacity constraint is mostly violated and then computing the smallest amount of necessary increase \( \Delta \mu_t^i \) for the multiplier \( \mu_t^i \) such that the solution of the new Lagrangian relaxation where \( \mu_t^i \) is replaced by \( \mu_t^i + \Delta \mu_t^i \) satisfies the t-th capacity constraint. This is done by determining the amount of necessary increase in \( \mu_t^i \) required to turn on an available unit that is originally offline in interval \( t \). When \( \mu_t^i \) is increased the commitment of all online units remains unchanged. This procedure is carried out recursively until the reserve constraint is satisfied for all intervals.

The 0-1 variables are now fixed and the generation levels \( p_t^i \) are adjusted by a final economic dispatch for each time interval. For this purpose we use the algorithm described in [3].
3 Computational Results

<table>
<thead>
<tr>
<th>number of periods</th>
<th>number of binary variables</th>
<th>number of real variables</th>
<th>termination parameter</th>
<th>number of BT-iterations</th>
<th>relative duality gap</th>
<th>CPU-time</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>6</td>
<td>108</td>
<td>72</td>
<td>1D-3</td>
<td>14</td>
<td>1.7 %</td>
</tr>
<tr>
<td>12</td>
<td>6</td>
<td>108</td>
<td>72</td>
<td>1D-4</td>
<td>35</td>
<td>1.1 %</td>
</tr>
<tr>
<td>12</td>
<td>6</td>
<td>108</td>
<td>72</td>
<td>1D-5</td>
<td>62</td>
<td>1.1 %</td>
</tr>
<tr>
<td>24</td>
<td>25</td>
<td>900</td>
<td>600</td>
<td>1D-3</td>
<td>8</td>
<td>0.4 %</td>
</tr>
<tr>
<td>24</td>
<td>25</td>
<td>900</td>
<td>600</td>
<td>1D-4</td>
<td>25</td>
<td>0.2 %</td>
</tr>
<tr>
<td>24</td>
<td>25</td>
<td>900</td>
<td>600</td>
<td>1D-5</td>
<td>56</td>
<td>0.2 %</td>
</tr>
<tr>
<td>96</td>
<td>50</td>
<td>7200</td>
<td>4800</td>
<td>1D-3</td>
<td>5</td>
<td>0.7 %</td>
</tr>
<tr>
<td>96</td>
<td>50</td>
<td>7200</td>
<td>4800</td>
<td>1D-4</td>
<td>23</td>
<td>0.1 %</td>
</tr>
<tr>
<td>96</td>
<td>50</td>
<td>7200</td>
<td>4800</td>
<td>1D-5</td>
<td>57</td>
<td>0.1 %</td>
</tr>
<tr>
<td>168</td>
<td>50</td>
<td>12600</td>
<td>8400</td>
<td>1D-3</td>
<td>3</td>
<td>0.08 %</td>
</tr>
<tr>
<td>168</td>
<td>50</td>
<td>12600</td>
<td>8400</td>
<td>1D-4</td>
<td>25</td>
<td>0.04 %</td>
</tr>
<tr>
<td>168</td>
<td>50</td>
<td>12600</td>
<td>8400</td>
<td>1D-5</td>
<td>50</td>
<td>0.05 %</td>
</tr>
<tr>
<td>168</td>
<td>100</td>
<td>25200</td>
<td>16800</td>
<td>1D-3</td>
<td>3</td>
<td>0.09 %</td>
</tr>
<tr>
<td>168</td>
<td>100</td>
<td>25200</td>
<td>16800</td>
<td>1D-4</td>
<td>13</td>
<td>0.01 %</td>
</tr>
<tr>
<td>168</td>
<td>100</td>
<td>25200</td>
<td>16800</td>
<td>1D-5</td>
<td>53</td>
<td>0.01 %</td>
</tr>
</tbody>
</table>

Tab. 1

The algorithm described in Section 2 is implemented in FORTRAN 77 and tested on a set of small- and mid-size unit commitment problems. Test runs are performed for problems of various dimension and for several values of the termination parameter of the BT-algorithm. Table 1 gives the corresponding number of BT-iterations, the computational relative duality gap and the CPU-times (on a HP-workstation). The termination criterion in step 2 of the BT-iteration is realized by multiplying the termination parameter from Table 1 with the estimate of the optimal primal function value obtained during the initialization procedure. It is known that the computing time for each Lagrangian relaxation given by (2.2) depends linearly on the number of periods and – assumed the average minimal up- and down-times are similar – on the number of units. This behaviour is not observed for the whole procedure due to the uncertain number of BT- and RFS-iterations. Compared to subgradient-type methods, which are used in [8, 9,15], an essential advantage of bundle-trust methods is the reliable stopping criterion (without calculating primal feasible solutions in between). Different termi-
nation criteria in the literature (bounds for the relative duality gap, maximal number of iterations) complicate the comparison of test results and CPU-times. However, the CPU-times of our test runs show that the algorithm is comparatively fast and, in particular, suitable for large unit commitment models.

Figure 2 contains the curves for the demand constraints, capacity constraints, and the sum of the lower and upper capacity limits of all online units, respectively, for the 100 unit and 168 hour problem.

Acknowledgement. The authors wish to thank H. Balzer and G. Schwarzbach (Vereinigte Energiewerke AG, Berlin) for the fruitful cooperation.

References


