2-1987

Fast Thermal Generation Rescheduling

F. Eugenio Villaseca
Cleveland State University, f.villaseca@csuohio.edu

B. Fardanesh
Cleveland State University

Follow this and additional works at: https://engagedscholarship.csuohio.edu/enece_facpub

Part of the Power and Energy Commons

How does access to this work benefit you? Let us know!

Publisher's Statement
© 1987 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Original Citation

Repository Citation
https://engagedscholarship.csuohio.edu/enece_facpub/101

This Article is brought to you for free and open access by the Electrical Engineering & Computer Science Department at EngagedScholarship@CSU. It has been accepted for inclusion in Electrical Engineering & Computer Science Faculty Publications by an authorized administrator of EngagedScholarship@CSU. For more information, please contact library.es@csuohio.edu.
FAST THERMAL GENERATION RESCHEDULING

F. Eugenio Villaseca, Member, IEEE  Behruz Fardanesh, Student Member, IEEE

Department of Electrical Engineering
Cleveland State University

Abstract - A new dynamic programming algorithm for fast rescheduling thermal generation is presented. The savings in computational times are brought about by the introduction of two new techniques: the variable truncation dynamic programming and the limitation of the solution space to be searched. Several examples on a 20 machine system are used to illustrate the application of the algorithm and to show that optimal solutions are obtained at significantly reduced computational times.

INTRODUCTION

A non-optimal generation schedule is most likely to offset the savings that are expected through economic dispatch efforts, thereby resulting in an inefficient system operation. Even if an optimal production schedule for the day is available, rescheduling may become necessary due to unexpected deviations from the forecasted load and changes in the availability of system components. This updating and adjustment of the generation schedule should be performed promptly, keeping the total operating cost at a minimum. Resolving the entire scheduling problem again will result in a new optimal generation schedule, but it may not become available in time for the system operator to effect the proper changes, forcing manual changes which may lead the system to operate on a course far from the optimal.

The generation scheduling problem, also referred to as unit commitment or predispatch [1], has been the subject of considerable discussion in the power literature [2,3,4].

Generally, solution techniques in present unit commitment routines can be categorized either as heuristic or as mathematical programming methods. Heuristic methods [5,6,7] start with an initial feasible solution and use a logical approach to reduce the operating costs in successive refinements. In each iteration, possible savings resulting from the advancement or delay of startup or shutdown of units is sought. When no further improvements are possible, the process ends. Although heuristic methods are flexible and allow for the consideration of actual system operation constraints, there is no guarantee that their solutions are optimal.

Mathematical programming techniques such as dynamic programming [8-16], linear and non-linear programming [17,18], Lagrangian relaxation [19,20] and branch-and-bound methods [19,21] have been proposed to solve different formulations of the problem.

In dynamic programming based solutions, for each time interval (usually an hour), different combinations of units which render feasible solutions to the problem are considered. Then economic dispatch is performed for each combination. Some dynamic programming formulations also yield a solution to the economic dispatch problem [9,12]. Others have implemented approximate economic dispatch subroutines [13] in order to save computation time. System losses are usually neglected. Some studies include system losses in the total forecasted demand [20].

Application of non-linear programming techniques is limited only to greatly simplified problems [8]. Consideration of the pertinent constraints results in a non-convex and non-differentiable problem which is not amenable to non-linear programming solutions. Linear programming methods do not generally present satisfactory solutions due to the level of approximation introduced by linearizing the problem [15,17].

More recent studies obtain a solution for the dual of the original problem and combine this approach with the branch-and-bound technique and solve a mixed-integer programming problem [19,20]. This approach, also referred to as Lagrangian relaxation, exploits the separable structure of the problem to keep computation time and memory requirements for large systems in a manageable range, and results in near optimal solutions [20].

All these studies consider the generation rescheduling problem as an entirely new scheduling problem, thus ignoring valuable information available in the previous schedule.

Here, a new algorithm for the fast rescheduling of thermal generation is presented and several examples on a 20 machine system are used to illustrate the application of the algorithm and to show that optimal solutions are obtained at significantly reduced computational times. The savings in computational times are the result of the introduction of two new techniques: the variable truncation dynamic programming and the limitation of the solution space to be searched. Both techniques are discussed in the following sections.

DYNAMIC PROGRAMMING FORMULATION

Consider a system with n generating units over a time horizon of M hours. The time horizon is divided into smaller time intervals, normally of one hour each.

An integrated hourly demand is used over each hourly interval. For each hourly interval, all unit combinations that satisfy the constraints are considered. A single unit combination is a set of units that satisfies all the required constraints. The total generation cost of all unit combinations for the entire horizon constitutes a cost function which is to be minimized. Mathematically, this is expressed as
Minimize $Z(a) = \sum_{t} \left[ \sum_{j \in Y_t} C(p) + \sum_{j \in Y_t} T(p, p) \right]$

A glossary defining all symbols used is listed in the Appendix.

This function is subject to the following set of constraints and considerations:

a. **Power balance constraint**

$$\sum_{j \in Y_t} p = P_t + P_L$$

b. **Spinning reserve constraints**

$$\sum_{j \in Y_t} \min(p - P, R) > Q$$

$$\sum_{j \in Y_t} \min(p - P, R) > A$$ for $t = 1, 2, \ldots, L$

c. **Unit capacity limits**

$$p_j \leq p \leq p_{jmax}$$

d. **Unit startup costs**

$$C_{SU} = C_{CS}(1 - e^{-Ut}) + C_{F}$$

In addition to the above constraints, the following must also be considered:

e. **Unit minimum up and minimum downtimes**

f. **Must run units**

g. **Units assigned to fixed generation status**

h. **Crew constraints on plants**

i. **Changes in unit status and characteristics**

The problem is decomposed into time intervals which constitute the stages of the problem. At each stage the problem is completely defined by a combination of generating units which satisfies all the constraints and the operating cost associated with that combination when its units are loaded to share the generation most economically. This is defined as the state of the problem. Since the most economic loading of the generating units, or economic dispatch, is inherent in the process of evaluation of the production cost for each combination of on-line units, in the following discussion "combination" and "state" are used interchangeably. Therefore, whenever "state" is replaced by "combination" the associated minimum cost of production is implicit.

At each stage the problem might have several allowable states. The transition from a unique state in a stage to another state at another stage occurs by effecting feasible controls which are the decisions made in regard to startup or shutdown of generating units. A given combination of units at each state is dependent on a state in the previous stage and the applied controls in the transition. The problem can be solved recursively using forward dynamic programming as follows:

$$\text{CUMC}(T, K) = \text{MIN}[\text{PODST}(T, K) + \text{TCOST}(T-1, T, K) + \text{CUMC}(T-1, 1)]$$

The hourly cost characteristics of thermal units are represented by a quadratic equation of the form:

$$C_j(p_t) = a_j(p_t) + b_j p_j + c_j$$

The traditional economic dispatch problem for any interval $t$ is formulated as:

$$\min \sum_{j \in Y_t} C(p_j)$$

subject to

$$\sum_{j \in Y_t} p_j = P_t + P_L$$

$$\sum_{j \in Y_t} \min(p - P, R) > A$$ for $t = 1, 2, \ldots, L$

The solution to this problem is well known in the literature [2, 22] and is obtained by using the classical Lagrange multipliers. To minimize the cost of supplying a given load $P_L$ with a given set of on-line generating units we must have

$$\lambda_t = 3C_{Cj} \sum_{j \in Y_t} p_j < P_j^{max}$$

The total transmission loss $P_L$ is generally represented as a quadratic function of the generation levels of units in the system using the system $B$s-constants [22].

The distribution of reserve resulting from traditional economic dispatch may not be practical because of the limitations in load pickup capability or response rate of the units. If all the reserve capacity is on one of the on-line units, load pickup capability of the system will be a minimum. For maximum response rate the reserve must be distributed among the greatest number of units.

To achieve a proper distribution of the reserve, a technique proposed by Wood [16] is used. In effect, this adds a set of unit maximum reserve constraints.

**VARIABLE TRUNCATION DYNAMIC PROGRAMMING**

Pang and Chen [13] introduced the concept of truncated dynamic programming for optimal scheduling of thermal units. In their method, priorities are assigned to different units and a unit selection list is formed. A higher priority corresponds to a more economic unit. After formulation of the unit priority list, the minimum number of units with highest priority that satisfy load plus reserve requirements are assigned to must-run status. Then a search window is placed around the lowest priority unit assigned to must-run status, which could include a number of units directly above and a number of units directly below this unit in the priority list. The number of units in the search window is specified by the user. Furthermore, a control over the computational requirements is suggested by limiting the number of feasible states saved at each hourly interval. However, once the level of truncation is chosen, it is uniformly applied to all hourly intervals in the horizon. In other words, the truncation level is fixed over the entire horizon.

The variable truncation method proposed here dif-
fers from the fixed truncation approach in that it is based on the forecasted load profile and two user-specified parameters. These parameters define a high and a low value for the number of strategies that may be saved, which correspond to the number of states that will be considered for transition from any time interval to the next. A criterion which identifies the periods where a high or low number of strategies must be saved is defined based on the following consideration.

It is apparent that during the periods of little or no ramping (low rate of increase or decrease of the demand), the set of states (and the associated strategies) which contains the optimal strategy, will often consist of the minimum cumulative cost state and a few of the states with cumulative costs closest to the minimum. In other words, during periods of little or no ramping, a small subset of the set of all strategies needs to be saved and examined during the solution process. The opposite is true for the periods of high ramping. This seems quite logical considering the fact that it is more likely to have generating units started up or shut down during the periods of high ramping.

The algorithm developed to identify the high ramping periods is simple. For each hourly interval a "demand change step" which is the absolute value of the change in demand from the present hour to the next hour is obtained. Then, an average "demand change step" is calculated for the entire horizon. If the demand change step from any interval t to interval t+1 is larger than the average "demand change step", the specified high value is used for the number of strategies that must be saved in time interval t. Otherwise the specified low value is used.

Although the above discussion is centered on the number of saved strategies at each interval, the same algorithm may be used to effect variable search window sizes [13] and increase the possibility that the combinations that constitute the optimal states be included in the searched solution space.

No direct procedure for the appropriate selection of the high and low parameter values can be recommended at this point, particularly since they appear to be system dependent. An experimental procedure for their determination is as follows.

Testing several fixed truncation levels (i.e. same high and low number of saved strategies) will identify the lowest value that most often results in the optimum solution. This can be selected as the high value for the number of saved strategies. Then with the selected high value and experimentation with several values for the low number of saved strategies, the proper low value that most often renders the optimal solution is determined.

Although the success of chosen parameters in obtaining the optimum solution cannot be guaranteed, a diverse class of studies indicated that once they have been identified their use consistently renders satisfactory solutions to the problem. The stability of these parameters justifies the preliminary efforts required for their experimental determination.

LIMITATION OF THE SEARCH SPACE

Another technique that can be used to further reduce the computational effort is to limit the solution space to be searched based on a previously obtained optimal schedule.

Consider two distinct situations, which may arise when changes to a schedule need to be introduced: extra capacity needs to be scheduled or scheduled capacity needs to be reduced. Clearly, both actions may be required during the horizon of the schedule, but for convenience, they will be treated separately in the discussion below.

If extra capacity needs to be scheduled, it is highly probable that the original set of on-line units plus off-line units, if necessary, will satisfy the requirements. In other words, it is highly likely that no unit in the original set of on-line units will be replaced by any off-line available unit. This was confirmed through a detailed study of the solutions obtained for many different cases. This low probability of unit replacement can be taken advantage of to effect a controlled reduction in the computational efforts.

To allow the possibility of unit replacement, the solution space to be searched will be limited only to the available off-line units plus a selected number of the units originally scheduled to be on-line. These selected units referred to as the transitional units will be examined for possible replacement. The transitional units to be considered for possible replacement are well identified in the available schedule, and are those that have been most recently started up or shut down.

Let IND = 0, 1, 2, ..., represent the index measuring the increase in computational effort. If IND = 0, the search space includes only the off-line but available units. That is, no unit from the original set is to be replaced. If IND = 1, 2, ..., the shutdown of each transitional unit will be advanced by 1, 2, ... hours and the startup of each transitional unit will be delayed by 1, 2, ... hours.

Now consider the case when the scheduled capacity needs to be reduced. The solution space to be searched is limited to the units scheduled to be on-line plus a selected number of off-line available units.

If IND = 0, the solution space to be searched includes only the on-line units. That is, only the on-line units are considered for possible shutdown. If IND = 1, 2, ..., the shutdown of each transitional unit will be delayed by 1, 2, ... hours and the startup of each transitional unit will be advanced by 1, 2, ... hours.

The recursive search algorithm to minimize the cost increase due to the new requirements at hour T with combination X is:

\[ \text{ACC}(T,K) = \min \{ \Delta C(T,K) + \text{ACC}(T-1,1:T,K) + \text{ACC}(T-1,1) \} \]

**COMPUTER SIMULATION AND RESULTS**

To examine the merits of the proposed algorithm, three computer programs were developed in FORTRAN and run on an IBM-3081 [23]: the full dynamic programming (FDP) code, the variable truncation dynamic programming (VTDP) code and the proposed algorithm for fast generation rescheduling (FAGRR). The FDP program was used to establish the optimal solutions of the problems studied. The VTDP was used to test the variable truncation technique, and to provide the schedule upon which the FAGRR algorithm would operate.

The power system utilized consists of 20 thermal units whose characteristics are given in Table 1. The parameters for the cost and start-up functions are
listed in Table 2. An hourly load profile for a 24-hour period is specified in Table 3. This is used as a base case to which alterations are introduced to force rescheduling. The system reserve requirements are assumed to be 600 [MW].

To demonstrate the savings in CPU time obtained by using variable truncation versus fixed truncation, three cases are presented in Table 4. Case 0 is the selected base case described above. Case 1 is identical to case 0, except that the hourly demand is increased by 100 MW. Case 2, on the other hand, corresponds to a 100 MW decrease in the hourly demand.

Although VTDP could be used for rescheduling whenever needed, it is not efficient since it solves the entire problem again. That is, it ignores valuable information contained in the previous schedule. FAGER, on the other hand, by using variable truncation and the limitation of the search space is a much more efficient tool, as illustrated below.

<table>
<thead>
<tr>
<th>CASE</th>
<th>Total Costs [$$]</th>
<th>CPU [sec]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>186,040</td>
<td>47.22</td>
</tr>
<tr>
<td>1</td>
<td>192,089</td>
<td>39.59</td>
</tr>
<tr>
<td>2</td>
<td>180,049</td>
<td>24.45</td>
</tr>
</tbody>
</table>

Table 5 shows the results obtained by FAGER for cases 1 and 2, and several others. All these cases use the case 0 schedule as reference. Most FAGER's CPU savings are obtained by the technique of limiting the search space. To show this, the results of FAGER, using no truncation at all, are also included. VTDP results are shown for comparison.

<table>
<thead>
<tr>
<th>CASE</th>
<th>Total Costs [$$]</th>
<th>CPU [sec]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>192,089</td>
<td>39.59</td>
</tr>
<tr>
<td>1</td>
<td>192,089</td>
<td>24.45</td>
</tr>
<tr>
<td>2</td>
<td>192,089</td>
<td>24.45</td>
</tr>
</tbody>
</table>

Eight case studies are presented below, where cases 1 through 7 arbitrarily assume no crew constraints and zero unit shut-down costs. However, case 8 includes both constraints. Transmission losses are neglected in all cases.
Before the significance of the case studies is discussed, it is important to note at the outset that, for the rescheduling cases shown in Table 5, use of FAGER results in CPU time reductions ranging from 77 to 91% compared to VTDP.

Cases 3 and 4 involve a change in the scheduled reserve. In case 3, the reserve capacity requirement is increased by 100MW over case 0, that is from 600MW to 700MW, and case 4 is for a 100MW reduction in the scheduled reserve capacity. If there is a choice between importing 100MW of power over the tie-lines from a neighboring company at a certain cost, and bringing more units on-line, a fast evaluation of choices will identify the best operating strategy. Also, sharing of reserve capacity among neighboring utilities is of particular interest since an inherent reserve capacity in excess of the scheduled level of reserve can be used profitably. Furthermore, several transactions for exchange of power over the tie-lines may be in effect at any time and the amount of power to be sold or purchased is determined by the operators. The ability of the system operator in making prompt decisions to keep the total production costs to a minimum is of great importance and it can be considerably enhanced by using FAGER.

Another case of interest is a unit outage or when an available unit becomes unavailable. In case 5, it is assumed that unit number 15 is not available and the schedule is updated using FAGER and VTDP.

As case 6, the peak shaving problem when a system has pumped-hydro or compressed-air storage capability is considered. To evaluate a certain peak-shaving pattern, FAGER can be used as a fast means for computing the savings resulting from a specific change in the load pattern. It should be noted that in this case, both increase and decrease of the demand are occurring over the considered horizon as opposed to cases 1 and 2 where considered only increase and decrease, respectively, in the demand. The chosen peak-shaving pattern is a reduction in generation of 200MW from hours 1 to 4 and from 22 to 24, and an increase in generation of 200MW from hour 11 to 18, on the base case load profile.

In addition to determining the best pumping and storage policy, it is possible using FAGER to control the load level and negotiate delay or advance of service to certain loads when direct load control or other load management options are available.

In case 7, a 400MW reduction in demand from hour 1 to 6 is considered. Using IND=0 does not result in the optimum solution and the next larger search range corresponding to IND=1 is needed to find it. An important observation which adds to the merits of FAGER is that in most cases IND=0 does result in the optimum solution and larger search spaces are not necessary. However, large step changes in demand or generation such as in case 7 are more likely to need larger search ranges.

Case 8 is the last case considered. In this case, the crew constraints which were neglected in all previous cases are considered and the number of unit startups at each plant to one, in any hourly interval. Also, a cost of $50 per unit shutdown is assumed. Using VTDP, the new base case with the crew constraints and with non-zero shutdown costs, a 24-hour schedule is obtained with a minimum total cost of $186,324. This total cost is $284 higher than the previous base case (case 0). This schedule can be updated and adjusted to accommodate for any required change in the generation schedule. Here, a 100MW reduction in the scheduled reserve capacity is considered. The results are presented in Table 5.

Another interesting result which adds to the versatility of FAGER is that it can also be used to optimally adjust a given optimal schedule by taking into account an ignored constraint or a new constraint. For example, using the results of the original base case (case 0) and considering the crew constraints and non-zero shutdown costs, FAGER with IND=4 results in the optimum schedule with the total cost of $185,929. This run uses high and low truncation levels of 15 and 4 and requires 7.97 seconds of CPU time.

It is important to note that all the studies performed, namely changes in demand, changes in scheduled reserve, unit outages, and peak shaving, all used the same high and low parameters in the variable truncation procedure. The stability of these system dependent parameters for such a diverse class of studies justifies the effort required for their experimental identification.

CONCLUSIONS

A new approach to the solution of the optimal thermal generation rescheduling problem has been presented. The FAGER algorithm, using a 20 machine system, is shown to provide a fast means of updating a previously obtained optimal schedule. The savings in CPU times are the result of the introduction of two new techniques: the variable truncation dynamic programming and the limitation of the solution space to be searched. The experimental determination of the high and low parameters for the variable truncation procedure and the search space index is discussed. These are shown to be stable under a diverse class of studies and the effort required for their determination seems justified.

APPENDIX

Glossary of terms:

\[ z = \text{Total cost of generation for the considered horizon.} \]
\[ M = \text{number of hourly intervals in the horizon.} \]
\[ t = \text{power set of } n \]
\[ c = \text{cost function for unit } j \text{ in time interval } t \text{ [$/h].} \]
\[ T = \text{transitional costs for unit } j \text{ (i.e., startup and shutdown costs) in going from a state in time interval } t-1 \text{ to a state in time interval } t \text{ [$/h].} \]
\[ a = \{y_1, y_2, ..., y^n\} = \text{set of unit combinations for the entire horizon.} \]
\[ y \in \gamma(n) = \text{a single unit combination in time interval } t. \]

\[ n = \{1, ..., n\} = \text{set of all units.} \]
\[ \delta(n) = \{\delta | \delta \in n\} = \text{power set of } n. \]
\[ n = \text{total number of units.} \]
\( P_t \) = total system demand in time interval \( t \) [MW],
\( D_t \) = total system transmission losses in time interval \( t \) [MW],
\( P_{j\max}^t \) = maximum generation level (capacity) of unit \( j \) in time interval \( t \) [MW],
\( R_{j\max}^t \) = maximum spinning reserve allowed on unit \( j \) in time interval \( t \) [MW],
\( Q_t \) = system reserve requirement in time interval \( t \) [MW],
\( A_t \) = area spinning reserve requirement in time interval \( t \) [MW],
\( \delta_N \) = set of units in area \( N \) (\( \delta_N \) \( n \)),
\( L \) = number of areas in the system,
\( P_{j\max}^t \) = maximum capacity of unit \( j \) in time interval \( t \),
\( P_{j\min}^t \) = minimum capacity of unit \( j \) in time interval \( t \),
\( C_{bu} \) = unit start up cost [$],
\( C_{cs} \) = cost of cold start [$],
\( v \) = Cooling rate [h\(^{-1}\)],
\( C_f \) = fixed costs involved in startup (e.g., maintenance costs and crew expenses).
\( C_j(P_j) \) = Hourly cost of operating unit \( j \) at \( P_j \) megawatts in the interval \( t \) [$/h],
\( a_j, b_j, d_j \) = Constants in cost function for unit \( j \).
\( P_j \) = Power output level of unit \( j \) in time interval \( t \) [MW],
\( \lambda_t \) = incremental cost of received power in time interval \( t \) [$/MWh],
\( \Delta P^t \) = incremental transmission loss for unit \( j \) in time interval \( t \) [$/MWh],
\( \Delta c_j \) = incremental cost of power generated by unit \( j \), in time interval \( t \) [$/MWh].
\( \text{CUMC}(T,K) \) = minimum total cost to arrive at state \( K \) in hour \( T \),
\( \text{PCOST}(T,K) \) = production cost in hour \( T \) with state \( K \),
\( \text{TCOST}(T-1,I:T,K) \) = transitional (start up and shut down) costs from state \( I \) at hour \( T-1 \) to state \( K \) at hour \( T \).


Post-print prepared by MSL Academic Endeavors, the imprint of the Michael Schwartz Library at Cleveland State University, 2014