MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Factored Markov Decision Process Models for Stochastic Unit Commitment

Daniel Nikovski, Weihong Zhang

TR2010-083 October 2010

Abstract

In this paper, we consider stochastic unit commitment problems where power demand and the output of some generators are random variables. We represent stochastic unit commitment problems in the form of factored Markov decision process models, and propose an approximate algorithm to solve such models. By incorporating a risk component in the cost function, the algorithm can achieve a balance between the operational costs and blackout risks. The proposed algorithm outperformed existing non-stochastic approaches on several problem instances, resulting in both lower risks and operational costs.

2010 IEEE Conference on Innovative Technologies

Factored Markov Decision Process Models for Stochastic Unit Commitment

Daniel Nikovski and Weihong Zhang, {nikovski,wzhang}@merl.com, phone 617-621-7510
Mitsubishi Electric Research Laboratories, 201 Broadway, Cambridge, MA 02139, USA, fax 617-621-7550

***Abstract*—In this paper, we consider stochastic unit commitment problems where power demand and the output of some generators are random variables. We represent stochastic unit commitment problems in the form of factored Markov decision process models, and propose an approximate algorithm to solve such models. By incorporating a risk component in the cost function, the algorithm can achieve a balance between the operational costs and blackout risks. The proposed algorithm outperformed existing non-stochastic approaches on several problem instances, resulting in both lower risks and operational costs.**

I. Introduction

Given a collection of generating units, demands, and operational constraints over a time horizon, the problem of unit commitment is concerned with finding the optimal schedules and amounts of generated power for each generator. In the past, the generators have typically been assumed to be fully controllable (e.g. fossil-burned, nuclear), and the future electrical power demand has been assumed to be completely known. Under these assumptions, various combinatorial optimization methods have been proposed, including ones based on dynamic programming, Lagrange relaxation, mixed integer programming, etc. [1].

In reality, though, these assumptions are not correct. Future power demand can rarely be predicted with errors less than 2% on prediction horizons of 24 hours or longer, so demand is in fact a random variable with at least that much standard deviation. Moreover, the rapidly increasing penetration of renewable power sources, such as photovoltaic panels and wind turbines, makes these assumptions less and less reasonable. The highly variable and intermittent output of renewable power sources affects significantly the net amount of power that has to be generated by means of controllable generators, by influencing both the demand and supply side. On the demand side, the power generated by renewables owned by customers changes the amount such customers would request from the electrical utility in order to supplement their needs. On the supply side, the amount generated by the renewables owned by the utility would similarly change the amount it needs to generate in order to meet demand. The effects on the demand and supply sides do not balance each other, but rather act in sync to increase the volatility of power demand for controllable generators, and thus exacerbate the planning problem for utilities.

One traditional way to plan for deviations from expected demand and supply has been to include a safety margin of extra capacity (for example, 3%) to be committed for production. This results in operating more and/or larger units than are necessary to meet expected demand. This approach is largely heuristic, and is not likely to work in the future, when renewable energy sources become even more widespread.

An alternative approach is to recognize that the uncertainty in power demand and generator supply makes the decision problem a stochastic one. A stochastic operational scheduler computes a schedule that is robust to future variations of supply and demand, and provides a safety margin implicitly, by planning for all possible contingencies. One significant difficulty associated with this approach has been how to represent all such possible contingencies, and how to plan for them. One proposal organizes all future possible realizations of the system (called scenarios) as a tree of scenario bundles [2]. However, this model for representing stochasticity is limited to only the few scenarios included in it, whereas in a practical system the future evolution can be realized in an infinite number of ways. Our work aims to expand this approach by improving the probabilistic modeling of system evolution.

We propose a method for finding the optimal conditional operational schedule of a set of power generators under stochastic demand for electrical power and stochastic output of some generators. Unlike traditional operational schedules, which are fixed in advance, a conditional operational schedule depends on the future state of the observable random variables (demand and output), and can result in different actual schedules depending on the observed outcomes for these variables. The scheduler explicitly balances the operational cost of electricity generation with the risk of not being able to meet future electricity demand. We represent the stochastic dynamics of the components of the system as a factored Markov decision process (MDP) model, and propose efficient approximate algorithms for computing suitable conditional operational schedules.

II. Factored Markov Decision Processes for Stochastic Unit Commitment Problems

A. Stochastic Unit Commitment

A stochastic unit commitment (UC) problem is an optimization problem under uncertainty. Let N be the number of available controllable generator units, and T be the total length of the planning horizon, in suitable units (typically, one hour). The objective function is presented in Equation II.1, where $u_t^i \in \{0,1\}$ represents the commitment status (on or off) for unit i at step t, x_t^i represents the number of time steps that unit i has been on/off, and d_t are the realizations of the random demand D_t, $1 \le t \le T$, assumed to be coming from a known stochastic process. Similarly, if there are K uncontrollable generators, we assume that the realizations y_t^k of their random output amounts Y_t^k also come from known stochastic processes. The configuration of all controllable units at time t is $u_t = [u_t^1, u_t^2, \ldots, u_t^N]$, and respectively the operating times of all controllable units at that time are $x_t = [x_t^1, x_t^2, \ldots, x_t^N]$. The vector of realizations of all uncontrollable generators at time t is $y_t = [y_t^1, y_t^2, \ldots, y_t^K]$.

$$\begin{aligned} J^* &= \min_{u_1,u_2,\ldots,u_T} \mathbf{E}_{u_0,x_0,y_t,d_t} \{ \textstyle\sum_{t=0}^{T-1} [\sum_{i=1}^{N} f_i(x_t^i, u_t^i, y_t, d_t) \\ &+ \textstyle\sum_{i=1}^{N} h_i(x_t^i, u_t^i, u_{t+1}^i) + g_t(u_t, y_t, d_t)] \} \end{aligned} \tag{II.1}$$

Here $f_i(x_t^i, u_t^i, y_t, d_t)$ denotes the operating cost of operating unit i in configuration u_t^i and state x_t^i for one time step in order to meet demand d_t when the uncontrollable generators output electricity amount y_t. The function $h_i(x_t^i, u_t^i, u_{t+1}^i)$ denotes the cost of switching to configuration u_{t+1}^i at the end of the step. The third cost component, $g_t(u_t, y_t, d_t)$, denotes the equivalent cost of the risk of not being able to meet demand d_t under output of uncontrollable generators y_t with the chosen configuration of all units u_t. This cost is proportional to the probability that the total capacity of the committed units in u_t plus what the uncontrollable generators produce (y_t) is less than the demand d_t:

$$g_t(u_t, y_t, d_t) = \alpha Pr(\sum_{i=1}^{N} u_t^i cap^i + \sum_{k=1}^{K} y_t^k < d_t),$$

where cap^i is the maximal generation capacity of unit i. A suitably chosen proportionality coefficient α specifies the relative preference between minimizing operating cost and risk of failure to meet demand. In Equation II.1, $\mathbf{E}\{\}$ denotes the expectation operator with regard to the initial configuration u_0, operational time x_0, and the realizations of demand d_t and output y_t. By adding the operating cost and risk compensation cost together, the objective function represents a trade-off between fuel costs and risk.

A UC problem has to observe several constraints in minimizing the total cost. The load balance constraint states that the total generation must be equal to the demand d_t at any time step. If p_t^i is the generation of unit i at hour t, then

$$\sum_{i=1}^{N} p_t^i u_t^i + \sum_{k=1}^{K} y_t^k - d_t = 0, \text{ for } \quad t = 1, 2, \ldots, T. \tag{II.2}$$

In most markets, a specific amount of spinning reserve is explicitly required by regulators. The positive excess spinning reserve constraint indicates that the total committed spinning capacity should be greater than the sum of the load and the required spinning reserve. Let r_t^{req} be the required spinning reserve. Then:

$$\sum_{i=1}^{N} cap^i u_t^i + \sum_{k=1}^{K} \bar{Y}_t^k - d_t - r_t^{req} \ge 0, \tag{II.3}$$

where $\bar{Y}_t^k$ is the expected output of uncontrollable generator k at time t. Given an existing commitment status u_{t-1}^i, operation time x_{t-1}^i, and new commitment status u_t^i, the new operational time x_t^i can calculated as in Equation $(II.4)$ where T_i^{cl} is the "cold start" time of unit i, T_i^{dn} is the minimum down time of unit i, and T_i^{up} is the minimum up time of unit i [3].

$$x_t^i = \begin{cases} 1 & \text{if } -T_t^{cl} \le x_{t-1}^i \le -T_i^{dn} \text{ and} \\ & u_t^i = 1 \text{ (start up)} \\ x_{t-1}^i + 1 & \text{if } 1 \le x_{t-1}^i \le T_i^{up} - 1 \\ & \text{(up and must stay up)} \\ T_i^{up} & \text{if } x_{t-1}^i = T_i^{up} \text{ and } u_t^i = 1 \\ & \text{(up and available to shut down)} \\ -1 & \text{if } x_{t-1}^i = T_i^{up} \text{ and } u_t^i = 0 \\ & \text{(shutting down)} \\ x_{t-1}^i - 1 & \text{if } -T^{dn} + 1 \le x_{t-1}^i \le -1 \\ & \text{(down and must stay down)} \\ & \text{or } -T_i^{cl} + 1 \le x_{t-1}^i \le -T_i^{dn} \text{ and} \\ & u_t^i = 0 \\ & \text{(down and available to start up)} \\ -T_i^{cl} & \text{if } x_{t-1}^i = -T_i^{cl} \text{ and } u_t^i = 0 \end{cases} \tag{II.4}$$

Other constraints that this paper considers include the mini-

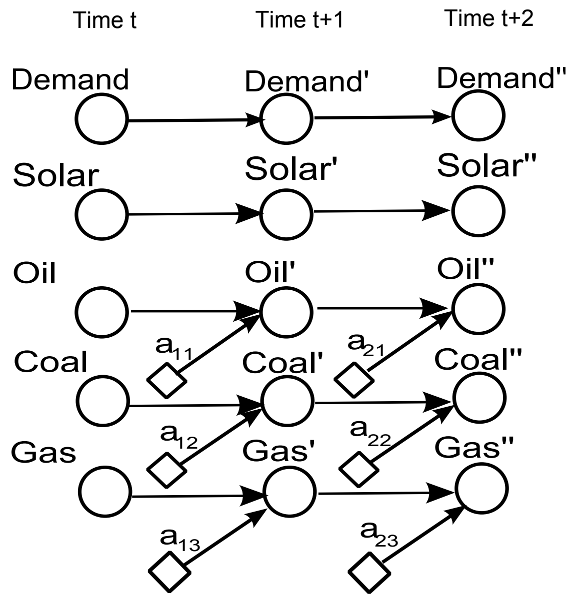

Figure II.1. DBN for a power generation problem with three controllable and one uncontrollable power generators.

mum and maximum generation capacity of the units, minimum up and down time constraints, and unit availability constraints. Additional constraints, such as maximal up/down times, can be accommodated by suitable modifications to Equation $(II.4)$.

B. Stochastic UC as a Factored Markov Decision Process

We propose to represent a power generation system consisting of multiple generators of the type described above by means of a *factored Markov decision process* (fMDP), and find the optimal conditional operational schedule by means of approximate dynamic programming [4]. The fMDP is usually expressed graphically as a *dynamic Bayesian network* (DBN). A DBN consists of circles that represent random variables, diamonds that represent decision variables, and directed edges connecting the circles and diamonds that represent the statistical dependence between the corresponding variables. A random variable is conditioned probabilistically on all variables from which a directed edge is connected to it. Conversely, lack of edges between variables signifies statistical independence between them. When dealing with a time-dependent system, each time period (e.g. one hour) is represented by its own set of random variables. Three time slices of the DBN for an example stochastic unit commitment problem with four generators, one of which uncontrollable, are shown in Fig. II.1.

In Fig. II.1, one random, real-valued, and uncontrollable variable represents the power demand. Another random, real-valued, and uncontrollable variable represents the output of a photovoltaic generator. Three conventional controllable generators are shown, too; their discrete variables x_t^i take on $T^{cl}+T^{up}$ different values, and represent the operational time of the respective generator. Three decision variables (shown as diamonds) represent the individual decisions u_t^i to turn on/off the corresponding generators, and thus commit them for power production. The probabilistic evolution of the system is described by local conditional probability tables for each variable, where the conditional dependence is defined only on the parents of that variable in the graph of the DBN. Thus, the DBN serves as a compact representation of a large Markov decision process whose state space is exponentially large in the number of states of the individual variables over which it is factored.

In order to specify a factored MDP, the state, action, and transition model for each individual variable must be defined. This is done differently for controllable variables (generators) and uncontrollable generators and the demand variable, as described below. The reward/cost model is described jointly for all variables in the fMDP.

State: The complete state of a controllable generation unit i is described by its commitment status u_t^i and its operational time x_t^i. It is denoted by (u_t^i, x_t^i). The state of all controllable units at time t is the cross product of the state of all units at that step. It is denoted by (u_t, x_t) where u_t (respectively, x_t) is the vector of u_t^i (respectively, x_t^i) over all units i. The state of the variable representing demand D_t is a discrete variable that corresponds to the deviation of demand from its expected value $\bar{D}_t$ at time t, normalized by the standard deviation σ_t of demand at that time. A suitable number of discrete bins for that variable can be chosen, for example with width one standard deviation. The state of a random variable Y_t^k representing the output of an uncontrollable generator can be represented similarly, as a deviation from its expected value $\bar{Y}_t^k$, suitably normalized with respect to its standard deviation.

The state of the system at one step is the cross product of the state of all units (controllable and uncontrollable) and the demand variable. Therefore, the system state at time t is represented as (u_t, x_t, Y_t, D_t). A particular state with concrete instantiation for the demand $D_t = d_t$ and the output of non-controllable generators $Y_t^k = y_t^k$, $1 \leq k \leq K$, will be denoted by (u_t, x_t, y_t, d_t).

Action: Actions change both the commitment status and operational time of the units. An action with respect to a unit i at time t is its intended commitment status u_t^i at the next step. An action with respect to all units, denoted by u_t, is the vector of such settings over all units. Note that a state of a unit indicates its commitment status and operational time at the current step, while an action indicates their intended status at the next step. Note also that given a state (u_t^i, x_t^i) and an action u_{t+1}^i, the operational time x_{t+1}^i of a controllable generator is uniquely determined by Equation $(II.4)$.

State transition: Given a state (u_t^i, x_t^i) and an action u_{t+1}^i, the operational time x_{t+1}^i of a controllable generator is uniquely determined by Equation $(II.4)$. So, for controllable units, state (u_t^i, x_t^i) transitions deterministically to (u_{t+1}^i, x_{t+1}^i) according to that equation.

For the demand variable D, we assume that the next demand D_{t+1} depends only on the current demand D_t (Markovian property of the underlying stochastic process) with transition probability $Pr(D_{t+1} = d_{t+1}|D_t = d_t)$. For the uncontrollable generators, we make similar assumptions that Y_{t+1}^k depends only on Y_t^k, with probability $Pr(Y_{t+1}^k = y_{t+1}^k|Y_t^k = y_t^k)$. For most generators, these are reasonable assumptions that capture the major part of the transition dynamics of their outputs. For example, whether a photovoltaic unit would produce less than one standard deviation below its expected value depends mostly on whether it was producing that much in the previous time step; this is the typical situation under cloudy conditions, which tend to persist over fairly long periods of time. Conversely, its output is only very weakly coupled to the output of other uncontrollable generators, for example a wind turbine. These transition probabilities can be estimated either from statistical data, or by means of discretizing a suitable continuous stochastic Markov process, such as the auto-regressive process of order 1 (AR(1) process).

We can also compute the joint transition probability for the entire system $Pr(u_{t+1}, x_{t+1}, y_{t+1}, d_{t+1}|u_t, x_t, y_t, d_t)$, from the transition probabilities of the individual random variables as

$$Pr(y_{t+1}, d_{t+1}|y_{t+1}, d_t) = Pr(d_{t+1}|d_t)\prod_{k=1}^{K} Pr(y_{t+1}^k|y_t^k), \quad \text{(II.5)}$$

if u_t, x_t, u_{t+1} and x_{t+1} conform to Equation $(II.4)$, or zero otherwise (impossible transition). It can be observed that although the MDP has a very large joint state space, its transition structure is very sparse.

Cost: Unlike transition probabilities, which can be specified separately for each individual variable, the transition cost is specified for the entire MDP. Given a joint MDP state (u_t, x_t, y_t, d_t) and an action u_{t+1}, the immediate one-step cost $c(u_t, x_t, u_{t+1}, y_t, d_t)$ is computed as

$$\begin{aligned} c(u_t, x_t, u_{t+1}, y_t, d_t) \quad &= \quad \textstyle\sum_{i=1}^{N} f_i(x_t^i, u_t^i, y_t, d_t) + \sum_{i=1}^{N} h_i(x_t^i, u_t^i, u_{t+1}^i) \\ &+ \quad g_t(u_t, y_t, d_t) \end{aligned} \quad \text{(II.6)}$$

where the switching costs $h_i(x_t^i, u_t^i, u_{t+1}^i)$ and risk cost $g_t(u_t, y_t, d_t)$ are computed as described above, and the fuel costs $f_i(x_t^i, u_t^i, y_t, d_t)$ are computed by solving the following economic dispatch problem:

minimize $\sum_i F_i(p_t^i)$ subject to the generation limits for all generators and the load balance constraint for this particular realization of the uncontrollable variables y_t and demand d_t:

$$\sum_{i=1}^{N} u_t^i p_t^i + \sum_{k=1}^{K} y_t^k - d_t = 0$$

where $F_i(p_t^i)$ is the cost of producing p_t^i units of electricity by generator i; typically, this function is quadratic in p_t^i, and the economic dispatch problem can be solved by means of quadratic programming. The objective of economic dispatch is to find the optimal generation amounts p_t^i of the committed units so that the cost of generation is minimized for a specific realization of the random variables. After the optimal generation amounts $[p_t^1, p_t^2, \ldots, p_t^N]$ are found, the individual generation costs can be calculated as $f_i(x_t^i, u_t^i, y_t, d_t) = F_i(p_t^i)$, $1 \leq i \leq N$.

Given such an MDP, we can define its cost-to-go functions J_t for each step t and each joint state of the MDP. For the terminal step T, when no further decisions will be made, $J_T(u_T, x_T, y_T, d_T) = 0$.

For all other steps, the cost-to-go function $J_t(u_t, x_t, y_t, d_t)$ is defined iteratively by means of a Bellman equation, as follows [5]:

$$\begin{aligned} &J_t(u_t, x_t, y_t, d_t) = \min_{u_{t+1}}\{c(u_t, x_t, u_{t+1}, y_t, d_t) \\ &+ \textstyle\sum_{d_{t+1}, y_{t+1}} Pr(d_{t+1}, y_{t+1}|d_t, y_t) J_{t+1}(u_{t+1}, x_{t+1}, y_{t+1}, d_{t+1})\} \end{aligned} \quad \text{(II.7)}$$

Note that the transition probabilities $Pr(d_{t+1}, y_{t+1}|d_t, y_t)$ are factored conveniently as per Equation II.5, due to the conditional independence relations in the DBN of the MDP. The cost-to-go function $J_0(u_0, x_0, y_0, d_0)$ of the initial state of the generators and demand would then correspond to the minimum in Equation II.1: $J* = J_0(u_0, x_0, y_0, d_0)$, which is the minimal expected cost of the conditional operation scheduler. This cost can be found be computing the costs-to-go of all states in the MDP.

Furthermore, if these costs are computed and stored, the optimal decision $u_{t+1} = \pi_t(u_t, x_t, y_t, d_t)$ for time step t and state

(u_t, x_t, y_t, d_t) can be identified as the one that minimizes the right-hand side of the Bellman equation II.7:

$$\pi_t(u_t, x_t, y_t, d_t) = argmin_{u_{t+1}}\{c(u_t, x_t, u_{t+1}, y_t, d_t) + \sum_{d_{t+1}, y_{t+1}} Pr(d_{t+1}, y_{t+1}|d_t, y_t) J_{t+1}(u_{t+1}, x_{t+1}, y_{t+1}, d_{t+1})\} \tag{II.8}$$

This policy is conditioned upon the current realizations of the random variables y_t and d_t, so it represents a conditional scheduler. By observing the outcomes y_t and d_t for each consecutive time step, different actual operating schedules will be obtained.

III. Solving fMDP Models with Aggregated Net Demand

The objective of solving the stochastic unit commitment problem represented by the fMDP is to find the optimal policy that maps the states of the fMDP onto the decision variables that signify which generators will be turned on/off in the next period, where optimality is defined in terms of jointly minimizing production cost and risk of failure. The straightforward method of solving fMDPs is to expand the factored state and solve the resulting flat MDP by means of dynamic programming, applying equation II.7 repeatedly, starting from the terminal step and proceeding backwards to the first step [5]. However, for most practical problems, e.g. when $T^{cl} = T^{up}$, the number of generators $N = 20$, the number of one-hour time periods $T = 24$, the expanded MDP will have $|X| = T(T^{cl} + T^{up})^N = 24 \cdot 10^{20}$ distinct states for the controllable generators only, and would be impossible to solve.

One practical simplification of the problem is to aggregate the output of the uncontrollable generators Y_t into the demand variable, by subtracting these outputs from the total demand D_t to arrive at the net demand D'_t. If all uncontrollable random variables are Gaussian processes, then D'_t is a Gaussian process, too, with expected value (mean) $\bar{D}'_t$ and variance σ_t for each time period t. Henceforth, we will assume that D_t denotes the net demand. For planning purposes, the net demand D_t can be computed by subtracting the expected values $\bar{Y}_t$ at the time of planning ($t = 0$). When executing the policy, the actually observed realizations y_t at time t can be used to estimate the distribution of the random variable D_{t+1}, so that the estimates of the transition probabilities $Pr(d_{t+1}|d_t, y_t)$ will in fact be based on y_t, when determining the optimal configuration u_{t+1} by means of Equation II.8.

Another computational simplification of the problem is to reduce the size of the MDP in a reasonable manner. Intuitively, if forecasts for the values of the continuous random variables D_t and Y_t are known in advance, and the assumption that these are Gaussian processes holds true, most of the configurations of the generators u_t at time t would be irrelevant to satisfying demand at that time. Some of them will have capacities too low to meet demand, and others will use unnecessarily many generators to meet demand economically. By considering only configurations u_t of the controllable part of the MDP whose maximal committed capacity (MCC) is close to the expected net demand $\bar{D}_t$, we can drastically reduce the size of the space of the MDP.

A practical way of identifying such suitable configurations is to run a fast deterministic algorithm for unit commitment for several possible values of target reserve β such that the target demand is $(1+\beta)\bar{D}$. Suitable schedules S_β are identified for each β, and the generator configurations u_t present in S_β are included in the reduced state space of an approximate solver, which essentially switches between individual segments from multiple schedules S_β, depending on the time evolution of power demand and uncontrollable generators.

Hence, the fundamental idea of the solution algorithm is to identify suitable configurations for representative demands, and then use them to produce schedules for any possible realization of demand. We use an AND/OR tree ([6]) to represent all selected configurations of the generators and possible realizations of future demand. The AND/OR tree is then used for planning for any demand instances.

A. Building the AND/OR tree

An AND/OR tree has two types of nodes — AND nodes and OR nodes. An AND/OR tree is a tree where (1) its root is an AND node, (2) it has alternating levels of AND and OR nodes, and (3) its terminal nodes are AND nodes [6]. An AND/OR tree is shown in Figure III.1 where the AND/OR nodes are respectively in rectangular/circular shapes. Note that in this case the outputs of the uncontrollable generators Y_t have been aggregated into the net demand variable D_t, and are not included in the AND/OR tree.

An AND node for the UC problem is associated with a system state (u_t, x_t, d_t) at time step t, whereas an OR node is associated with the action u_t at that time. The root node corresponds to the initial state of the UC system. The values of the nodes are evaluated bottom-up. For an OR node u_{t+1}, if its parent AND node is (u_t, x_t, d_t) and its children (AND) nodes are $\{(u_{t+1}, x_{t+1}, d_{t+1})|d_{t+1}\}$, then the value of the OR node is evaluated as

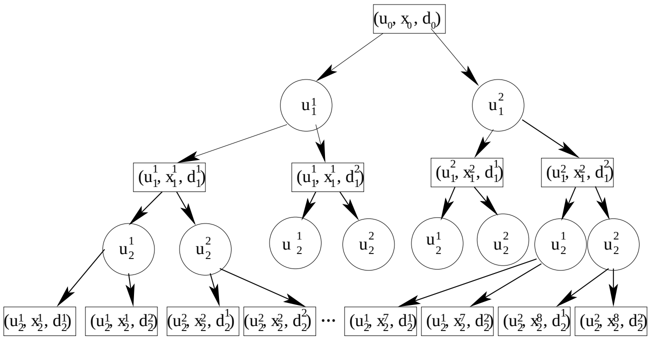

Figure III.1. An AND/OR tree example

$$V_t(u_{t+1}|u_t,x_t,d_t) = c(u_t,x_t,u_{t+1},d_t) + \sum_{d_{t+1}} p(d_{t+1}|d_t)V_{t+1}(u_{t+1},x_{t+1},d_{t+1}) \quad \text{(III.1)}$$

Note that the notation $V_t(u_{t+1}|u_t,x_t,d_t)$ means that the value of OR node is conditional on its parent AND node. For an AND node (u_t,x_t,d_t), its value $V_t(u_t,x_t,d_t)$ is evaluated as follows:

$$V_t(u_t,x_t,d_t) = \begin{cases} c(u_T,x_T,u_T,d_T), & \text{if } t=T \\ \min_{u_{t+1}} V_t(u_{t+1}|u_t,x_t,d_t) & \text{otherwise} \end{cases} \quad \text{(III.2)}$$

Note that the minimization in $\min_{u_t}$ is over all children OR nodes u_{t+1}, and that no configuration switching cost is incurred at the last step, since the continuation of the schedule at that time is yet unknown.

1) The root node (u_0,x_0,d_0) is the initial state of the system, and is the first node of the tree, at level 0. Since the levels correspond to the time steps in a UC problem, we use steps to refer to levels. The rest of the tree is built over time steps: the AND nodes at Step t and an OR node are used to build AND nodes at Step $t+1$. The OR nodes correspond to the suitable configuration selected earlier. For a node (u_t,x_t,d_t), we use every candidate configuration to produce AND nodes at the next time step. Let the OR node be u_{t+1}. The status of the units at next time step is (u_{t+1},x_{t+1}). Since the demand at $t+1$ is uncertain, we generate an AND node for every possible demand, suitably discretized into several discrete variables. So, the set of next AND nodes is $\{(u_{t+1},x_{t+1},d_{t+1}|d_{t+1})\}$. The process of tree constuction repeats until completion at Step T.
2) The values of the AND and OR nodes in the tree are computed by means of Equations (III.1) and (III.2), applied in a bottom-up manner. In evaluating the non-terminal AND nodes, there must be an OR node that achieves the minimum in Equation (III.2). The action represented by that OR node is the best action of the system state represented by the parent AND node.

B. Evaluating MDP Policies

Once a policy has been computed and stored in the AND/OR tree, we adopt a sampling approach to evaluate its operational cost and risk under future random demand D. For this purpose, we draw a suitable number of samples $d = [d_1, d_2, \ldots, d_T]$ from the demand variable D (e.g., 1000 samples). For each sample, we start from the root of the tree and execute the actions specified by the tree. Such an execution results in a path in the tree. Specifically, an execution path is a sequence of system states and actions $\{(u_0,x_0,d_0),u_1,(u_1,x_1,d_1),\ldots,u_T,(u_T,x_T,d_T)\}$ that are prescribed by the initial system state, the AND/OR tree, and the demand realization $d = [d_1, d_2, \ldots, d_T]$. The cost of a path can be accessed by solving the economic dispatch problem for each step, given the prescribed configurations u_t, while its risk can be calculated using the committed capacity u_t and the realization of demand d_t. The overall risks and costs are the average across the paths associated with the demand samples.

The most expensive part of the algorithm is the building and evaluation of the AND/OR tree, because it is exponential in the planning horizon T, where the base of the expeonent is the number of discrete levels of discretization for the demand variable D_t. However, when an AND node is added to the tree, a feasibility check is performed first: a node is added only when it meets all temporal constraints plus the demand and load constraints, and the economic dispatch associated with the node has a feasible solution.

IV. Experimental results

We experimented with the proposed method on a test problem adopted from [3], extended with the introduction of uncertainty in the demand. The standard deviation of demand was assumed to be 2% of expected demand: $\sigma_t = 0.02\bar{D}_t$. No uncontrollable generators were used, so the net demand is equal to the total demand. The approximate algorithm from the previous section was implemented and compared against two existing algorithms: one of them was based on a priority list ([1]), and the other one was the decommitment algorithm proposed in [3]. Our results showed that the approximate solution method provides a good balance between generation cost and risk of failure to meet demand. We performed experiments on two UC examples: one with 4 units, and another one with 20 units. We were able to calculate the truly optimal MDP solution for the 4-unit UC example, so we were able to investigate the accuracy of our approximation scheme on that problem, too. The experiments were performed on a computer with Intel Core 2 Duo E6600 CPU (2.40GHz). The algorithm was implemented in MATLAB 7.9.0 (R2009b).

A. Experimental Conditions

The generation cost of a committed unit i at time t is computed as a quadratic function of the produced amount of power by the unit: $f_i(x_t^i, u_t^i, d_t) = c_0^i + c_1^i p_i^t + c_2^i (p_t^i)^2$. The unit switching and start-up cost is expressed as $h(x_t^i, u_t^i, u_{t+1}^i) = tcst_i + bcst_i(1 - \exp(-\gamma x_t^i))$, if $u_t^i = 0$ and $u_{t+1}^i = 1$, and zero otherwise. In the start-up cost, the fixed component $tcst_i$ represents the cost of starting generator i, while the second term $bcst_i$ represents the cost of starting the boiler and varies exponentially with the length of the time that the unit has been off.

Under a Gaussian assumption for demand ($D_t \sim N(\bar{D}_t, \sigma_t^2)$), the risk compensation cost $g_t(u_t, d_t)$ is given by

$$\alpha' \cdot C_{FSO} \cdot \int_{\sum_i u_t^i cap_i}^{\infty} \frac{1}{\sqrt{2\pi\sigma_t^2}} exp(-\frac{(D - \bar{D}_t)^2}{2\sigma_t^2}) \cdot dD$$

where α' is the proportionality constant, C_{FSO} is the full system operating costs (the cost of the system in which all units are turned on and generate according to their maximum capacity), and the integral is the failure probability (risk). Failure happens when the actual demand D is greater than the Maximum Committed Capacity (MCC) $\sum_i cap_i u_t^i$ of all operating units. By increasing the constant α, the weight of the risk component in the objective function is increased, thus favoring configurations with higher MCC, at the expense of a higher operational cost for running such configurations.

B. A 4-unit example

The decision horizon of the 4-unit UC problem was 24 hours. The coefficients $tcst_i$ and $bcst_i$ of the start-up costs for the four units were [200,2000;500,20000;100,700;44,100]. The fuel cost coefficients $[c_0, c_1, c_2]$ for the four units were [0.00211,16.51,02.7; 0.00063,21.05,1313.6; 0.00712,22.26,371.0; 0.00413,25.92,660.8], in chosen cost units. The minimum up and minimum down times were $[3,3,2,2]$ and $[4,4,3,3]$. The minimum and maximum capacities were $[10,10,10,10]$ and $[100,90,80,60]$, here and henceforward, in chosen power units. The expected demand vector was $\bar{D}$ =[105,85,65,140,100,105,125,145,165,185,205,245,265,285, 200,140,100,105,125,145,165,185,205,225]. The spinning reserves were 15 for all time steps. The initial operational times were $x_0 = [5, -5, 5, -5]$.

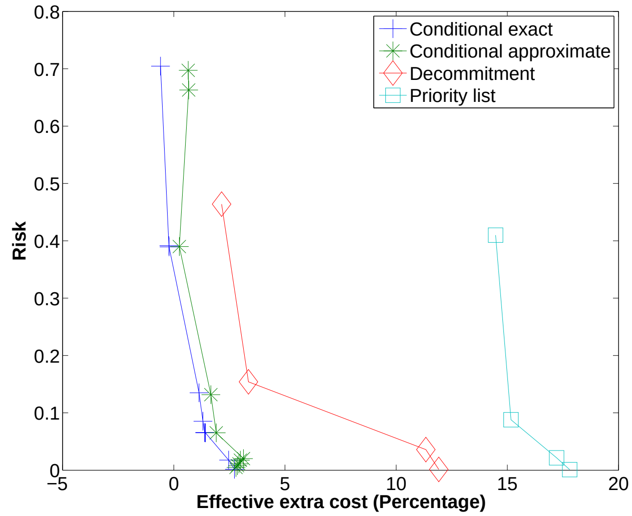

Figure IV.1. Performances of the algorithms on a 4-unit problem

The risk versus cost curves for various methods are presented in Figure IV.1. "Conditional exact" refers to the algorithm that solves the MDP exactly, i.e., all Bellman backups (Equation II.8) were performed. "Conditional approximate" refers to the algorithm proposed in the previous section. In the figure, the horizontal axis is the percentage of the extra operational cost with respect to a reference operational cost, taken to be the lowest experimentally obtained operational cost for any scheduler on this problem. For this problem instance, it can be seen that the solution of the proposed algorithm is very close to optimality (the conditional exact solution), and the algorithm outperforms significantly both the priority list and the decommitment algorithms in balancing operational costs and risks. For the lowest levels of risk, which are probably close to the desired cost/risk trade-off point of an actual generation system, the loss of optimality is less than 1%, whereas the gain in costs with respect to deterministic schedulers is greater than 9%.

C. 20-unit example

In this experiment we used all 20 generators described in [3]. The expected demand vector was [2133.3, 2133.3, 2066.7, 2066.7, 2133.3, 2133.3, 2266.7, 2400.0, 2400.0, 2400.0, 2333.3, 2200.0, 2133.3, 2133.3 ,2200.0, 2266.7, 2400.0,

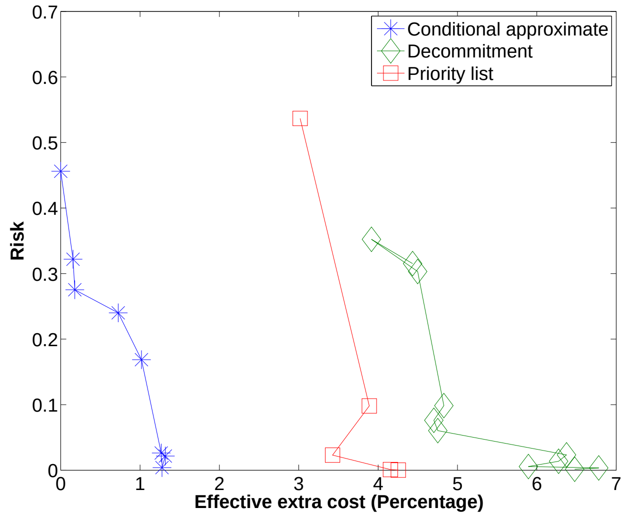

Figure IV.2. Performances of the algorithms on a 20-unit problem

2400.0, 2400.0, 2400.0, 2333.3, 2200.0, 2200.0, 2066.7]. The required spinning reserve was 133.3. It was no longer possible to find the truly optimal conditional schedules, but it is possible to compare the performance of the conditional approximate, priority list, and decommitment algorithms (Figure IV.2). Again, the results show that the proposed novel algorithm uniformly achieved a much better risk/cost balance than the priority list and the decommitment approaches, with operational cost savings around 4% for the lowest levels of risk.

V. Conclusion

We have described a general method for representing the stochastic dynamics of power generation systems under multiple sources of uncertainty such as variable power demand and intermittent renewable energy sources, and have introduced a class of conditional generation schedules where the unit commitment decisions are conditioned upon the state of observable random variables. The proposed factored Markov decision process models represented in the form of dynamic Bayesian networks are compact and are also easy to specify, maintain, and extend with new power sources.

We have also proposed a concrete algorithm for finding such conditional operational schedules for power generation that depend on a single random variable — the net demand that aggregates in itself all sources of randomness. The algorithm focuses on small subsets of all possible configurations of generators in order to compute the schedule efficiently. Experimental results suggest that the resulting conditional schedules are close to the truly optimal ones, and provide a much better trade-off between generation cost and risk of failure to meet demand than two known non-stochastic unit commitment algorithms that compute fixed schedules.

In the proposed solution algorithm, we use AND/OR trees to represent, find, and evaluate the optimal conditional schedule. However, this algorithm is by no means the only possible way to solve stochastic generation problems represented by means of fMDPs and DBNs. In future work, we plan to investigate other solution methods based on approximate dynamic programming that could result in much better computational complexity. Furthermore, the current solution aggregates the variability of all stochastic variables into the net demand to the controllable power generators, for the sake of computational efficiency. This simplifies the planning problem, because the branching in the AND/OR tree is based only on that single variable. However, even higher efficiency might be possible if the conditional schedule is conditioned on the values of each individual stochastic component. This would significantly increase the complexity of the planning process, and would depend critically on finding more computationally efficient solution methods for the underlying fMDP models.

The formulation of the fMDP described in the paper assumes that all generators would switch to their intended configurations u_t^i without fail. This allows us to use the decision variables u_t^i as components of the state of the system, thus simplifying the planning process. If the possibility of equipment failure must be taken into account, the actual configuration U_t^i of the generators should be included as a random state variable in the DBN, and its probabilistic dependence on the intended configuration u_t^i can be modeled according to the failure probabilities of individual generators. Such an extension is completely compatible with the proposed modeling formalism of factored Markov decision processes.

References

[1] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation, and Control*. New York: Wiley-Interscience, 1996.

[2] S. Takriti, J. R. Birge, and E. Long, "A stochastic model of the unit commitment problem," *IEEE Transactions on Power Systems*, pp. 1497–1508, 1996.

[3] C. Li, R. B. Johnson, and A. J. Svoboda, "A new unit commitment method," *IEEE Transactions on Power Systems*, pp. 113–119, 1997.

[4] C. Boutilier, R. Dearden, and M. Goldszmidt, "Stochastic dynamic programming with factored representations," *Artificial Intelligence*, pp. 49–107, 2000.

[5] M. L. Puterman, *Markov Decision Processes—Discrete Stochastic Dynamic Programming*. New York, NY: John Wiley & Sons, Inc., 1994.

[6] A. Martelli and U. Montanari, "Additive AND/OR graphs," in *Proceedings of the Third International Joint Conference on Artificial intelligence*, 1973, pp. 1–11.