CHAPTER II
LITERATURE REVIEW ON GENERATION SCHEDULING

2.1 Introduction

The determination of most economical generation scheduling, which satisfies all operating constraints, is one of the major problems in power system operation. In most utilities, electric power is dominantly generated from thermal units. In the current situation in which the fuel price is tend to continuously increases, the economic consequence of operating schedule of thermal units becomes much more important.

Scheduling generating units over a short-term horizon is a complex, large scale, mixed-integer, and non-linear optimization problem. Generation scheduling problem in the short-term horizon sometimes also called as unit commitment (UC) problem [39, 47], which has been the subject of a large number of publications [26, 47, 48] for decades.

In a conventional vertically integrated structure, the UC problem concerns with the determination of committed units for supplying the forecasted demand and reserve at minimal total cost while satisfying the prevailing constraints. Many papers which utilizes several methods for obtaining the solution under this old structure have been reported in [28, 48, 49]. This problem became much more complex since deregulation in the electricity industry introduced in 1990's. Several methods has also been proposed for solving the UC problem in the deregulated environment as reported in [26, 50]. One of the models of deregulation is the combination of vertically integrated system, called as the original system in this dissertation, with bilateral contracts. This model is mainly adopted by developing countries by allowing some GENCOs to directly sell power to consumers via bilateral transactions.

This chapter reviews several aspects which are usually considered in the unit commitment problem. In the beginning, the formulation of UC problem based on both conventional and MILP method are presented. Then, issues which have strong correlation with the UC problem, e.g. demand uncertainty and reliability, are discussed. The review of optimization techniques currently available to solve UC problem are also presented. Finally, literature study on decision analysis and its previous application is described.

2.2 Conventional Unit Commitment Formulation

The UC problem can be formulated as minimization of the objective function subjected to variety of system and unit constraints. The objective function is usually non-convex and representing the total cost of producing electricity. The system constraints comprise demand and spinning reserve as of the classical UC problem in vertically integrated structure. The unit constraints include physical limitation of generator, e.g. minimum and maximum delivered power, minimum up and down-time constraints.

2.2.1 Objective Function

The objective of UC problem in a vertically integrated structure is to minimize the total generation cost which is the sum of the production and the start-up costs of all committed units over a scheduling period. This objective function can be formulated as

$$\min F\left(P_i^t, U_i^t\right)=\sum_{t=1}^{T}\sum_{i=1}^{I}\left[F_i\left(P_i^t\right)+ST_i^t\left(1-U_i^{t-1}\right)\right]U_i^t \quad \text{.............................. (2.1)}$$

where

P_i^t : output power of unit i at time t (MW),

U_i^t : status of unit i at time t (1=on; 0=off),

$F_i(P_i^t)$: fuel cost function of unit i when its output power is P_i^t (\$),

ST_i^t : start-up cost of unit i at time t (\$),

I : number of generating units,

T : number of scheduling periods (hours).

Generally, the fuel cost of a thermal generating unit can be modeled as a quadratic function [39] as follows:

$$F_i\left(P_i\right)=a_i+b_iP_i+c_iP_i^2 \quad \$/hr \quad \text{.. (2.2)}$$

where

a_i, b_i, and c_i : the cost coefficients for the generating unit i,

P_i : generated power of generating unit i.

The UC objective function in (2.1) is subjected to some system and unit constraints which will be explained in the following sub-section.

2.2.2 System Constraints

1) Demand requirements: Since the short-term generation scheduling is an activity to determine the committed units to meet the forecasted demand, hence the total amount of generated power has to balance with the forecasted load demand in every time interval (P_{load}^t). Generally, this constraint can be formulated as

$$P_{load}^t - \sum_{i=1}^{I} P_i^t U_i^t = 0, \forall t \in [1, T] \quad (2.3)$$

2) Spinning reserve requirements: Spinning reserve is the term used to describe the total amount of generation available form all units synchronized on the system minus the load demand and losses being supplied. The UC schedule should guarantee that sufficient generating spinning reserve is available to ensure the reliable and secure operation of the power system during emergency conditions. This means that in the event of a contingency where one or more units are disconnected from the system, there must be enough reserve in the remaining units to make up for the generating outage. The spinning reserve requirement can be represented as follows:

$$P_{load}^t + SR^t - \sum_{i=1}^{I} P_{\max,i} \leq 0, \forall t \in [1, T] \quad (2.4)$$

where SR^t and $P_{max,i}$ is the spinning reserve requirement at time t and maximum output capacity of unit i respectively.

Practically in most utilities, the spinning reserve requirement is deterministically set as equal to or grater than the maximum capacity of the biggest unit scheduled in one particular time. Other criteria for setting the reserve requirement are based on a percentage of the peak load or even as a hybrid function of these two quantities [39].

While the deterministic criteria are easy to implement, they do not represent the stochastic nature of the operation of power system and do not take into account the intrinsic reliability of the generating units. Probabilistic techniques for computing spinning reserve requirements, where a UC "risk index" determine the probability of failing to meet demand at some instant in time, has been proposed in [25]. The integration of the concept of a "risk index" with an LR-based UC program is described in [44], where it is proposed that the optimal risk index is determined by a balance between the cost of carrying a certain amount of reserve and the expected cost of energy not supplied. The first factor of the cost/benefit analysis is a by-product of the UC package while the second is simply the product of the expected energy not supplied by the value of loss load (VOLL).

2.2.3 Unit Constraints

1) Generating limits: the generating units should only be scheduled to supply power within the limits, set by their minimum synchronized generation level ($P_{min,i}$) and maximum capacity ($P_{max,i}$). This constraint can be formulated as follows:

$$P_{\min,i} U_i^t \le P_i^t \le P_{\max,i} U_i^t,\ \forall\, i \in [1, I]; t \in [1, T] \qquad (2.5)$$

2) Minimum up and down time: once a generating unit is committed, it should be remaining on-line for a minimum period of time (MUT_i). Similarly, one a generating unit is de-committed, it should not be re-committed before a minimum amount of time (MDT_i). These two constraints can be formulated as follows:

$$U_i^t = \begin{cases} 1, & \text{if } T_{on,i}^t < MUT_i, \\ 0, & \text{if } T_{down,i}^t < MDT_i, \\ 1 \text{ or } 0, & \text{otherwise} \end{cases} \qquad (2.6)$$

where

$T_{on,i}^t$: continuously online time of unit i at time t (hours)

$T_{down,i}^t$: continuously down time of unit i at time t (hours)

MUT_i : minimum up time of unit i (hours)

MDT_i : minimum down time of unit i (hours)

3) Start-up cost: The start-up costs are usually represented by exponential function of the point the units have been shut-down, $T_{down,i}^t$, such as

$$ST_i^t = \alpha_i + \beta_i(1 - \exp\left(\frac{-T_{down,i}^t}{\tau_i}\right) \quad \text{(2.7)}$$

The longer this period, the colder the unit is and the more expensive is to start-up. In (2.7), the crews start-up cost and equipment maintenance costs, which are in part proportional to the number of start-ups, are integrated in the term α_i, the cost associated with the required fuel to start-up the unit from the cold condition are represented by the term β_i, and the unit cooling rate is expressed by τ_i.

The start-up cost function of the unit is usually can be approximated by assuming that there are two start-up conditions, i.e. cold and hot start-up [49]. This approximation is formulated as

$$ST_i^t = \begin{cases} HST_i, \text{ if } MDT_i \le T_{down,i}^t \le T_{cold,i} + MDT_i, \\ CST_i, \text{ if } T_{down,i}^t > T_{cold,i} + MDT_i \end{cases} \quad \text{(2.8)}$$

where

HST_i : hot start-up cost of unit i ($)

CST_i : cold start-up cost of unit i ($)

$T_{cold,i}$: cold start hour of unit i (hours)

2.3 UC formulation based on MILP Method

2.3.1 Objective Function

The objective function in (2.1) can be reformulated in the form of mixed-integer linear programming (MILP), which can be presented as follows:

$$\min \sum_{t=1}^{T}\sum_{i=1}^{I} CP_i^t + CU_i^t \quad \text{(2.9)}$$

where CP_i^t and CU_i^t represent production cost and start-up cost respectively.

Two components of the objective function (2.9) mentioned above are explained in the following.

1) Production cost: As shown in Figure 2.1, the fuel cost function in (2.2) can be approximated by a set of piecewise blocks [51]. The number of piecewise block can be denoted as NL. In this research, two piecewise blocks are employed. F_{1i} and F_{2i} represent the slop of block l=1 and l=2 of unit i respectively. Meanwhile, the δ represents the generated power in the considered block of the piecewise linear cost function. The constant value A_i shows the constant cost at minimum power output, $P_{min,i}$. The value of A_i can be calculated as

$$A_i = a_i + b_i P_{\min,i} + c_i P_{\min,i}^2 \quad (2.10)$$

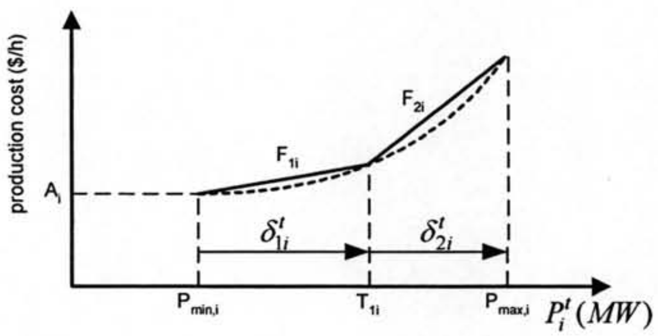

Figure 2.1 Piecewise linear production cost

The analytical representation of this linear approximation is

$$CP_i^t = A_i U_i^t + \sum_{l=1}^{NL} F_{li} \delta_{li}^t, \forall i \in I, \forall t \in T \quad (2.11)$$

$$P_i^t = \sum_{l=1}^{NL} \delta_{li}^t + P_{\min,i} U_i^t, \forall i \in I, \forall t \in T \quad (2.12)$$

$$\delta_{1i}^t \le PT_{1i} - P_{\min,i} U_i^t, \forall i \in I, \forall t \in T \quad (2.13)$$

$$\delta_{2i}^t \le PT_{2i} - PT_{1i}, \forall i \in I, \forall t \in T \quad (2.14)$$

$$\delta_{2i}^t \le P_{\max,i} - PT_{1i}, \forall i \in I, \forall t \in T \quad (2.15)$$

$$\delta_{li}^t \ge 0, \forall i \in I, \forall t \in T, \forall \in 1 \ldots NL \quad (2.16)$$

2) Start-up cost: Figure 2.2 shows the approximation of a typical exponential start-up cost function [39] with a stairwise cost function. Since the time span has been discretized into hourly periods, the start-up cost is also a discrete function, as shown in Figure 2.2 with blackened circles. The discrete start-up cost can be asymptotically approximated by a stairwise function (solid line in Figure 2.2).

A mixed-integer linear formulation for the stairwise start-up cost was proposed in [52].

$$CU_i^t \ge K_i^q \left[U_i^t - \sum_{r=1}^{q} U_i^{t-r} \right], \forall i \in I, \forall t \in T, \forall q \in 1 \ldots ND \quad (2.17)$$

$$CU_i^t \geq 0, \forall i \in I, \forall t \in T \quad (2.18)$$

Regarding the two start-up conditions as stated in (2.8), the coefficient K_i^q can be described as

$$K_i^q = \begin{cases} HST_i, \text{if q} = 1 ... T_{cold,i} + MDT_i, \\ CST_i, \text{if q} = T_{cold,i} + MDT_i + 1 ... \text{ND} \end{cases} \quad 2.19$$

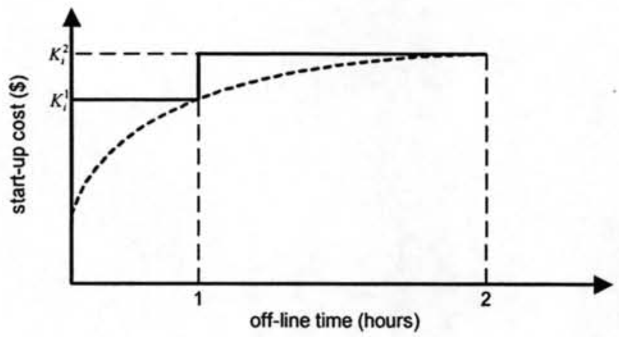

Figure 2.2 Exponential and stairwise start-up cost functions.

2.3.2 Minimum Up and Down Time Constraints

The new expressions for minimum up and down time constraints based on [51] can be formulated as

$$\sum_{t=1}^{V_i} \left[1 - U_i^t\right] = 0, \forall i \in I \quad (2.20)$$

$$\sum_{q=t}^{t+MUT_i-1} U_i^q \geq MUT_i \left[U_i^t - U_i^{t-1}\right], \forall i \in I, \forall t = V_i + 1 ... T - MUT_{i+1} \quad (2.21)$$

$$\sum_{q=t}^{T} \left[U_i^q - \left(U_i^t - U_i^{t-1}\right)\right] \geq 0, \forall i \in I, \forall t = T - MUT_i + 2 ... T \quad (2.22)$$

where V_i is the number of initial periods during which unit i must be online. V_i is mathematically expressed as $V_i = Min\left\{T, \left(MUT_i - T_{on,i}^0\right) U_i^0\right\}$.

Analogously, the new expression of minimum down time constraints can be formulated as shown in (2.23)-(2.25).

$$\sum_{t=1}^{W_i} U_i^t = 0, \forall i \in I \quad (2.23)$$

$$\sum_{q=t}^{t+MDT_i-1} \left[1 - U_i^q\right] \geq MDT_i \left[U_i^{t-1} - U_i^t\right], \forall i \in I, \forall t = W_i + 1 ... T - MDT_i + 1 \quad (2.24)$$

$$\sum_{q=t}^{T}\left[1-U_i^q-\left(U_i^{t-1}-U_i^t\right)\right]\geq 0, \forall i\in I, \forall t=T-MDT_i+2...T \text{ .... (2.25)}$$

where W_i is the number of initial periods during which unit i must be offline. W_i is mathematically expressed as $W_i = Min\{T, \left(MDT_i - T_{off,i}^0\right)\left(1-U_i^0\right)\}$.

2.4 Demand with Uncertainty Consideration

Since the expected product of UC process is a set of committed units to satisfy the demand for each time interval within the scheduling periods without violating the prevailing constraints hence it is necessary to provide a time chronological of the forecasted demand. The time interval is normally half hour or one hour. The accuracy of the short-term load forecast, as reported in [19], is generally in the range of 2 to 5 percent of the actual demand. It is found in general that, the longer lead time is, the higher uncertainty in the forecasted value [19, 53] will be embedded. A typical forecast demand of one day period by incorporating forecast error can be represented by low, medium, and high load level as shown in Figure 2.3.

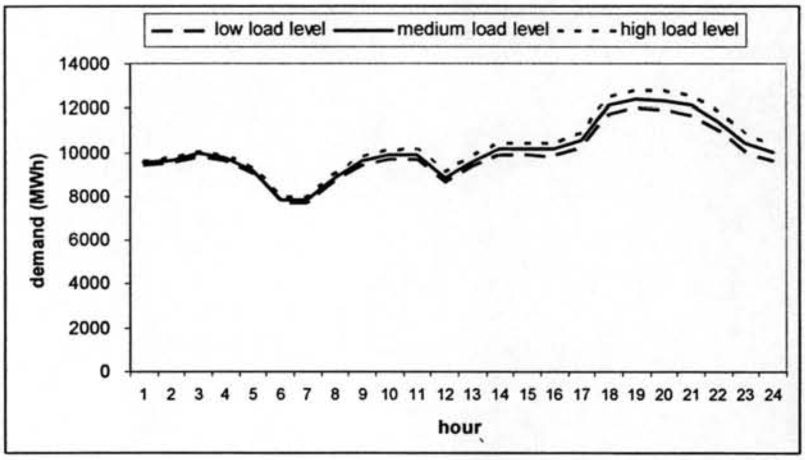

Figure 2.3 A typical of demand uncertainty as function of lead time

2.5 Economic Dispatch

In the term of generation scheduling, economic dispatch is a process of allocating output power to each committed unit, assuming that the thermal units at each time interval have been determined, so that the system demand at the considered time interval be supplied entirely subjected to the considered unit constraints at lowest production cost. Suppose that NI_t generating units are committed at time interval t, hence the ED problem can be formulated as

$$\min \sum_{i=1}^{NI_t} F_i\left(P_i^t\right) \qquad (2.26)$$

subject to

$$\sum_{i=1}^{NI_t} P_i^t = P_{load}^t \qquad (2.27)$$

$$P_{\min,i} \le P_i^t \le P_{\max,i} \qquad (2.28)$$

Since a linearized fuel cost characteristic is used in this research as has been explained previously in subsection 2.2, hence the objective function of (2.26) now becomes minimization of linear function. This linear objective function subjects to the system demand constraint (2.27) and the individual unit output constraint (2.28). Therefore the equations (2.26)-(2.28) constitute a linear programming problem which can be solved by various methods, e.g. equal incremental cost criteria [39], simplex method of Dantzig [54], or by using Interior Point Method (IPM) [55, 56].

2.6 Reliability Consideration

In short-term operation, it is important to ensure that electric power can be supplied to customer both at minimum production cost and at most reliable. The minimum cost can be obtained if utilities allocate the total online capacity at 'close to the edge' with the demand. In other hand, the responsibility of utilities to maintain the reliability at an acceptable risk level can be achieved if utilities provide enough spinning reserve. However, operation with too much reserve is not economic since it incurs higher cost. Therefore, it is important to obtain the trade-off between reliability and economic criteria.

Many papers in the previous works [49, 57] set the spinning reserve requirement based on deterministic criterion, e.g. percentage of forecasted demand. In practice, the amount of percentage reserve is defined based on the previous experience. Other approach in determining the spinning reserve requirement with consideration of stochastic nature of each generating unit is based on probabilistic technique. The technique for computing the spinning reserve requirement have been proposed in [25, 44]. In [44] risk index is used as stopping criteria to measure how much reserve should be scheduled. Recent works which incorporating reliability issues in determining the UC decision by using LOLP and EUE constraints is presented in [45]. The total cost under this technique is composed of production cost and reliability cost.

Since most utilities still use deterministic criteria, i.e. specific percentage of demand, it is necessary to evaluate the impact of different amount of percentage criteria to the total cost, which is defined as the summation of production cost and cost of expected unserved energy (EUE). The Expected Unserved Energy (ENS) cost per-kWh, which may be called as value of loss load (VOLL). In Argentina, the VOLL is set at $US 2/kWh [1, 58], whereas it was set at £ 2/kWh in the previous UK pool market. Studies identified VOLL varying between 1-3 $US/kWh depending on different consumer class [1].

In this thesis, the considered model in which some part of the contracted demand need a certain level of back-up capacity creates new problems in related to the determination of appropriate spinning reserve to maintain acceptable reliability level of the overall obligated demand.

2.7 Solution Technique of Unit Commitment Problem

The thermal unit commitment problem has been traditionally solved in centralized power systems to determine when to start-up or shutdown thermal generating units and how to dispatch online generators to meet system demand and spinning reserve requirements while satisfying unit constraints over a specific short-term time span, so that the overall operation cost is minimized [39].

For several decades, this large-scale, mixed-integer, combinatorial, and nonlinear programming problem has been an active research topic because of potential savings in operation costs. As a consequence, several solution techniques have been proposed such as heuristics [36, 59], dynamic programming [60-62], mixed-integer linear programming (MILP) [51, 63, 64], Lagrangian relaxation [49, 57, 65], simulated annealing [45, 66-68], and evolution-inspired approaches [69-72]. A recent extensive literature survey on unit commitment can be found in [26]. Among the aforementioned methodologies, Lagrangian relaxation is the most widely used approach because of its capability of solving large-scale problems. The main disadvantage of this method is that, due to the non-convexities of the unit commitment problem, heuristic procedures are needed to find feasible solutions, which may be suboptimal.

In contrast, MILP guarantees convergence to the optimal solution in a finite number of steps [73] while providing a flexible and accurate modeling framework. In addition, during the search of the problem tree, information on the proximity to the optimal solution is available. Efficient mixed-integer linear software such as the branch-and-cut algorithm has been developed, and optimized commercial solvers with large-scale capabilities are currently available [74, 75]. As a consequence, a great deal of attention has been paid to MILP-based approaches.

In [63], MILP was first applied to solve the unit commitment problem. The formulation in [63] was based on the definition of three sets of binary variables to, respectively, model the start-up, shut-down, and on/off states for every unit and every time period.

This mixed-integer linear formulation was extended in [64] to model the self-scheduling problem faced by a single generating unit in an electricity market. Non-convex production costs, time-dependent start-up costs, and inter-temporal constraints such as ramping limits and minimum up and down times were accounted for at the expense of increasing the number of binary variables. For realistic power systems comprising several tens of generators, the models of [63] and [64] require a large number of binary variables. Thus, the resulting MILP problems might be computationally intensive for state-of-the-art implementations of branch-and-cut algorithms [74, 75] and current computing capabilities.

The latest improvement of MILP formulation proposed in [51] which only requires a single set of binary variable (one per unit and per period). This alternative mixed-integer linear formulation is denoted by MILP-UC. Unlike previous MILP approach [63, 64], the lower number of binary variable proposed in MILP-UC yields a reduction in the number of node of the search tree used by branch-and-cut algorithm, as well as reduction in the number of constraint thus decreasing the computation time. Moreover, MILP-UC accurately model thermal unit commitment states, intertemporal constraints, and time-dependent start-up cost, thereby improving the modeling capabilities.

2.8 Decision Analysis for Generation Scheduling

2.8.1 Basic Concept of Decision Analysis

Electric utilities have faced with increasing uncertainty in engineering and management decisions, which include unpredictable energy supply, capital market chaos, demand uncertainty, and other regulatory statutes. These significant uncertainties have motivated some electric utilities to consider using decision analysis as a tool to determine engineering and management decisions. Decision analysis is a formal procedure for analyzing and relating the factors relevant to a decision analysis on an explicit and logical basis, with special emphasis on the treatment of uncertainty. A review of wide range applications of decision analysis on electric power system engineering and management are presented in [76]. The basic concept of decision making process [77, 78] by employing decision analysis method can be summarized below.

The first step of decision analysis is to list viable options for gathering information, experimentation, and actions. After the list of possible events is generated, the list is sorted into chronological order. Next, the consequences for each possible state have to be identified as well as the probability of each event. The critical component is to include all costs, benefits, and risks. It is always better to do no decision analysis than to do a poor analysis.

The second step in applying decision analysis is to draw the tree. The decision tree includes three types of nodes: decision, natural and result. The decision nodes have connections to natural nodes, one connection for each possible decision. The natural nodes are connected to other natural nodes or, finally to result nodes. Each connection leaving a natural node identifies the probability (subjective or estimated) of traversing to the next state based on all other probable outcomes for that natural node. The final nodes are the result nodes which give the benefit minus the cost for the alternative given the present state defined by all of the probability connections to natural nodes back to the first decision node. All decision trees start with a decision node. All decision trees end with result nodes.

The next step is to evaluate the tree. The common approach is to assume that all natural nodes are Bayesian (normally distributed). Then the variable to be calculated at each node, the expected monetary value is simply a summation of the weighted costs from the next node (natural or result). The alternative selected at each decision node is made by selecting the path, which yields the maximum value in the case of maximization problem or minimum value in the case of minimization problem.

The final step is to perform a sensitivity analysis. How much do the probabilities have to vary in order to change the decision(s)?

2.8.2 Decision Analysis for Solving Generation Scheduling Problem

Regarding to generation scheduling problem, the application of decision analysis to analyze the risk of short-term power system operational planning in presence of electrical load forecast uncertainty is reported in [22]. Another work [78] proposed the application of decision analysis for solving the dispatch problem of GENCOs under competitive generation market structure. In this paper, decision analysis is employed to determine the best dispatching strategy which constitutes the balance between risk and expected profit. The implementation of decision making steps presented in the previous subsection for solving generation scheduling problem can be described below.

The first step is to gather all required data for determining committed units in generation scheduling problem. The required data for solving generation scheduling problem based on decision analysis can be summarized as follows:

a. Generating unit's data: fuel cost characteristic, minimum and maximum generation limit, minimum up and down time constraints, start-up cost characteristic, and reliability data.

b. Demand data: forecasted demand for 24-hour and its uncertainty parameter, i.e. standard deviation and occurrence probability of load level.

c. Other related parameters, i.e. value of loss load, spinning reserve requirement (as a percentage of demand).

Based on the demand uncertainty, a set of demand scenarios for the 24-hour period has to be developed. Each of the scenario represents information about the possible demand to be considered. Each scenario is assigned with a weight of, PL_j, which reflects the possibility of its occurrence in the considering period. Each demand scenario is used as input in a conventional unit commitment problem with a specified spinning reserve requirement. This unit commitment problem can be solved by any developed unit commitment technique, e.g. MILP method. Based on the obtained committed units, then the consequence of generation cost and risk cost for each possible load level can be calculated.

Suppose that the number of load scenarios or levels is J. The decision tree of generation scheduling problem can be illustrated as shown in Figure 2.4. The decision tree is developed by connecting decision nodes to the J number of natural nodes. Each natural node represents unit commitment (UC) decision of each considered load level. The consequence cost of a specified scenario as a summation of all considered costs, i.e. generation cost and risk cost after weighted by each occurrence probability, can be calculated based on the resulted committed units from the natural node. The best alternative is then selected among the UC strategies which give minimum expected total cost. The impact of some investigated parameters to the selected scenario can be performed by performing sensitivity analysis. All the details and test results will be presented in Chapter 3.

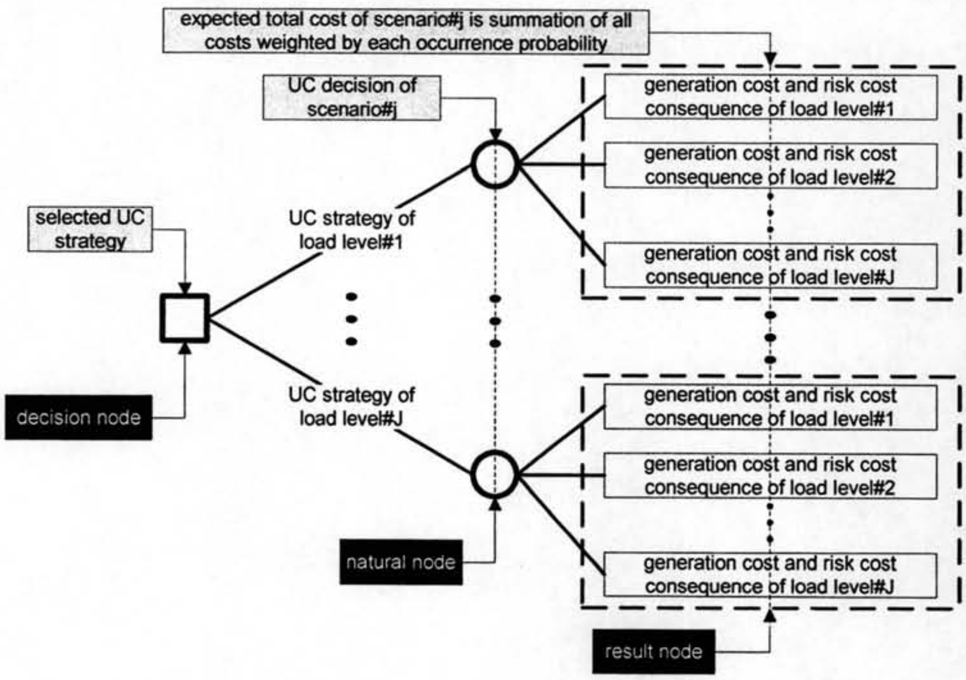

Figure 2.4 Decision tree of generation scheduling problem