A GENETIC ALGORITHMS APPROACH FOR GENERATION EXPANSION PLANNING OPTIMIZATION

Y. M. Park | **J. B. Park** | **J. R. Won**

Seoul National University, Electrical Engineering Depart., 56-1, Shilimdong, Kwanakgu, Seoul, Korea
TEL : 82-2-880-7251
FAX : 82-2-871-5974

Korea Electric Power Corp., Electricity Economics Depart., 167, Samsungdong, Kangnamgu, Seoul, Korea
TEL : 82-2-550-6234
FAX : 82-2-550-6259

Seoul National University, Electrical Engineering Depart., 56-1, Shilimdong, Kwanakgu, Seoul, Korea
TEL : 82-2-880-7251
FAX : 82-2-871-5974

Abstract: This paper shows how a genetic algorithms (GA) approach can be applied to solve a least-cost generation expansion planning problem. Many methods have been suggested to solve this problem. However, most of previous algorithms have the defect of falling in a local optimal valley. Furthermore, the integer problem still has remained as a difficult issue. The GA can search for the global optimum and treat integer problem naturally. This paper presents a new GA which adopts two methodologies. The first is to use a composition mapping for the evaluation of each string's fitness. The second is to incorporate a heuristic-rule into GA procedure.

Keywords: Generation expansion planning, Genetic algorithms, Local and global optimum, Composition mapping, Heuristic-rule

1. INTRODUCTION

Traditionally, optimal generation expansion planning (GEP) is to determine the yearly capacity additions of each generator type and the corresponding number of units *etc.*, with least-cost while satisfying the given reliability criteria (Jenkins and Joy, 1974; Caramanis *et al.*, 1982). Thus, the objective function of a GEP problem has been the discounted total cost which incorporates the construction & operating costs, and salvage costs *etc.*

Many outstanding mathematics-based algorithms have been suggested to solve a least-cost GEP problem. These approaches include linear and nonlinear programming, integer and mixed integer programming, dynamic programming (Oatman and Hamant, 1973; David and Zhao, 1989) maximum principle (Park *et al.*, 1985), and artificial intelligence approaches using fuzzy and expert systems (David and Zhao, 1989, 1991). Since a GEP problem has highly nonlinearities resulting from production costing simulation (Lin *et al.*, 1989) and many other physical constraints, these mathematical algorithms have a common drawback of falling in the local optimal trap.

Another difficulty inherent in a GEP problem is the integer problem. The control variables (the number of each candidate generator in our case) should have the integer values with the upper and lower bounds. Since LP, NLP, and maximum principle approaches deal with the control variables in continuous space, it is difficult to assign continuous values as integer ones accurately. Though DP and integer programming can overcome this integer problem, they still have the problem of 'the curse of dimensionality'. Thus, the commercial packages such as WASP (Jenkins and Joy, 1974) and EGEAS (Caramanis *et al.*, 1982) search the reduced solution domain, which result in a local optimal solution.

The genetic algorithm is a search algorithm based on the mechanics of natural selections and natural genetics (Goldberg, 1989). Until now, GA has been applied to some areas of power system such as economic load dispatch (David and Gerald, 1993), reactive power planning (Iba, 1994), unit commitment(Dasgupta and McGregor, 1994). However, most of the researches are confined to static optimization problems.

The following advantages can be obtained by applying GA to a GEP problem. First, GA has the possibility of finding the global optimum through multiple path searching. Second, the dimensionality problem related with a GEP can be easily eliminated since it only needs memory size corresponding to the number of population. Third, as GA uses binary number {0,1} to represent each string, it can treat integer problems and constraints on control variables naturally.

This paper presents a new GA approach which adopts the following two methodologies. Firstly, a composition mapping is used for the evaluation of each string's fitness, which scales each fitness value to prevent the premature convergence. Secondly, a heuristic rule is incorporated into GA under the assumption that the global solution is probably located in the neighbourhood of reliability boundary. In this rule, a set of good solutions with higher fitness values is selected. After that, a target plant type is randomly chosen in each year and move the selected solutions toward reliability boundary regions.

2. MATHEMATICAL FORMULATION

The standard GEP problem can be formulated mathematically by solving the following equations. The objective of an electric utility's GEP is to decide the least-cost supply plan that meets future demand over a planning horizon (typically 10 to 20 years).

$$\min_{U_1,\dots,U_T} J=\sum_{t=0}^{T}\left(f_1^t(\mathbf{U}_t)+f_2^t(\mathbf{X}_t,\mathbf{U}_t)-f_3^T(\mathbf{X}_t,\mathbf{U}_t)\right) \quad (1)$$

$$s.t.\quad \mathbf{X}_{t+1}=\mathbf{X}_t+\mathbf{U}_t \quad \text{for } t=0,\cdots,T-1 \quad (2)$$

$$\mathrm{LOLP}(\mathbf{X}_t)\le\varepsilon \quad \text{for } t=1,\cdots,T \quad (3)$$

$$0\le\mathbf{U}_t\le\overline{\mathbf{U}_t} \quad \text{for } t=1,\cdots,T \quad (4)$$

$$\underline{P_t}\le\mathbf{X}_t\le\overline{P_t} \quad \text{for } t=1,\cdots,T \quad (5)$$

where,

T : the planning horizon(year)
$\mathbf{X}_t$: the vector which denotes cumulative capacity by plant types in year t
$\mathbf{U}_t$: the vector which denotes capacity addition for year t
$\overline{\mathbf{U}_t}$: the upper bound of each plant in year t
$\mathrm{LOLP}(\mathbf{X}_t)$: the loss of load probability of $\mathbf{X}_t$
ε : the reliability criteria of LOLP
$\underline{P_t}$, $\overline{P_t}$: the lower and upper bound of total installed capacity
$f_1^t(\mathbf{U}_t)$: the discounted construction cost with $\mathbf{U}_t$ in year t
$f_2^t(\mathbf{X}_t,\mathbf{U}_t)$: the discounted operating cost with $(\mathbf{X}_t,\mathbf{U}_t)$ in year t
$f_3^T(\mathbf{X}_t,\mathbf{U}_t)$: the discounted salvage cost with $(\mathbf{X}_t,\mathbf{U}_t)$ in year T

Thus, the objective of this problem is to find a set of optimal control vector ($\mathbf{U}^*_t$ *for* $t=1,\ldots,T$) which minimizes the objective function under the equality and inequality constraints. Although there exist many algorithms to calculate the operating cost and reliability measure (Lin *et al.*, 1989), piecewise linear approximation (PLA) is used.

Originally, the state($\mathbf{X}_t$) and control vector($\mathbf{U}_t$) in the year t means the sum of cumulative and additive capacity(MW) vector in the year t, respectively. These vectors can be converted equivalently into the following vectors which need only the information about the number of each generator type, *i.e.*, x_t, u_t.

$$\mathbf{X}_t'=(x_t^1,x_t^2,\cdots,x_t^N)' \quad (6)$$

$$\mathbf{U}_t'=(u_t^1,u_t^2,\cdots,u_t^N)' \quad \text{for } t=1,\cdots,T \quad (7)$$

3. GENETIC ALGORITHMS

The mechanics of GA are quite simple, involving nothing more complex than copying strings and swapping partial strings (Goldberg, 1989). GA starts with a population of strings generated at random, and then generates successive populations by processing the genetic operations. Most of conventional algorithms such as LP, NLP, and DP *etc.*, are single path search. However, GA performs multiple path search for the global optimum.

In GA, the fitness function is a procedure to evaluate fitness of each string in a population. The fitness value is a measure that shows how well a string can survive in a certain environment.

Genetic operations are the procedure to produce a population of offspring. Unlike the conventional mathematics-based algorithms, GA adopts the stochastic transition rules. These operations are applied to each string during each generation to generate a new and improved population from the old one. GA is composed of three basic operations: reproduction, crossover, and mutation.

4. DETAIL OF SUGGESTED ALGORITHM

4.1 String Structure of Control Variables

Since the objective of a GEP problem is to find a set of optimal control vectors with integer values that covers the whole stages, these control variables are represented as a string in this paper. If the number of power plant types is I, and the number of stages is T, then a string U is encoded as the following structure.

$$\begin{aligned} U &= [\quad TP\,1 \quad | \cdots | \quad TP\,I \quad] \\ &= [\ TP\,1_1, \cdots, TP\,1_T \mid \cdots \mid TP\,I_1, \cdots, TP\,I_T\ ] \end{aligned} \tag{8}$$

where,
TPi: a set of i type plant over the whole stage
TPi_j: the number of i type plants added in stage j

The suggested encoding scheme utilizes the binary alphabet. A string is divided by each plant type(*TPi*). Each *TPi* is also composed of TPi_j that contains the information on the number of added plants in each stage.

4.2 Random Creation of First Generation

To construct a string, each bit which has binary number {0,1} is created using uniform random variables. This process is repeated until a string is completely filled. In this process, if any TPi_j of a string violates the inequality constraint of eq. (4), then it is eliminated from the list. Furthermore, a string which violates the LOLP criteria is also eliminated from the initial population. If a set of strings, which satisfy the reliability and inequality constraints, is generated, it becomes the first generation.

4.3 Fitness Function Using Composition Mapping

After translating the information of strings to integer state variables, the cost (eq. 1) of each string can be obtained by probabilistic simulation and direct calculations. In general, the fitness function is defined as follows (Goldberg, 1989);

$$f = \frac{\alpha}{1 + J} \tag{9}$$

where,
α : constant, J : the total cost (eq.1)
f : fitness function

However, this kind of mapping occasionally brings about duplication in a population since strings with higher fitness values occupy large parts in a roulette wheel. Accordingly, this results in premature convergence due to

TYPE1 TYPE2 TYPE3
|xx|xx|xx|xxx|xxx|xxx|xx|xx|xx|

STAGE1 STAGE2 STAGE3

TYPE2
| x x x | x x x | x x x |
$2^2\ 2^1\ 2^0$
(The Number of TYPE2 Units
to be Added in Stage 2)

Fig. 1. Encoding Illustrations

proportional selection scheme (Grefenstette, 1986; Goldberg, 1989; Craig Potts *et al.*, 1994).

In this paper, the following composition mapping is adopted, which is a kind of fitness value scaling in order to prevent above-mentioned problems (Goldberg, 1989; David and Gerald, 1993). Thus, the fitness value of string i can be determined by the following composition mapping.

$$\hat{f}(i) = \frac{f(i) - f_{min}}{f_{max} - f_{min}} \tag{10}$$

where,
f(i): fitness value of string i (eq.9)
f_{max}, f_{min}: maximum and minimum fitness values in a certain generation

By using the modified fitness function, occupied portions among strings with different fitness values can be somewhat levelized in the roulette wheel. Thus, the object to prevent the premature convergence and duplication among strings can be obtained.

4.4 Crossover and Mutation

In the conventional genetic operations, crossover and mutation are performed with a whole string. But in this paper, they are accomplished by each plant type(*TPi*). The reason to deal with each *TPi* for genetic operations results from the long-length of a string and its structure. Although these genetic operations result from the empirical situation, they do not deviate from the original concepts of GA, and showed good convergence characteristic throughout the case studies.

4.5 Heuristic Rule

The Heuristic-rule considered in this paper is based on the following assumption.

Assumption : The optimal solution of a GEP problem is probably located in a local neighbourhood of reliability boundary.

In this paper, a population is composed of a set of strings which satisfy the reliability criteria. By this reason, strings will encounter two conflicting environments, *i.e.*, fitness maximization (cost minimization) and reliability satisfaction. Accordingly, strings will evolve in both conflicting directions. One of which is to minimize cost function, and the other is to satisfy the LOLP criteria enough to survive as times go by. Furthermore, it needs much exploitation time to get one of local optimums in conventional GA. To overcome these obstacles, the following heuristic rule is incorporated into the GA procedure.

Step 1. Perform this heuristic rule for every M generation ($M \leq N$, N : total number of generations)
Step 2. Sort strings in the descending order of their fitness values, and choose the best L strings (*i.e.*, an elite group).
Step 3. Select one plant type randomly each year for each string in the elite group and reduce its number by 1.
Step 4. Check reliability criteria. If satisfied, then update the strings with new ones. Otherwise, the strings remain as old ones. Increase the year by 1 and go to step 3. If all the years and all the strings in the elite group are considered, then go to step 5.
Step 5. Stop the application of this rule, and go to next generation for GA procedure.

5. CASE STUDIES

5.1 Input Data

The initial configuration existing in the reference year, 1994, is given in Table 1. The planning horizon extends over 10 years which are divided into five stages, at two yearly intervals. Also it is assumed that the future peak demands have the deterministic ones as shown in Table 2. The technical and economic data of candidate plant types are presented in Table 3. There are five types of candidate plant options available, four of them are the same as the existing plant types, while a PHWR plant is a new option for the future. Each five item constitutes a dimension in the control and state space of our optimization problem.

Note that there are other important input data such as LOLP criteria, and discount rate. The criteria of LOLP is set as 0.02, and the discount rate as 8.5%.

Table 1 Existing System Data

Type	No. of Units	Capacity (MW)	FOR (%)	Operating Cost ($/kWh)
Oil	2	200	7.0	0.024
LNG C/C	2	400	10.0	0.033
Coal	1	500	9.0	0.019
Nuclear(PWR)	1	1,000	9.0	0.0048
Total Capacity : 2,700 MW				

Table 2 Future Peak Demand

Stage	Year	Peak(MW)
0	1994	2,500
1	1996	3,500
2	1998	4,500
3	2000	5,500
4	2002	6,500
5	2004	8,000

5.2 Numerical Results

The proposed GA was applied to the sample system. Although GA depends on some parameter values such as population size, probability of crossover and mutation, and generation number (Goldberg, 1989; Craig Potts *et al.*, 1994), most of these values were determined throughout empirical tests. The probability of crossover was set as 0.9, population size as 150, and generation number as 70, while mutation probabilities were changed from 0.001 to 0.009. Also, the heuristic rule is applied every 1 to 10 generations.

To show its effectiveness and efficiency, the results of the suggested GA were compared with those of truncated DP and DP. The result of truncated DP, which is the optimization technique in WASP (Jenkins and Joy, 1974) package, is considered as a local optimum of the GEP problem. Also, the result of DP which searches the whole state space is considered as the global optimum. The results of conventional GA using eq. 9, GA using composition mapping of eq. 10, and the suggested GA with composition mapping and heuristic rule are compared with those of DP, respectively.

Conventional GA failed to pass through the local optimum obtained by WASP due to duplication among strings and premature convergence (Table 4). However, the GA using composition mapping succeeded to find other local solutions.

Table 3 Future Candidate Plant Data

Type	Upper Limit for Construction	Capacity (MW)	FOR(%)	Operating Cost ($/kWh)	Capital Cost ($/kW)	Life Time (Years)
Oil	6	200	7.0	0.024	812.5	25
LNG C/C	5	400	10.0	0.033	500.0	20
Coal	3	500	9.0	0.019	1062.5	25
Nuclear (PWR)	2	1,000	9.0	0.0048	1625.0	25
Nuclear (PHWR)	2	700	7.0	0.0029	1750.0	25

Table 4 Simulation Results

Case	Parameter		Cumulative Discounted Cost (10^6 $)		
	PC	PM	Conventional GA	GA with Composition Mapping	Suggested GA
DP	-			4908.3	
Truncated DP	-			4960.6	
Case 1	0.9	0.001	5139.4	4983.4	4977.1
Case 2	0.9	0.005	5200.3	4965.8	4952.7
Case 3	0.9	0.006	5128.8	4956.2	4949.5
Case 4	0.9	0.007	5107.1	4958.9	4947.3
Case 5	0.9	0.008	5208.9	4963.3	4954.4
Case 6	0.9	0.009	5118.8	4951.9	4933.5

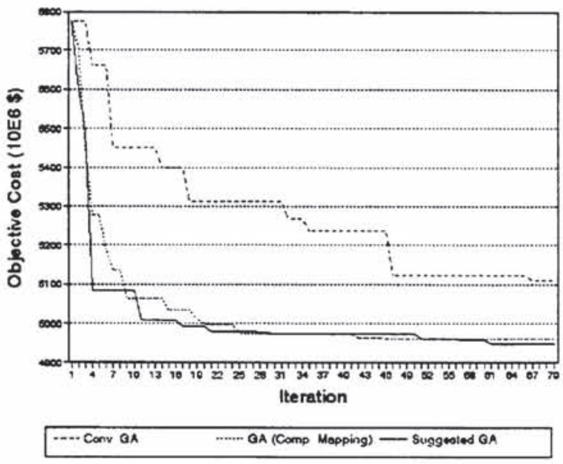

Fig. 2. Convergence Characteristics

Table 5 CPU Times

Algorithms	Elapsed CPU Time
Truncated DP	45[mins]
DP	14[hrs] and 30[mins]
GA with Fitness Function (Eq. 9 & 10)	3[hrs] and 20[mins]
Suggested GA	4[hrs] and 10[mins]

Furthermore, every solution of the suggested GA except case 1 is better than the local optimum of WASP(Table 4). Errors between the global optimum and solutions of the suggested GA ranges from 0.5 percent to 1.4 percent, while errors between the global solution and those of conventional GA reaches from 4.1 to 6.1 percent.

Fig. 2 shows the convergence characteristics of case 4 in table 4. The convergence characteristics of GA using composition mapping and GA with composition mapping and heuristic rule are very good. The GAs presented in this paper, truncated DP, and DP were implemented on 80486 50MHz IBM PC. The required average CPU times for each algorithm are given in Table 5.

In terms of the global optimal solution, it is difficult to find the global optimum by using the suggested GA. However, the proposed method can always get many quasi-optimums, which are important information when making decisions for capacity investment, stably and surely with less computation time.

Additionally, the premature convergence of conventional GA can be overcome by using the composition mapping. Also, the convergence speed to get a local solution is accelerated by adopting the heuristic rule.

6. CONCLUSION

This paper presents a new approach for the generation expansion planning optimization. The advantages of the suggested GA lie in its ability to handle integer problems and to circumvent 'the curse of dimensionality' of a GEP problem naturally, which are impossible to consider both simultaneously in traditional mathematical algorithms.

The proposed GA utilizes the composition mapping for the evaluation of each string's fitness to prevent the premature convergence and duplication among strings. Another main feature of the suggested GA is the incorporation of heuristic rule into GA procedure. The incorporation of this heuristic rule into GA resulted in fast convergence to better solutions. Throughout case studies, it was shown that the suggested GA could search many quasi-optimal solutions stably and easily with less computation time.

Despite of some problems inherent in GA, it can be a powerful search algorithm in the coming computer era especially in the area of the GEP optimization.

REFERENCES

Caramanis, M.C., F.C. Schweppe and R.D. Tabors (1982). *Electric Generation Expansion Analysis System.* EPRI EL-2561, Vol. 1, Vol. 2, RP1529-1, Final Report.

Craig Potts, J., T.D. Giddens and B.Y. Surya (1994). The development and evaluation of an improved genetic algorithm based on migration and artificial selection. *IEEE Trans. on SMC*, **24**, 73-86.

Dasgupta, D. and D.R. McGregor (1994). Thermal unit commitment using genetic algorithms. *IEE Proc.-Gener. Transm. Distrib.*, **141**, 459-465.

David, A.K. and R.D. Zhao (1989). Integrating expert systems with dynamic programming in a generation expansion planning. *IEEE Trans. on Power Systems*, **4**, 1095-1101.

David, A.K. and R.D. Zhao (1991). An expert system with fuzzy sets for optimal planning. *IEEE Trans. on Power Systems*, **6**, 59-65.

David, C.W. and B.S. Gerald (1993). Genetic algorithm solution of economic dispatch with valve point loading. *IEEE Trans. on Power Systems*, **8**, 1325-1332.

Goldberg, D.E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*, Addition-Wesley Publishing Company, Inc.

Grefenstette, J.J. (1986). Optimization of control parameters for genetic algorithms. *IEEE Trans. on SMC*, **16**, 122-128.

Iba, K. (1994). Reactive power optimization by genetic algorithm. *IEEE Trans. on Power Systems*, **9**, 685-692.

Jenkins, R.T. and D.S. Joy (1974). *Wien Automatic System Planning Package (WASP)*, ORNL-4954, Oak Ridge National Laboratory.

Lin, M., A. Breipohl and F. Lee (1989). Comparison of probabilistic production cost simulation methods. *IEEE Trans. on Power Systems*, **4**, 1326-1334.

Oatman, E.N. and L.J. Hamant (1973). A dynamic approach to generation expansion planning. *IEEE Trans. on PAS*, **92**, 1888-1982.

Park, Y.M., K.Y. Lee and L.T.O. Youn (1985). New analytical approach for long-term generation expansion planning based on maximum principle and gaussian distribution function. *IEEE Trans. on PAS*, **104**, 390-397.