Journal of Intelligent Manufacturing, 15, 701–709, 2004

Solution of trim-loss problem by an integrated simulated annealing and ordinal optimization approach

CHIA HUANG YEN, DAVID SHAN HILL WONG and SHI SHANG JANG*

Department of Chemical Engineering, National Tsing Hua University, Hsinchu, Taiwan 300
E-mail: ssjang@che.nthu.edu.tw

Received March 2003 and accepted February 2004

This work presents a novel optimization method capable of integrating ordinal optimization (OO) and simulated annealing (SA). A general regression neural network (GRNN) is trained using available data to generate a ''rough'' model that approximates the response surface in the feasible domain. A set of ''good enough'' candidates are generated by conducting a (SA) search on this ''rough model''. Only candidates accepted by the SA search are actually tested by evaluating their true objective functions. The GRNN model is then updated using these new data. The procedure is repeated until a specified number of tests have been performed. The method (SAOO + GRNN) is tested the well-known paper trim loss problem. SAOO + GRNN approach can substantially reduce the number of function calls and the computing time far below those of simple ordinal optimization method with such as horse race selection rule, as well as straightforward simulated annealing.

Keywords: Ordinal optimization, simulated annealing, general regression neural network, integrated approach, trim-loss problem

1. Introduction

The solution of large scale combinatorial optimization, or mixed integer programming (MIP) problem is very important in process and product development (Davis, 1999). Whereas many process designs constitute networking and scheduling problems, product developments such as drug design, catalyst synthesis and solvent selections also involve combinations of molecular building blocks.

One category of solutions to IP problems is to solve the relaxed LP or NLP problem; these solutions use branch-and-bound or another tree-search method to search the integer variable space (Floudas, 1999; Edgar *et al*., 2001). Such methods are usually very efficient if provided that the number of integer variables is not too large and the relaxed LP and NLP problems have a well-behaved response surface. The solution obtained is a local maximum and global optimality can be guaranteed in some cases.

Another category of solutions is using heuristic methods such as simulated annealing (SA, Kirkpatrick *et al*., 1983), and genetic algorithms (GA, Holland, 1975). Heuristic methods are often used to solve large scale combinatorial optimization problems. Heuristic based methods are simple to program and normally effective in locating an adequate solution. However, the optimality of the solution cannot be guaranteed.

Ho and coworkers (e.g., Ho *et al*., 1992) introduced ordinal optimization to assess the effectiveness of the heuristic based decision methods. Using probability theory, they explained why heuristic methods are effective in locating a ''good enough'' solution.

*Author for correspondence.

Moreover, they demonstrated how a ''rough'' model, despite its lack of precision, can improve the efficiency of a search for a ''good enough'' solution.

Considered a numerical experiment that includes a set of $N = 1000$ candidates for the optimality. Let Z be the number of candidates selected by some model. The alignment number K is the actual number of points ranked among the top G candidates that are actually selected. Lau and Ho (1997) found that K depends only on the distribution of the fitness values (the shape of the ordinal performance curve OPC), and the accuracy of the model. Five categories of OPC (Flat, Neutral, U-shaped, Bell-shaped and Steep) are presented by the original works of Ho and co-workers (e.g., Lau ad Ho, 1997). A model accuracy parameter W is defined. When $W = 0.5$, the error of the model span the whole range of distribution values. There is non-zero probability that the poorest sample is predicted to be a good enough solution. They also showed that even with such poor models, the alignment number K can be increased significantly when a horse race selection rule is used.

However, ordinal optimization theory does not specify how such a rough model can be built. Various methods for constructing such a ''rough'' model were proposed in the literature. For example, Ho *et al.* (2001) employed heuristic knowledge of the nature of the solution to improve solution to the Witsenhausen problem. The model used is problem specific. Luo *et al.* (2001) suggested a method that combined the concept of ordinal optimization and the genetic algorithm. Solutions that are sufficiently fit are crossbred to generate candidates of good solutions.

Simulated annealing (SA) is a highly effective stochastic method for solving combinatorial optimization problems based on ideas from statistical mechanics. The theory has been extensively researched (Collins *et al.*, 1988) and has many applications to different problems as discussed in the literature (Johnston *et al.*, 1989; Eglese, 1990; Ku and Karimi, 1991; Koulamas *et al.*, 1994; Painton, 1994; Falcioni and Deem, 2000). In a typical SA scheme, a reference solution is given initially. A ''trial'' solution in the neighborhood of this reference solution is generated. The objective function of this trial solution is calculated. If it is lower than the objective function of the reference solution, then the reference solution is replaced by the trial solution. Otherwise, an exponential type transition probability is applied to determine whether the solution should be updated:

$$P = \begin{cases} 1 & Y(x^t) < Y(x^r) \\ \exp\left(\dfrac{-(Y(x^t) - Y(x^r))}{T}\right) & \text{otherwise} \end{cases} \quad (1)$$

where T is the annealing temperature in the SA, search procedure, $Y(x^t)$ is a trial solution and $Y(x^r)$ is a reference solution in sequential search procedure. The temperature is reduced gradually according to a predetermined annealing schedule to ensure that the system is not trapped in a local minimum. In the SA scheme, the actual objective functions are evaluated for all trial solutions. In a practical engineering problem, the evaluation may involve lengthy calculations or even actual experiments. However, the whole scheme is used to ensure that more potentially ''good'' candidates are sampled. Therefore, the ''actual'' fitness function need not be applied. A ''rough'' prediction model can be used instead for generating these ''potentially good'' candidates. Furthermore, according to ordinal optimization theory, only a small fraction of these generated candidates need be tested to ensure alignment.

Genetic algorithm (GA) is also an important branch in the area of stochastic approach method for solving combinatorial optimization problems. Combination of GA and ordinal optimization has been shown to be effective for sampling strategies (e.g., Zhang *et al.*, 2002; Luo *et al.*, 2001). Further improvement in sampling is made possible by including information theory sampling (Tsujimura and Gen, 1998). There is no doubt that if we implement GA instead of SA in our integrated approach mentioned in the following sections, the result should be of interest to combinatorial optimization researchers. Due to the limit of our research resources, we leave this topic for our future research.

The ''trim-loss'' or ''cutting-stock'' problem is an integer programming problem that arises in the paper cutting industry. The main task is to cut paper products of different sizes from a large roll of paper to meet customers' orders. A set of orders can generally not be met without throwing away some of the raw paper. The optimum cutting scheme minimizes the waste paper or trim loss. The problem involves only integer and binary variables, which appear in a linear objective function, and linear as well as bilinear constraints. Depending on the individual problem, the search space can be very large. Many methods have been proposed for solving the problem. For example, Wäscher (1990) combined linear programming methods with heuristic methods to round the continuous LP-solutions into realizable integer ones. Goulimis (1990) proposed a branch and bound and a cut-plane approach. Harjunkoski *et al*. (1996) suggested an *a priori* pattern generation method to reduce the size of the problem and to combine it with a MILP strategy. Foerster and Wäscher (1998) employed simulated annealing. Östermark (1999) used a genetic hybrid algorithm. Floudas *et al*. (1999) applied a global optimization method known as α-branch-and-bound. Harjunski *et al*. (1998) demonstrated that the ease of finding a solution depends on how the formulations of the objective function and the constraints.

This work presents a novel optimization procedure by integrating the concepts of ordinal optimization and simulated annealing. The effectiveness of this approach was tested using the paper trim loss problem.

2. Methodology

2.1. *The SAOO + GRNN procedure*

In this paper, we proposed to use SA as the selection rule in ordinal optimization and a generalized regression neural network (GRNN) as the meta-model. Note that the well-known stochastic approach—the GA can also be implemented as we mentioned above. The entire procedure (SAOO + GRNN) is described in the steps shown in Fig. 1. A set of N_0 initial configurations are sampled

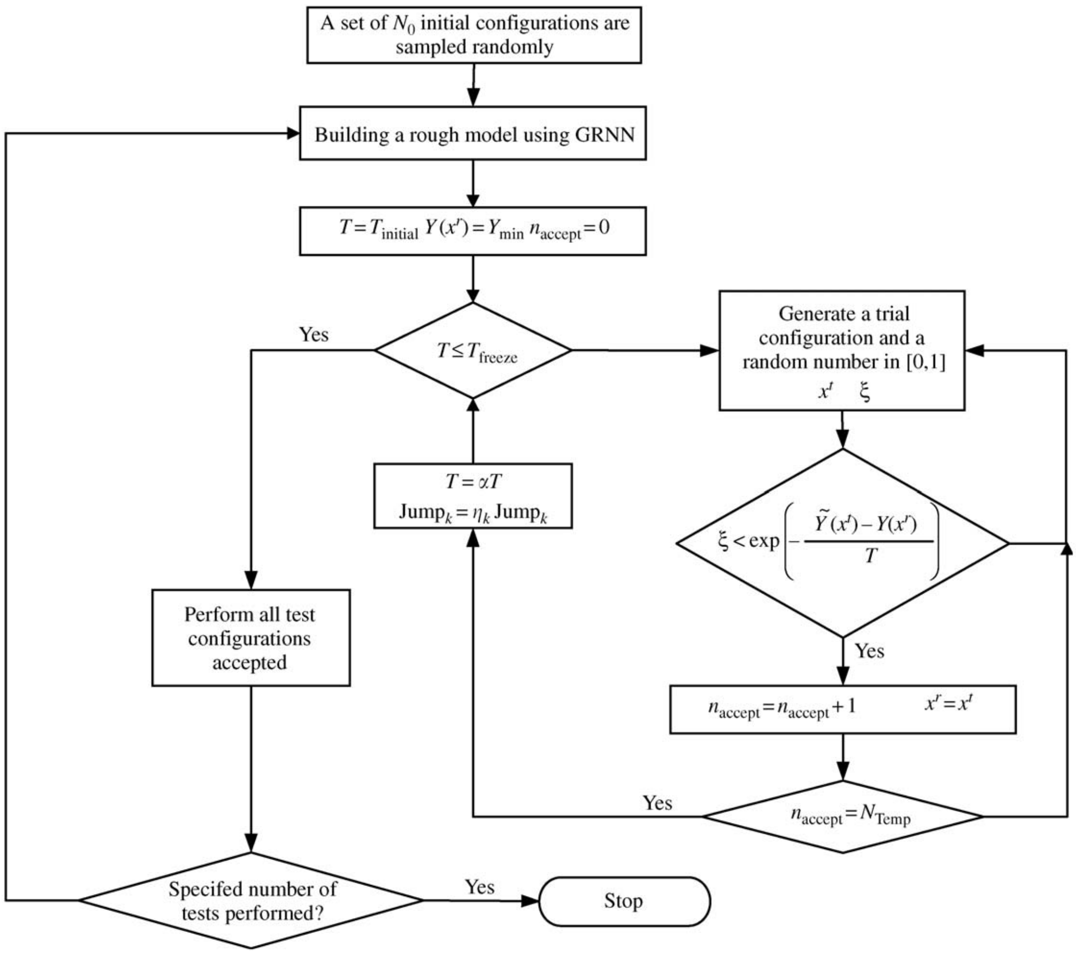

Fig. 1. Flowchart of the SAOO + GRNN method.

randomly. The data are trained to generate a ‘‘rough’’ model using GRNN. A SA search is performed using this ‘‘rough’’ GRNN model. Note that GRNN model is a simple regression of the existed model as described below. At each temperature, trial configurations are generated and accepted according to the transition probability in Equation 1 but with the fitness value predicted by the ‘‘rough’’ GRNN model. N_{Temp} configurations are accepted at each temperature. The annealing schedule is adjusted so that N_{Test} configurations are accepted when the annealing process iscompleted and the annealed temperature T_{freeze} is reached. The actual objective functions of these test configurations are evaluated. These new data are then used to update the GRNN model. The modeling-selection-testing cycle is repeated until a specified maximum number of tests have been carried out.

The details of algorithm SA adopted in this study is shown in Fig. 2. The procedure is similar to the general procedure given by van Laarhoven and Aarts (1987); the Cauchy annealing schedule is employed.

In order that existing data can be efficiently modeled, the generalized regression network was used (Specht, 1991). GRNN is based on the estimation of probability density function from observed samples by Parzen window estimator. Give a number of around existing measurements (x^i, Y^i), $i = 1, \ldots, M$, the expected value of the output at x is given by

$$\tilde{Y}(x) = \frac{\sum_{i=1}^{M} \exp\left[-(x - x^i)^T S^{-1}(x - x^i)\right] Y^i}{\sum_{i=1}^{M} \exp\left[(x - x^i)^T S^{-1}(x - x^i)\right]} \quad (2)$$

where S is the smooth/bandwidth factor matrix of the GRNN. GRNN requires no training, only a one-pass input of all training data is required. Both integer and continuous inputs can be used in GRNN although different values of S may be needed. The predictions of GRNN always fall within the maximum and minimum of the training data set. Use of GRNN avoids unreasonable extrapolation.

2.2. *Blind pick (BP) and Horse Race (HR + GRNN)*

To demonstrate the effectiveness of our approach, results are compared with blind pick search (BP) and an ordinal optimization using horse race as selection rule and GRNN as the ‘‘rough’’ model (HR + GRNN). In the BP, feasible solutions are randomly generated. The HR + GRNN procedure is described in the flow chart shown in Fig. 3. A set of N_0

Set temperature equals to an initial temperature $T = T_0$, Jump=1;
Select a feasible configuration as the current configuration $\vec{x}$ and calculates in objective function $Y(\vec{x})$
For $T > T_{\text{freeze}}$ perform this following.
  For $n < N_{\text{Temp}}$:
    Select a feasible configuration as the current configuration $\vec{x}'$ so that
    $$x_k' = \begin{cases} x_k + \text{Jump} * r * (UB_k - x_k) & r < 0 \\ x_k + \text{Jump} * r * (x_k - LB_k) & r > 0 \end{cases} \qquad r = \text{random}[-1, 1]$$
    Calculate the objective function value $Y(\vec{x}')$ and $\Delta = Y(\vec{x}') - Y(\vec{x})$
    If $\Delta \le 0$
      $\vec{x} = \vec{x}'$
    else
      Generate a random number χ distributed uniformly between [0, 1]
      If $\chi > \exp\left(-\frac{\Delta}{T}\right)$
        $\vec{x} = \vec{x}'$
      end
    end
    $n = n + 1$
  end
  $T = \alpha T$, $\alpha \le 1$
  $\text{Jump}_k = \eta_k \cdot \text{Jump}_k \quad \eta_k \le 1$
end

Fig. 2. A SA algorithm.

1. A set of N_{initial} initial configurations are sampled randomly
2. The data are trained to generate a "rough" model using GRNN
3. A new set of 20 test configurations are selected using horse race selection rule
   i. 1000 trial configurations were generated randomly
   ii. Rank all trial configurations using GRNN
   iii. Use the top 20 configurations as test configurations
4. Obtain the actual objective functions for test configurations
5. If a maximum number of tests is exceeded, stop; otherwise, go to step 2

Fig. 3. Flowchart of the HR + GRNN method.

initial configurations are sampled randomly. The data are trained to generate a ''rough'' model using GRNN. N_{Trial} trial configurations were generated randomly and ranked using GRNN predictions. The top N_{test} configurations were tested by evaluating the true objective function. The new sampled data were combined with existing data to update the GRNN model. The modeling-selection-testing cycles were repeated until a specific number of tests have been performed.

3. The trim-loss problem

The trim-loss problem appears when a set of ordered product reels are to be cut from raw paper reels or other reels with specified widths. The cutting process is simply a winding process, where the raw paper is wound through the slitter and cut by a set knives positioned on the line, see Fig. 4. The product widths can rarely be combined to the exact raw paper width, therefore waste appears during the cutting process. The main objective is to minimize the trim loss while demand specifications are satisfied.

A raw paper roll with a width $B_{\max}$ must be cut to satisfy the following order specifications. There are

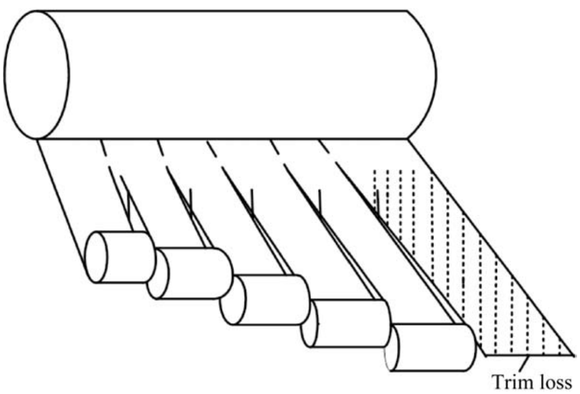

Fig. 4. A schematic illustration of the trim-loss problem.

$i = 1, \ldots, I$ different products, $n_{i,\text{order}}$ rolls of product i with a width b_i must be cut. All product rolls are assumed to be equal length. In order to identify the best overall scheme, a maximum number of $j = 1, \ldots, J$; different cutting patterns is postulated. A pattern is defined by the position of the knives. The number of repeats of pattern j is given by the integer variable m_j. The existence of a product in a given pattern is denoted by an integer variable n_{ij}. Another binary variable y_j is introduced to a change in pattern. If a new pattern is introduced $(m_j > 0)$, then y_j is equal to 1. A sample cutting pattern is shown in Fig. 5.

The actual cost of the trim loss is the total amount of raw materials used, that is, the sum of all repeated patterns multiplied by a cost factor C, in addition to the cost of changing knife positions between patterns. Let the pattern change be weighted by a coefficient c_j, thus a trim loss problem is solve the following minimization problem:

$$\min_{m_j, y_j, n_{ij}} \sum_{j=1}^{J} \left(C \cdot m_j + c_j \cdot j \cdot y_j \right) \tag{3}$$

restricted by the following constraints:

(1) The number of rolls of each product must be greater than customers' order:

$$\sum_{j}^{J} m_j \cdot n_{ij} \geq n_{i,\text{order}}, \quad i = 1, \ldots, I \tag{4}$$

(2) The width of each pattern must be less than the width of the raw paper roll:

$$\sum_{i=1}^{I} b_i \cdot n_{ij} \leq B_{\max} \quad j = 1, \ldots, J \tag{5}$$

(3) There must be at least one product in a pattern

$$\sum_{i=1}^{I} n_{ij} \geq y_j \quad j = 1, \ldots, J \tag{6}$$

(4) The total number of knives is limited to $N_{i,\max}$:

$$\sum_{i=1}^{I} n_{ij} \leq y_j \cdot N_{i,\max} \quad j = 1, \ldots, J \tag{7}$$

(5) There must be at least one pattern after a knife change

$$m_j \geq y_j \quad j = 1, \ldots, J \tag{8}$$

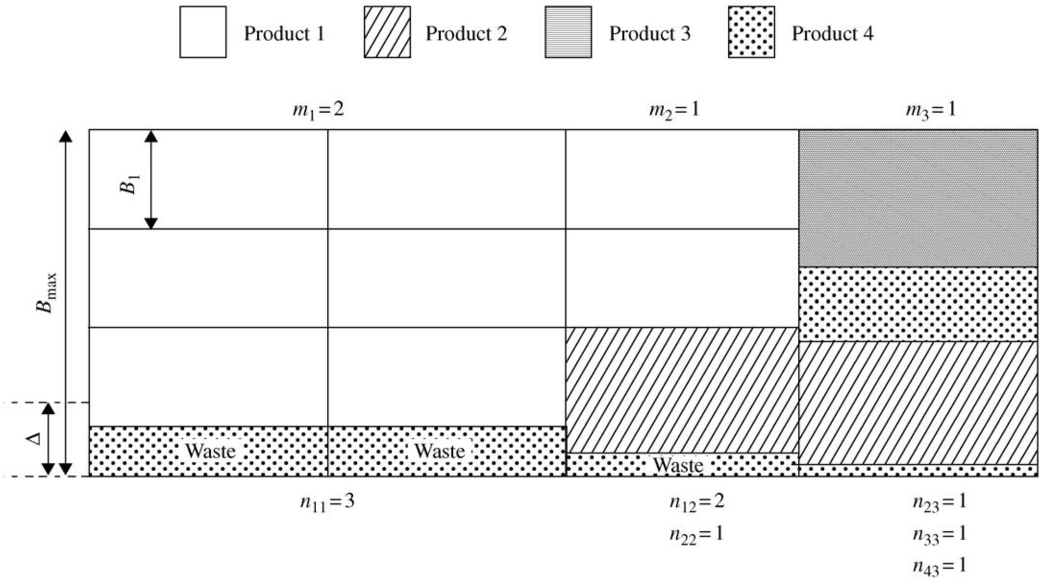

Fig. 5. The cutting pattern.

(6) The maximum number of pattern repetitions is limited to M

$$m_j \leq M \cdot y_j \quad j = 1, \ldots, J \tag{9}$$

(7) The width of cut product in each pattern must exceed a certain minimum Δ:

$$\sum_i^I b_i \cdot n_{ij} \geq (B_{\max} - \Delta) y_j \quad j = 1, \ldots, J \tag{10}$$

The example (Floudas *et al.*, 1999) presented in Table 1 is used. The problem is to satisfy the following order $b = (290, 315, 350, 455)$, with $n = (15, 28, 21, 30)$. The maximum width of a cutting pattern was $B_{\max} = 1850$ mm with a tolerance of $\Delta = 100$ mm. The maximum number of knives was $N_{\max} = 5$. The number of cutting patterns was taken form the number of products: $J = I = 4$, the cost of a pattern was determined to be $C = 1$ and the cost of a knife change was $c_j = 0.1$. m_j ranges from 0 to 30, y_j are either 0 or 1, and n_{ij} ranges 0–5. The size of the search space is $31^4 \times 2^4 \times 6^{16}$ of the order of 10^{19}. An objective function of 19.6 was obtained by Floudas *et al*. (1999).

Table 1. Parameters of the trim loss sample problem

I	4
(b_1, b_2, b_3, b_4)	(290, 315, 350, 455)
(n_1, n_2, n_3, n_4)	(15, 28, 21, 30)
M	30
$B_{\max}$	1850 mm
Δ	100 mm
J	4
$N_{\max}$	5
C	1
c	0.1

4. Results

The results are shown in Table 2. Average and standard deviations are statistics of 10 runs. If blind pick (BP) is used, we found the average of fitness functions is 22.4 after 5000 tests. A HR + GRNN procedure improved the results of a BP procedure. However, the optimal solution cannot be reached in 5000 tests.

A simulated annealing search is applied to the problem. The following parameters are adopted: the Cauchy temperature decay ratio is $\alpha = 0.6$. The jump "*range*" of all integer variables is one. The jump decay ratio of the integer variables m, n are $\eta_{m,n} = 0.8$. The jump decay ratio of the binary variable y is $\eta_y = 1$. The upper bound and lower bounds are $\mathrm{UB}_{m_j} = 30$, $\mathrm{UB}_{y_j} = 1$, $\mathrm{UB}_{n_{ij}} = 5$, $\mathrm{LB}_{m_j} = 0$, $\mathrm{LB}_{y_j} = 0$, $\mathrm{LB}_{n_{ij}} = 0$. This SA procedure was able to find the global optimum if the annealing

Table 2. Search results of the trim loss problem

Search methods	*Performance evaluation*			
		Fitness obtained		
	No. of tests	*Avg.*	*Std.*	*Time*
BP	200	29.2	2.3	0.2
	500	25.7	3.0	0.7
	1000	25.6	2.2	2
	2000	23.5	1.9	12
	5000	22.4	0.8	82
HROO + GRNN	200	23.3	1.3	13
	500	21.9	0.9	59
	1000	21.1	0.4	148
	2000	20.7	0.6	359
	5000	20.0	0.2	1873
SA	200	22.5	0.9	27
	500	21.5	0.9	81
	1000	20.1	0.7	113
	2000	19.8	0.2	262
	5000	19.6	0	493
SAOO + GRNN	120	22.4	1.5	7
	140	20.5	0.9	17
	160	19.7	0.2	31
	180	19.6	0	48
	200	19.6	0	58

schedule is slow enough that 5000 tests were performed. If the annealing schedule is hastened so that only 2000 tests were performed, then the optimal solution may not reached. Simulated annealing is a much better heuristic based search procedure than purely blind pick. By combining simulated annealing with concepts in ordinal optimization, we can increase the efficiency of SA by evaluating the objective functions of only those trial configurations that are high ranked by the ''rough'' model.

Using a SAOO + GRNN procedure, we found that global solution value of 19.6 is obtained within 200 tests. The Cauchy annealing parameter is the same as the above one. Two samples are taken at each temperature, and T_{freeze} is set so that 20 samples are accepted in each SA run. Since 100 tests were generated randomly to provide the initial model, the SAOO + GRNN reached the solution within five model-generation-testing cycles. More importantly, after about 180 tests, the SA + GRNN procedure actually converged since all runs produce the same solution.

Since model predictions were made and tests were

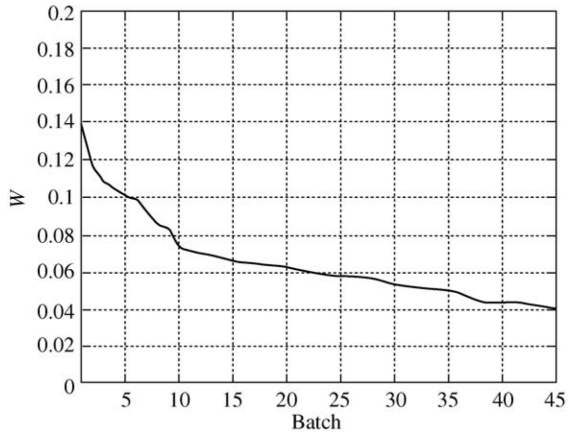

Fig. 6. The evolution of maximum model error *W* for HR + GRNN.

performed, we were able to calculate *W*, the error of our model as:

$$W = \frac{\max|Y - \tilde{Y}|}{Y_{\max} - Y_{\min}} \tag{11}$$

where during a HR + GRNN (Fig. 6) and SAOO + GRNN (Fig. 7) search. We can see that the model error is very small and decreases as more data are accumulated. Moreover, we found that the typical OPC curve of this trim-loss problem is of the Bell-shaped type (Fig. 8). According to Lau and Ho (1997), when 100 samples are selected from a set of 1000 randomly, only two are expected to be among the top 50. However, with the help of a rough model

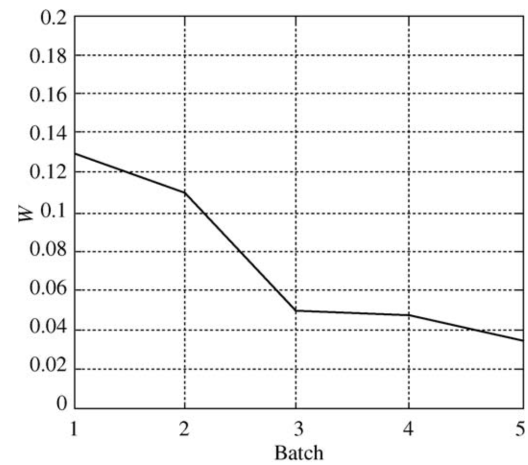

Fig. 7. The evolution of maximum model error *W* for SAOO + GRNN.

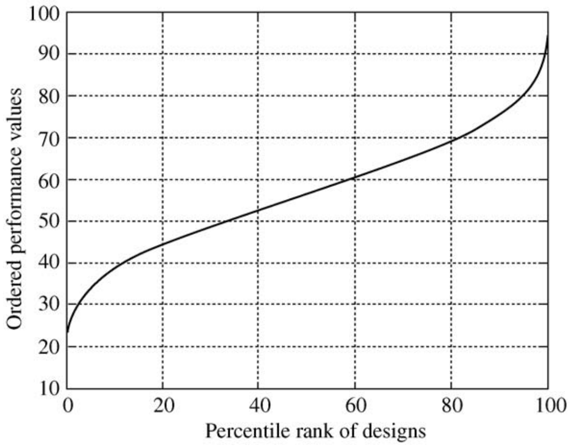

Fig. 8. Ordinal performance curve of the trim-loss problem.

$(W = 0.5)$ to rank the candidates, the number of alignment in the top 50 is increased to nine. Since the GRNN model is an adequate model $(W \lesssim 0.1)$, we can expect an increase in alignment of at least one order of magnitude. Indeed, the SAOO + GRNN method requires only about 1/10 of the function calls required by SA.

By doing so, we traded off number of tests (actual function calls) with time required in modeling construction and selection. Table 2 also compared CPU time expended to solve this problem on an IBM PC with Pentium 4 processor of 1.6 GHz. The time required for SAOO + GRNN to find the global optimum solution is less than 59 s while the time required for SA is more than 262 s. Such a trade-off is obviously very worthwhile.

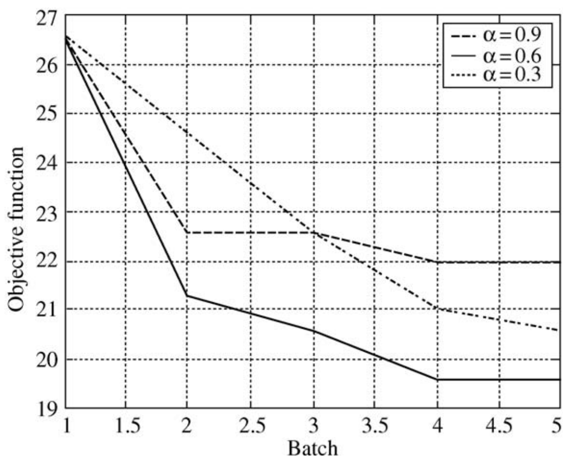

Fig. 9. Effect of annealing rate.

Figure 9 illustrates how the efficiency of the SAOO + GRNN procedure is affected by the annealing schedule. If the annealing rate is too slow, $\alpha = 0.9$, the procedure is no different from a blind pick procedure. The convergence rate is relatively slow. If the annealing rate is too fast, $\alpha = 0.3$, too much trust is put into the model, the solution procedure is easily trapped in a local optimum. The procedure is more like a horse race selection rule.

5. Conclusions

The optimization approach we proposed integrates concepts in ordinal optimization with SA. In the solution of large scale optimization problem by heuristic based method, it is of paramount importance to sample as many important candidates (potentially good solutions) as possible. Ordinal optimization theory explained that during this candidate selection process, the estimation and ranking of ''potential fitness'' does not have to be very accurate. We suggested that this can be done by a purely empirical ''rough'' model, the generalized regression neural network, which memorizes all existing data and performs smoothing interpolation and extrapolation. By performing SA search on this model, a good enough set trial configurations are generated. Actual objective function evaluations are required only for a small fraction of the top-rank candidates in this set to ensure that some good candidates are actually sampled. Using this approach, we found that the trim-loss problem arising from paper-cutting industry can be solved very effectively.

Acknowledgments

The authors thank the financial support from National Science Council, Taiwan and Top Information System Co. for this work through the grant NSC91-2622-E007-006.

References

Collins, N. E., Eglese, R. W. and Golden, B. L. (1988) Simulated annealing—an annotated bibliography. *American Journal of Mathematical and Management Sciences*, **8**(3), 209–307.

Davis, Mark E. (1999) Combinatorial methods: how will they integrate into chemical Engineering? *AIChE Journal*, **45**(11), 2270–2272.

Edgar, T. F., Himmelblau, D. M. and Lasdon, L. S. (2001) *Optimization of chemical processes*. McGraw-Hill, New York.

Eglese, R. W. (1990) Simulated annealing: a tool for operational research. *European Journal operational research*, **46**, 271–281.

Falcioni, M. and Deem, W. (2000) Library design in combinatorial chemistry by Monte Carlo methods. *Physical Review E*, **61**(5), 5948–5952.

Floudas, C. A., Pardalos, P. M., Adjiman, C. S., Esposito, W. R., Gumus, Z., Harding, S. T., Klepeis, J. L., Meyer, C. A. and Schweiger, C. A. (1999) *Handbook of Test Problems for Local and Global Optimization*, Kluwer Academic Publishers, The Netherlands.

Foerster, H. and Wäscher, G. (1998) Simulated annealing for order spread minimization in sequencing cutting patterns. *European Journal of Operational Research*, **110**(2), 272–281.

Goulimis, C. (1990) Optimal solution for the cutting stock problem. *European Journal of Operational Research*, **44**, 197–208.

Harjunkoski, I., Westerlund, T., Isaksson, J. and Skrifvars, H. (1996) Different formula for solving trim loss problems in a paper-converting mill with ILP. *Computers and Chemical Engineering*, **20**, s121–s126.

Harjunkoski, I., Westerlund, T., Pörn, R. and Skrifvars, H. (1998) Different transformations for solving non-convex trim-loss problems by MINLP. *European Journal of Operational Research*, **105**, 594–603.

Ho, Y. C., Sreenivas, R. and Vakili, P. (1992) Ordinal optimization of discrete event dynamic systems. *Journal of Discrete Event Dynamic Systems*, **2**(2), 61–88.

Holland, J. H. (1975) *Adaptation in Natural and Artificial System*, Ann Arbor, The University of Michigan Press.

Johnston, D. S., Aragon, C. R. and McGeoch, L. A. (1989) Optimization by simulated annealing: an experiment evaluation: Part 1, graph partitioning. *Operational Research*, **37**, 865–892.

Kirkpatrick, S., Gelatt, Jr. C. D. and Vecchi, M. P. (1983) Optimization by Simulated Annealing. *Science*, **220**, 671–680.

Koulamas, C., Anotony, S. R. and Jean, R. (1994) A survey of simulated annealing application to operations research problems. *International Journal of Production Research*, **30**(1), 95–108.

Ku, H. M. and Karimi, I. (1991) An evaluation of simulated annealing for batch process scheduling. *Industrial and Engineering Chemistry Research*, **30**, 163–169.

Lau, T. W. E. and Ho, Y. C. (1997) Universal alignment probability and subset selection for ordinal optimization. *Journal of Optimization Theory and Applications*, **93**(3), 455–489.

Lee, J. T., Lau, E. and Ho, Y. C. (2001) The Witsenhausen counter example: a hierarchical search approach for non-convex optimization problems. *IEEE Transactions Automatic Control*, **46**(3), 382–397.

Luo, Y. C., Guignard, M. and Chen, C. H. (2001) A hybrid approach for integer programming combing genetic algorithms, linear programming and ordinal optimization. *Journal of Intelligent Manufacturing*, **12**(5), 509–519.

Östermark, R. (1999) Solving a nonlinear non-convex trim loss problem with a genetic hybrid algorithm. *Computers and Operations Research*, **26**(6), 623–635.

Painton, L. A. and Diwekar, U. M. (1994) Synthesizing optimal design configurations for a brayton cycle power plant. *Computers and Chemical Engineering*, **18**(5), 369–381.

Specht, D. F. (1991) A general regression neural network. *IEEE Transactions on Neural Network*, **2**(6), 568–576.

Tsujimura, Y. and Gen, M. (1998) Entropy-based genetic algorithm for solving TSP. *Proceedings of Second International Conference on Knowledge Based Intelligent Electronic System*, April, Adelaide, Australia.

Van Laarhoven, P. J. M. and Aarts, E. H. L. (1987) *Simulated Annealing: Theory and Applications*, Rediel, Dordrecht, The Netherlands.

Wäscher, G. (1990) An LP-based approach to cutting stock problems with multiple objectives. *European Journal of Operational Research*, **44**, 175–184.

Westerlund, T. and Isaksson, J. (1998) Some efficient formulations for the simultaneous solution of trim-loss and scheduling problems in the paper-converting industry. *Transactions of the Institute of Chemical Engineers*, **76**(A6), 677–684.

Westerlund, T., Pettersson, F. and Grossmann, I. E. (1994) Optimization of pump configurations as a MINLP problem. *Computers and Chemical Engineering*, **18**, 845–858.

Zhang, L., Wang, L. and Tang, F. (2002), Order based genetic algorithm for flow shop scheduling. *Proceedings of First International Conference on Machine Learning and Cybernetics*, November, Bejing, China.