Optimal Multiloop Feedback Design Using Simulated Annealing and Neural Network

Julian Jeng Lin Lin and David Shan Hill Wong
Dept. of Chemical Engineering, National Tsing Hua University, Hsinchu, Taiwan 300, R.O.C.

Shuh Woei Yu
Dept. of Chemical Engineering, National Central University, Chung Li, Taiwan 320, R.O.C.

The essential design and analysis element of most control systems is a mathematical representation of the process dynamics. This model dependence may create serious problems when the process under consideration exhibits strong nonlinear behavior. For instance, Luyben (1987) reported that normal linearization and identification techniques may not yield suitable linear transfer functions for high-purity distillation columns. Despite recent developments in nonlinear control systems, the dominant force of control in the chemical process industries will continue to be conventional PID controllers (Bequette, 1991). The majority of the processing units still rely on a combination of feedbacks for control. This multiloop environment and robust stability requirement can make the design of simple feedbacks a formidable task. In addition, performance of this linear control scheme could become limited due to system uncertainty and nonlinearity. Certain attempts have been made to counter the effects of process nonlinearity on feedback control in nonlinear gains (Luyben, 1989) or gain scheduling (Tsogas and McAvoy, 1985). However, efficient and optimal design of multiloop feedbacks for nonlinear processes remains elusive for most control practitioners.

There has been a growing interest in using an artificial neural network (ANN) as a tool in process knowledge representation due to its learning and complexity handling ability. Diverse applications of ANN in process systems engineering include nonlinear modeling and control and many other areas (Bakshi and Stephanopoulos, 1993). A substantial amount of input/output data is generally required for training ANN. While abundance of data is usually assumed in pattern recognition problems, the cost and risk of obtaining necessary data in constructing an ANN model for a process plant cannot be overlooked. To alleviate the burden of data requirement, a response surface management scheme should be developed. Simulated annealing (SA) is a Markov chain based optimization strategy originated from statistical mechanics (Kirkpatrick et al., 1983). It has been used to solve various kinds of optimization problems in chemical engineering (Patel et al., 1991). Lin and Wong (1993) recently proposed that simulated annealing could be used as a response surface management strategy for constructing an ANN model.

In this study, ANN is utilized directly for feedback control system design. A neural network is trained on-line to learn the complex and highly nonlinear relationship between tuning parameters, process upsets, and the summation of time-weighted absolute error integrals of the quality control loops of a simulated high-purity distillation column. With the learning process directed by SA, the number of test runs can be greatly reduced. Optimal tuning parameters are obtained using the ANN model. Satisfactory control of both end products of the binary column is maintained with the proposed design technique.

Design Algorithm

Artificial neural network models for ITAE

Performance of dual-composition control of a distillation unit can be directly evaluated via summation of the integral of time-weighted absolute error of each loop:

$$\text{ITAE} = \sum_{i=1}^{2} \int_0^{\infty} |e_i(t)|\, t\, dt \tag{1}$$

For a given type of disturbance or set point change, optimal design of the feedback controllers basically involves the search for a set of tuning constants that would minimize the objective function defined above. In a multiloop environment the relation between loop constants and error integrals is quite complex and cannot be possibly expressed in any conventional mathematical form. ANN should be used instead. For simplicity, a back-propagation neural network is used together with a modified generalized delta rule proposed by Rigler et al. (1991) in this study. It is assumed that an ANN of one input layer (with 12 nodes), one output layer (with one node) and three hidden layers of 15 nodes is of sufficient complexity for our sample problem.

Correspondence concerning this article should be addressed to S. H. Wong.

Due to the large number of independent variables, a significant amount of data is required to identify functional forms for these relations that are quantitatively reliable to be used for optimization. Data acquisition by brute force approach is simply unrealistic. A systematic and effective approach must be employed. It is proposed in this work that simulated annealing be used to direct the data acquisition procedure.

Basic theory of simulated annealing

The basic idea of optimization by SA is to generate a Markov chain of test states using the Metropolis algorithm (Metropolis et al., 1953):

$$P_{ij} = \begin{cases} 1 & \text{if } C(\vec{x}_j) \le C(\vec{x}_i) \\ \exp\left[-\dfrac{C(\vec{x}_j) - C(\vec{x}_i)}{T} \right] & \text{otherwise} \end{cases} \tag{2}$$

where P_{ij} is the transition probability between state i and state j. In statistical mechanics, temperature is a measure of the kinetic energy. When the temperature is high, the system has enough kinetic energy to rumble around all possible states. At a lower temperature, the system will be "trapped" at a local minimum since the system does not have enough kinetic energy to overcome the "barrier." The object of finding a global optimum is a good "annealing strategy" (that is, a temperature reduction program that allows the system to survey enough possible states but rapidly settle into the global minimum). In this work, we have adopted the annealing strategy proposed by Dekkers and Aarts (1991). The "temperatures" of successive Markov chains are reduced by:

$$T' = T\left[1 + \frac{T \ln(1+\delta)}{3\sigma(T)} \right]^{-1} \tag{3}$$

where δ is an arbitrary parameter that controls the rate of convergence, and $\sigma(T)$ is the standard deviation of the energy of the states of the Markov chain at temperature T. In this study, δ is set equal to 0.5.

The stopping criterion of the annealing procedure is governed by:

$$\left| \frac{d\langle C \rangle_{T,s}}{dT} \frac{T}{\langle C \rangle_{T_0}} \right| \le \epsilon = 10^{-4} \tag{4}$$

where $\langle C \rangle_{T,s}$ is the average of energy of the states smoothed over last five Markov chains. Its derivative with respect to temperature, in statistical mechanical terms, is a measure of the fluctuation of energy of the states. $\langle C \rangle_{T_0}$ is the average of energy of states in the first Markov chain.

Model construction and optimization algorithm

The procedure proposed by Lin and Wong (1993) that uses simulated annealing to direct the data acquisition process is employed here to develop an ANN model for ITAEs. The "temperatures" of the Markov chains in simulated annealing thus become a measure of the trust one can put in the current model. When there are relatively few available data points, "temperatures" are high. SA annealing calls for additional experiments in a nearly random manner. As information accumulates, greater trust is placed in the models and "temperature" decreases. Transition to states with a higher cost function value becomes more and more unlikely. Simulated annealing directs new experiments toward the global optimum. The model construction and optimization procedure consists of the following steps:

(1) For a given disturbance or set point change, a number of initial closed-loop experiments are performed with arbitrarily selected values of (K_{C1}, τ_{I1}, K_{C2}, τ_{I2}).

(2) An ANN model for ITAE is constructed using input/output data of these on-line tests. A cost function is defined as:

$$C(K_{C1}, \tau_{I1}, K_{C2}, \tau_{I2}) = (\text{ITAE})_{\text{ANN}} \tag{5}$$

(3) An initial value of the "temperature" is determined by giving the Markov chain a 90% chance of overcoming the maximum difference in cost function among these experimental data.

(4) The global optimal of the ANN cost function surface is found by direct application of the SA method proposed by Dekkers and Aarts (1991).

(5) Starting with this global optimal as the initial state, the next experiment is performed at an operating state selected using Markov chain transition probability estimated by the ANN model.

(6) Update the ANN model with new data.

(7) Steps 5 and 6 are repeated until the Markov chain is truncated.

(8) If the stopping criterion of SA, Eq. 4, is satisfied, the model construction procedure is terminated. Otherwise the "temperature" is reduced and steps 4 to 7 are repeated until annealing.

Results and Discussion

The concept is illustrated in this study of using a simulated annealing and artificial neural network as a systematic trial-and-error testing and knowledge accumulation procedure in tuning the controllers of a high-purity distillation column. The rigorous tray-by-tray dynamic simulation model for a column, which separates methanol and water with overhead and bottom product purity specified at 99.9 mol % and 0.1 mol % with reflux and boilup rates as manipulated variables, is used as the "test plant." Detailed information concerning this column can be found in Georgiou et al. (1988).

For illustration purposes, application of SA + ANN in optimal design of two PI controllers for a decrease of 0.003 mole fraction in overhead product purity is described in detail. The proposed SA + ANN procedure required 76 closed-loop experiments to find an optimal set of tuning constants at (−27.5, 58.7, 30.0, 57.5) with ITAE equals to 9.66. Contours of ITAE at different K_{C2} and τ_{I2} with $K_{C1} = -27.5$, $\tau_{I1} = 58.75$ are presented in Figures 1a through 1d. Similarity between Figure 1a (obtained using rigorous simulation) and Figure 1b (learned by SA + ANN) indicates that SA + ANN had indeed learned the essence of the functional relation between ITAEs and the tuning parameters. If brute force approaches, such as conducting approximately same number of experiments at evenly or randomly distributed locations over a range of possible K_Cs

and τ_Is, the ANN models obtained are quite different from the actual ITAE surface (Figures 1c and 1d). And the tuning sets obtained are much inferior (Table 1). These results clearly demonstrate the merits of using simulated annealing as a response surface management tool. Evolution of the ANN model along the SA procedure is shown in Figure 2. Initially, the actual ITAE values at optimal locations predicted by the ANN are quite different from model values. This disparity, normally occurring for $\log(1/T)$ less than -1, gradually dissipated as more relevant information is accumulated by SA. The SA + ANN approach also compares favorably to direct gradient search procedure such as GRG2 which yields a slightly better tuning set, but needs 179 function evaluations.

Table 1. Comparison of SA + ANN to Brute Force Data Acquisition Approaches and Direct Optimization Methods

Design Method	No. of Data	K_{C1}	τ_{I1}	K_{C2}	τ_{I2}	ITAE
ANN + SA	76	−27.5	58.7	30.0	57.5	9.66
ANN + Even Placement	81	−39.3	59.9	0.01	39.9	1,497
ANN + Random Placement	81	−33.7	55.1	29.8	83.9	13.8
GRG2	179	−32.8	57.8	25.0	64.8	9.25

For comparison, results of conventional multiloop robust design techniques for the same column (Wang and Yu, 1993)

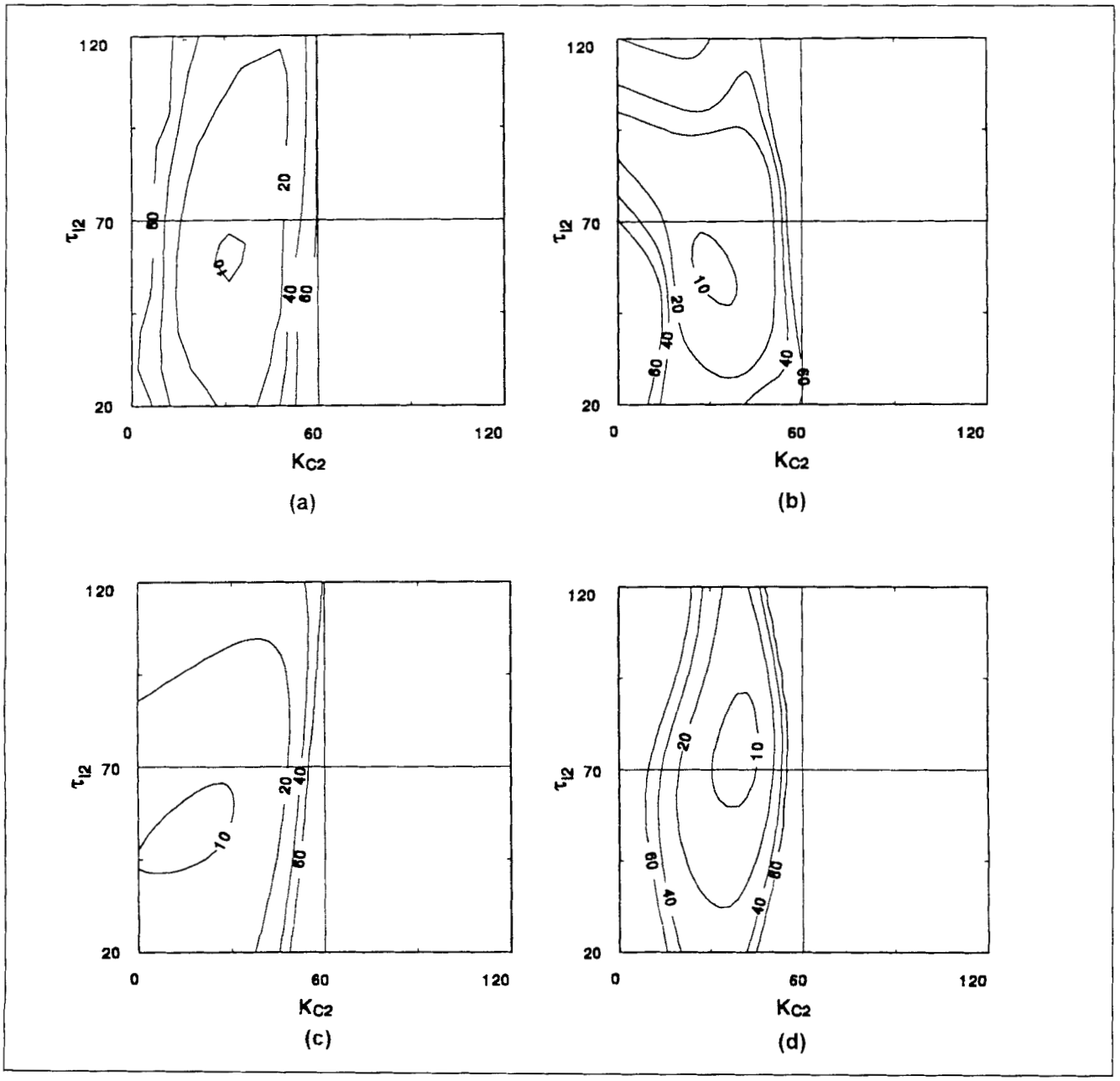

Figure 1. Contours of ITAE at $K_{C1} = -27.5$, $\tau_{I1} = 58.75$.

(a) "Actual plant"; (b) ANN obtained using SA; (c) ANN obtained using 81 evenly distributed data; (d) ANN obtained using 81 randomly distributed data.

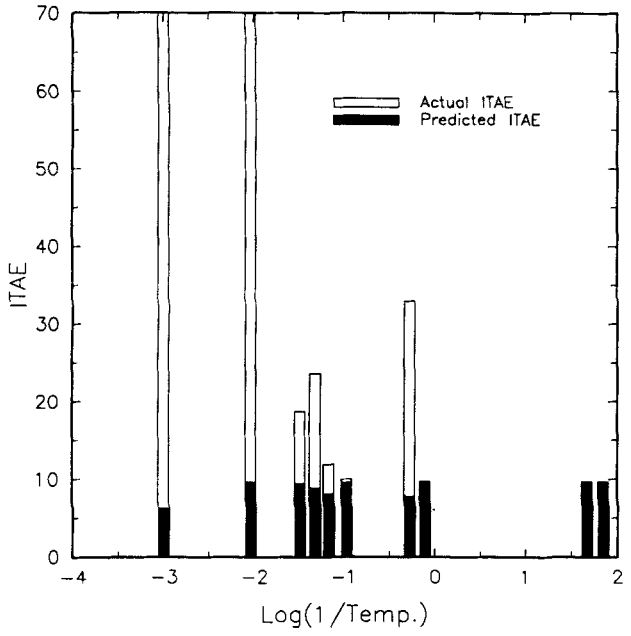

Figure 2. Evolution of ITAE at the optimal tuning predicted by ANN + SA.

are listed in Table 2. All methods investigated base the controller design on a nominal model obtained by ATV (Luyben, 1987) with different model uncertainty estimation criteria. While obtaining the nominal model may be tedious, reliable estimation of all possible uncertainties is even more difficult and depends greatly on the stability criterion used. If the range of uncertainty is overestimated, the design becomes excessively conservative and results in poor control. Insufficient robustness consideration, on the other hand, will lead to plant upset. Table 2 clarifies superiority of the proposed SA + ANN design technique. Dynamic responses with tuning parameters determined by SA + ANN, SSV and BLT for a decrease in overhead product purity are shown in Figure 3. Optimal control of both end products with the SA and ANN design is evident.

It is certainly logical to expect the SA + ANN procedure in providing the best control for x_D set point change from 0.999 to 0.996 mol fraction since this was a predetermined learning process. However, adequate control of this nonlinear column can still be achieved with the same set of tuning constants for other x_D set points changes even beyond the original training condition (such as, from 0.999 to 0.9992 mol fraction), as shown in Table 2. The ITAEs presented in Table 2 to a certain extent also confirm robustness of the designed control system despite the fact that robust stability and performance considerations are not included in the SA + ANN learning scheme.

Table 2. Comparison of Performance and Robustness of the SA + ANN Tuning and Other Design Methods for Set Point Changes in x_D

Design Method	0.999 to 0.996	0.999 to 0.998	0.999 to 0.9992	0.998 to 0.996
SA + ANN	9.66	4.50	3.81	4.72
SVA	13.2	7.16	6.40	7.62
SSV	12.5	4.64	2.36	6.80
HSE	22.0	3.20	1.19	11.8
BLT	42.6	13.1	4.07	13.1

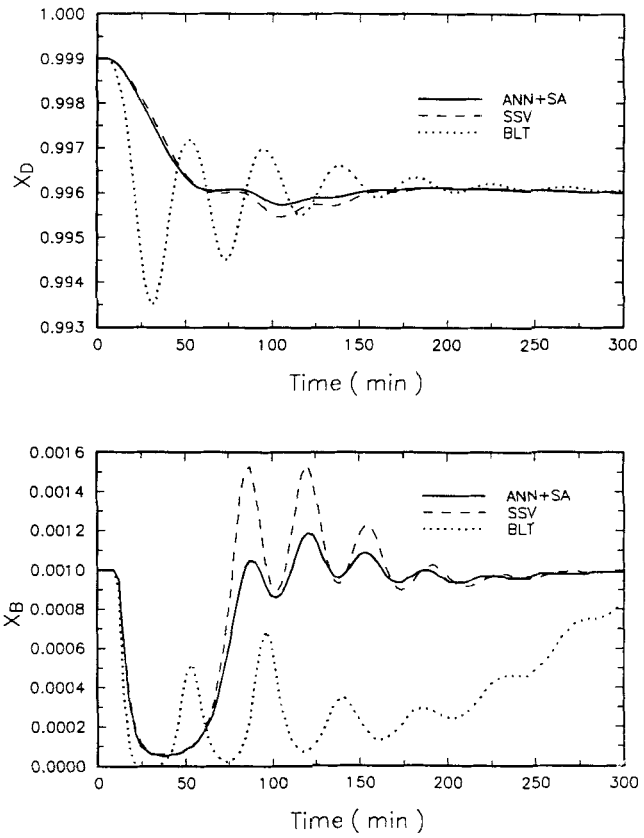

Figure 3. Closed-loop responses using tuning parameters obtained by ANN + SA, SSV and BLT for x_D set point change from 0.999 to 0.996 mol fraction.

A systematic and empirical procedure of elucidating the complex relation between performance of quality control loops of a high-purity distillation column, tuning constants, and set point changes using ANN and SA is demonstrated in this article. A robust and optimal set of tuning parameters can be obtained using an ANN model with minimal amount of closed-loop test data. Potential of the proposed design technique in handling column load disturbances or even unmeasurable disturbances is currently being investigated.

Notation

C = cost function
$\langle C \rangle_{T,s}$ = smoothed average of the energy of the states over five temperature levels
$C(\vec{x}_i)$ = energy at state i
e_i = error of loop i
K_{Ci} = proportional gain of loop i
t = time, minutes
T = temperature in annealing schedule
T' = new value of temperature in next Markov chain
x_B = liquid model fraction of light component in bottom product
x_D = liquid model fraction of light component in overhead product

Greek letters

δ = distance parameter that controls the temperature reduction rate

ϵ = stop parameter
τ_{Ii} = integral time constant of loop i, minute

Acronyms

ATV = autotune variation
BLT = biggest log-modulus tuning
HSE = highly structured E contours
SSV = structured singular value
SVA = singular value analysis

Literature Cited

Bakshi, B. R., and G. Stephanopoulos, "Wave-Net: A Multiresolution, Hierarchical Neural Network with Localized Learning," *AIChE J.*, **39**, 57 (1993).

Bequette, B. W., "Nonlinear Control of Chemical Processes: A Review," *Ind. Eng. Chem. Res.*, **30**, 1391 (1991).

Dekkers, A., and E. Aarts, "Global Optimization and Simulated Annealing," *Math. Programming*, **50**, 367 (1991).

Georgiou, A., C. Georgakis, and W. L. Luyben, "Nonlinear Dynamic Matrix Control for Highpurity Distillation Columns," *AIChE J.*, **34**, 1287 (1988).

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

Lin, J. J. L., and D. S. H. Wong, "Empirical Model Construction for Process Optimization Using Simulated Annealing and Artificial Neural Network," *Chem. Eng. Sci.*, submitted (1994).

Luyben, W. L., "Derivation of Transfer Functions for Highly Nonlinear Distillation Columns," *Ind. Eng. Chem. Res.*, **26**, 2490 (1987).

Luyben, W. L., *Process Modeling, Simulation and Control for Chemical Engineers*, 2nd ed., McGraw-Hill, New York (1989).

Metropolis, N., A. W. Rosenbluth, and A. H. Teller, "Equation of State Calculations by Fast Computing Machines," *J. Chem. Phys.*, **21**, 1087 (1953).

Patel, A. N., R. S. H. Mah, and I. A. Karimi, "Preliminary Design of Multiproduct Noncontinuous Plants Using Simulated Annealing," *Comput. Chem. Eng.*, **15**, 451 (1991).

Rigler, A. K., J. M. Irvine, and T. P. Vogl, "Rescaling of Variables in Back Propagation Learning," *Neural Network*, **4**, 225 (1991).

Tsogas, A., and T. J. McAvoy, "Gain Scheduling for Composition Control of Distillation Columns," *Chem. Eng. Commun.*, **37**, 275 (1985).

Wang, S. J., and S. W. Yu, "Robust Feedback Design for Nonlinear High-Purity Distillation Column Control," *Comput. Chem. Eng.*, **17**, 897 (1993).

Manuscript received Nov. 1, 1993, and revision received Jan. 24, 1994.