Genetic Algorithm for Identification of Time Delay Systems from Step Responses

Gang-Wook Shin, Young-Joo Song, Tae-Bong Lee, and Hong-Kyoo Choi

Abstract: In this paper, a real-coded genetic algorithm is proposed for identification of time delay systems from step responses. FOPDT (First-Order Plus Dead-Time) and SOPDT (Second-Order Plus Dead-Time) systems, which are the most useful processes in this field, but are difficult for system identification because of a long dead-time problem and a model mismatch problem. Genetic algorithms have been successfully applied to a variety of complex optimization problems where other techniques have often failed. Thus, the modified crossover operator of a real-code genetic algorithm is proposed to effectively search the system parameters. The proposed method, using a real-coding genetic algorithm, shows better performance characteristics when compared to the usual area-based identification method and the directed identification method that uses step responses.

Keywords: Genetic algorithm, identification, step response, time delay system.

1. INTRODUCTION

A major characteristic of industrial systems is that there are various constraints, including long dead times and time constants, multi-variable, nonlinear, and non-minimum phase systems An effective measurement method for industrial systems is needed for system identification of long-range dead-time systems among system constraints. The performance and stability of time-delay systems are influenced by dead time. To solve the constraints, the smith predictor, which compensates for dead time and a matched model, was proposed.

In a good model case, the merits of the smith predictor include the ability of improving the performance and of disregarding the dead time caused by the characteristics of a closed-loop system. But, to apply the smith predictor effectively, the model must be matched with a real system. If the model is mismatched with the system, the smith predictor will be difficult to apply in the real world. Thus, system identification is very important in time-delay systems. [1-4]

The systems for identification could be divided into linear and nonlinear types. However, this study considers two types of linear models: FOPDT (First Order Plus Dead Time) system and SOPDT (Second Order Plus Dead Time) system which are mainly used in the process industry.

For system identification methods of real world industrial systems, the input signal can have significant influence on identification results. Generally, test signals include pseudo-random binary sequences pulses, steps, ramps, and sinusoidal functions. However, the step test of all these tests is the simplest, needs little equipment, and can even be performed manually. Also, because a step test can be easily implemented on programmable logic controllers (PLC) or distribute control systems (DCS), this study used the step test function.

In this paper, we reviewed usual identification method, which are area-based identification and direct identification for time-delay systems. As well, we proposed an identification method using genetic algorithms that is better than the conventional identification methods for FOPDT and SOPDT systems.

2. SYSTEM IDENTIFICATION

Manuscript received November 27, 2005; revised May 8, 2006; accepted August 24, 2006. Recommended by Editorial Board member Eun Tai Kim under the direction of Editor Jin Young Choi.

Gang-Wook Shin is with the Korea Institute of Water and Environment, Korea Water Resources Corporation, 462-1 Jeonmin-dong, Yuseong-gu, Daejeon 305-390, Korea (e-mail: gwshin@kwater.or.kr).

Young-Joo Song is with the Institute of Electrical Facility, Hongik University, 300 Sinan-ri, Jochiwon-eup, Yeongi-gun, Chungcheongnam-do 339-701, Korea (e-mail: zealouslove @hotmail.com).

Tae-Bong Lee is with the Dept. of Electronics and Informations, Kyungwon College, San 65, Bokjung-dong, Soojung-gu, Sungnam-shi, Kyunggi-do 461-702, Korea (e-mail: tblee@kwc.ac.kr).

Hong-Kyoo Choi is with the Dept. of Electrical Engineering, Hongik University, 300 Sinan-ri, Jochiwon-eup, Yeongi-gun, Chungcheongnam-do 339-701, Korea (e-mail: hkchoi @wow.hongik.ac.kr).

2.1. Area-based identification

The FOPDT model, which describes a linear monotonic process quite well for most long-range dead-time systems, has been applied as an approximation to such systems. This model is represented as (1) [5,6].

$$Y(s)=\frac{K}{Ts+1}e^{-Ls}U(s), \tag{1}$$

where $U(s)$ and $Y(s)$ are control inputs and process outputs, and K, T, and L are process gains, time constants, and dead times respectively.

A graphical method is the simplest method of all the identification methods. The intercept of the tangent to the step response that has the largest slope with respect to the horizontal axis gives L. T is determined from the difference between L and the time when the step response reaches the value of 0.63K. Further, the times at which the process output reaches 28% of K and 63% of K, respectively, are measured, and used to estimate T and L. These two methods are simple, but quite sensitive to large measurement noise. Thus, an area-based method has been proposed as having better estimation robustness [7].

The process gain K is obtained from the steady states of the process input and output. The average residence time is computed from the area in Fig. 1. Thus, the average residence time T_{ar} can be expressed as (2).

$$T_{ar}=\frac{A_0}{K}=\frac{\int_0^{\infty}[y(\infty)-y(t)]dt}{K} \tag{2}$$

The time constant and the dead time can be obtained by the area A_1 under the step response up to T_{ar}.

$$T=\frac{eA_1}{K}=\frac{e\int_0^{T_{ar}}y(t)dt}{K} \tag{3}$$

$$L=T_{ar}-T=\frac{A_0}{K}-\frac{eA_1}{K} \tag{4}$$

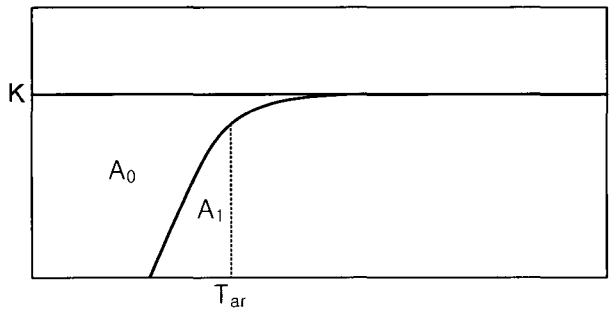


Fig. 1. Area-based method from step response.

The estimated accuracy of this method depends mainly on the area A_0. In order to increase the accuracy, one needs a testing span that is long enough for the process to enter a new steady state completely.

2.2. Direct identification

In the direct identification of FOPDT and SOPDT systems [8,9], the parameters of the model are obtained directly from a set of newly derived linear regression equations.

The direct identification supposes that a given system is in a zero initial state, before a step change with an amplitude of h at t=0 in the system input. Also, the system input and the resulting output response are logged until the process enters a new steady state.

For a system described by (1), its transient output after t=L is described by (5).

$$y(t)=hK(1-e^{-(t-L)/T})+\varepsilon(t), \tag{5}$$

where $\varepsilon(t)$ is the white noise in the measurement noise of $y(t)$.

Equation (5) is re-written for $t\geq L$ as

$$e^{-(t-L)/T}=1-\frac{y(t)}{hK}+\frac{\varepsilon(t)}{hK}. \tag{6}$$

Integrating output in (5) from $t=0$ to $t=\tau(\tau\geq L)$ yields (7).

$$\begin{aligned}\int_0^{\tau}y(t)dt&=hK(t+Te^{-(t-L)/T})\;\Big|_L^{\tau}+\int_0^{\tau}\varepsilon(t)dt\\ \int_0^{\tau}y(t)dt&=hK[\tau-L-T\frac{y(\tau)}{hK}]+[T\varepsilon(t)]\;\Big|_L^{\tau}+\int_0^{\tau}\varepsilon(t)dt\end{aligned} \tag{7}$$

Let $A(\tau)$ and $\delta(\tau)$ as (8).

$$A(\tau)=\int_0^{\tau}y(t)dt\;\;\delta(\tau)=[T\varepsilon(t)]\;\Big|_L^{\tau}+\int_0^{\tau}\varepsilon(t)dt \tag{8}$$

Equation (7) can be re-written as (9).

$$A(\tau)=hK[\tau-L-T\frac{y(\tau)}{hK}]+\delta(\tau)$$

or

$$\begin{bmatrix}h(\tau) & -h & -y(\tau)\end{bmatrix}\begin{bmatrix}K\\ LK\\ T\end{bmatrix}=A(\tau)-\delta(\tau). \tag{9}$$

By collecting (9) for all sampled $y(\tau)$ after $\tau\geq L$, a system of linear equations is obtained as

$$\Psi\theta=\Gamma\Delta \quad \text{for} \quad \tau\geq L, \tag{10}$$

where $\theta=[K \quad LK \quad T]^T$,

$$\Psi = \begin{bmatrix} hmT_s & -h & -y[mT_s] \\ h(m+1)T_s & -h & -y[(m+1)T_s] \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ h(m+n)T_s & -h & -y[(m+n)T_s] \end{bmatrix},$$

$$\Gamma = \begin{bmatrix} A[mT_s] \\ A[(m+1)T_s] \\ \bullet \\ \bullet \\ A[(m+n)T_s] \end{bmatrix}, \quad \Delta = \begin{bmatrix} -\delta[mT_s] \\ -\delta[(m+1)T_s] \\ \bullet \\ \bullet \\ -\delta[(m+n)T_s] \end{bmatrix}.$$

T_s is the sampling interval and $mT_s \geq L$. The estimation $\hat{\theta}$ of θ in (10) can be obtained using the least-squares method as

$$\hat{\theta} = (\Psi^T \Psi)^{-1} \Psi^T \Gamma. \tag{11}$$

3. GENETIC ALGORITHMS FOR SYSTEM IDENTIFICATION

3.1. Genetic algorithms

Genetic algorithms are global, parallel, and random search algorithms based on evolutionary theory, natural selection and the survival of the fittest, which Darwin first published in *On the Origin of Species* in 1895. In order to design a self-learning decision making system, John Holland first proposed genetic algorithms during modeling the biologic process of natural selection and individual genetics in 1975. Since genetic algorithms were applied by Goldberg to plants in 1989, the algorithms have been used as a new optimization method, replacing conventional optimization methods.

According to evolutionary theory, among the organisms living within a group under the same circumstance, the one having the most suitable genetic make-up for a given circumstance has the largest probability of survival and of evolving in a more suitable direction through crossover and mutation, while unfit organisms are gradually eliminated over the course of evolution.

As the evolution process repeats itself, the optimum organism is the dominant one. Genetic algorithms, a method that applies evolutionary theory to general optimization problems, distributes multi search points to search spaces and vests each search point with fitness according to the error of constraints and objective functions.

Larger fitness increases probability of inclusion in the next round of crossover and mutation, more search points having better genes form at the next stage, and the whole search process proceeds in the desired direction as he calculations continue.

The main characteristic of genetic algorithms is that it uses integer or binary coded strings converted from the actual values of parameters.

It is very simple and convenient to perform the crossover and the mutation because parameters of these string forms have the same form as the arrayed gene within the chromosome. Also, these strings can be effectively used for a hybrid optimization that uses integer or discrete parameters, because the strings have discrete quality.

While most of the optimization methods follow a local search course by moving from a point to a point, genetic algorithms simultaneously searches from the population of several grouped chromosomes. So, using information from a wider area, the convergence probability for the total optimal point is relatively larger than the existing optimization method.

Because a genetic algorithm uses a directed search method that uses objective functions and constraints and needs neither differential value nor other information, it is suitable of optimization under complex and various environments and is easily modified by changing the basic model [10,11].

Simple Genetic Algorithm (SGA) generally consists of individual initialization, fitness evaluation, and new population as shown in Fig. 2.

The way to represent points on a search space in SGA uses the binary encoding or the binary coding as most general methods.

According to this method, initialization methods, which make initial individual populations for simulated evolution, include random initialization and directed initialization based on prior knowledge or experience.

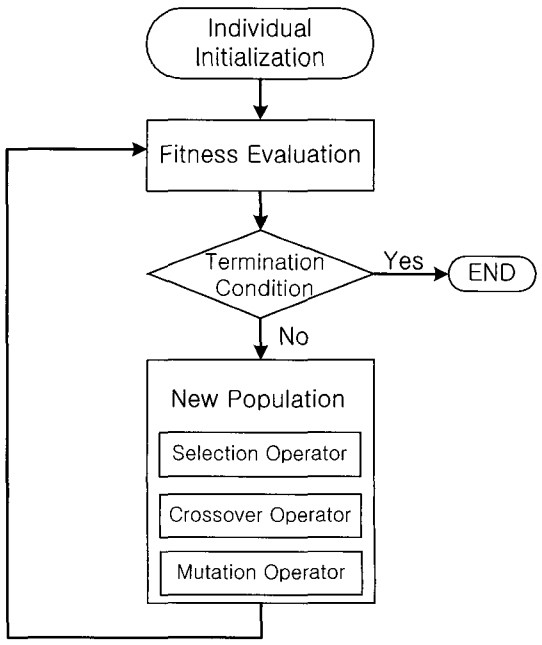


Fig. 2. Flow chart of simple genetic algorithm.

GA adaptability of changes is based on fitness evaluation of organisms, such as the adaptability of each life, is based on the circumstances of the natural world. Whenever a new population forms, the fitness of each individual is evaluated using objective functions.

To create new population we make new offspring for the next generation by selection, crossover, and mutation operations based on strings of chromosomes.

The selection operator is a procedure that reproduces the organism based on the fitness of each individual. In the natural world, offspring have genes transmitted from parents but their genetic composition differs from that of their parents. Imitation of the above procedure in the genetic algorithm yields the crossing procedure.

Mutation provides new information that is not previously in the present population. The optimization of SGA finishes whenever new organisms, through these procedures, satisfy proper convergence conditions or whenever proper iterations are completed [12].

3.2. Proposed real-coded genetic algorithms

3.2.1 Encoding

Actual industry problems often need high accuracy solutions, lack preliminary knowledge of solutions, and have complicated constraints.

A genetic algorithm needs much time to search for a solution and occasionally fails to find a solution, because long chromosomes makes for very large search spaces when one increases accuracy or lacks preliminary knowledge for the solution in a binary coded algorithm.

A real-coded algorithm is a typical expression that can solve the above problems by easily designing tools to treat constraints and for operators with knowledge related to the problems, providing a method that makes the chromosome expression approach closer to the solution space.

Also, compared to binary-coded, a real-coded algorithm needs no coding or decoding processes. Therefore, it can increase search speeds and fix very large domains in cases having no preliminary knowledge of the solution.

GA chromosomes could be generally represented as (12). The forms of each gene are binary in simple genetic algorithms and the real-number in real-coded genetic algorithm.

Also, drawbacks to the binary gene could be complemented by changing the binary to a real-number.

$$S_i = \{x_{i1} \quad x_{i2} \quad x_{i3} \quad \cdots \quad x_{in}\}, \tag{12}$$

where S_i is i-th chromosome of the population, x_{in} is n-th gene of i-th chromosome.

In some vector $x \in R^n$, the length l of chromosome S_i is as same as vector dimension n.

3.2.2 GA operators

Arithmetic crossover that relaxes discontinuity at the crossover point is used most frequently in real-coded crossover methods. However, the search space of the crossover can be reduced by limited ranges of the multiplier, which the main factor of the first combination between parent chromosomes.

So, (13) proposes a modified crossover operator to extend the multiplier range along with the fitness of parent chromosomes.

$$\hat{x}_1 = \lambda_s x_1 + (1-\lambda_s)x_2 \tag{13a}$$

$$\hat{x}_2 = \lambda_s x_2 + (1-\lambda_s)x_1 \tag{13b}$$

If chromosomes from the two parents are S_1 and S_2 and the relationship between their fitnesses, $f(S_1)$ and $f(S_2)$, is $f(S_1) \geq f(S_2)$, the range of λ_s could be set as (14). The initial search ability to seek for a global solution is increased by extending the search range of the chromosome which has a larger fitness.

$$-\frac{f(S_2)}{f(S_1)+f(S_2)} \leq \lambda_s \leq 1+\frac{f(S_1)}{f(S_1)+f(S_2)} \tag{14}$$

The cost function is calculated to minimize the variance between the step input response and the reference. Additionally, the fitness $f(S_i)$ evaluation is calculated by the cost function $J(S_i)$ of each chromosome S_i, and can be calculated by transforming into the maximization problem shown in (15).

$$f(S_i) = \frac{1}{J(S_i)+\alpha}, \tag{15}$$

where α is a constant with the fitness values always positive, α is set at 0.1 so as to not have it divided by 0 in this study.

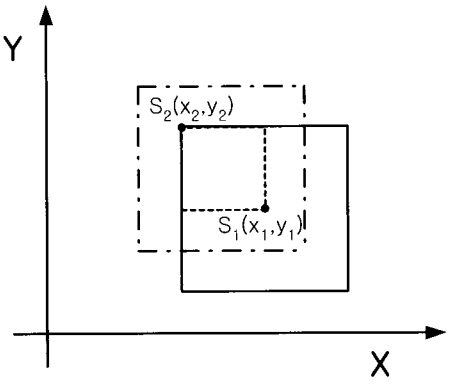


Fig. 3. Modified crossover.

4. SIMULATION

4.1. FOPDT system

The identification parameters of the FOPDT system are defined as three parameters, which are the genes of the GA chromosomes: system gain, time constant and dead time. Individual initialization is a method where each chromosome randomly takes a point in the domain to be a point in the proper area. In the FOPDT system identification, however, individual initialization randomly takes a point in the domain using the output response by the initial step input as follows.

$$\begin{aligned} &0.8 \times Y_{final} \le \Pr ocess\_Gain \le 1.2 \times Y_{final} \\ &0 \le Time\_Cons\tan t \le Rise\_Time/2 \\ &0 \le Dead\_Time \le (0.1 \times Y_{final}) \end{aligned} \tag{16}$$

To identify the FOPDT system, the detail design parameters of the proposed real-coded GA are represented in Table 1.

To verify the validity of system identification using GA, each of the two systems compares relationships among the results made by the three different identification methods: area-based identification, directed identification and proposed GA identification.

The systems considered for simulation are non-minimum phase and multi-lag systems.

To obtain the time-domain identification error of step input for each system, (17) is used to calculate the square means of error between the actual system output and the estimated output.

$$\varepsilon = \frac{1}{n+1}\sum_{k=m}^{m+n}[y(kT_s) - \hat{y}(kT_s)]^2, \tag{17}$$

where $y(kT_s)$ is the actual system output under a step change, and $\hat{y}(kT_s)$ is the response of the estimated system under the same step change.

Judging from the simulation results concerning error values for each system, response characteristics by area-based identification and those by directed identification were similar, but the results obtained by proposed identification were better than those by other methods.

Table 1. Parameters of real-coded GA.

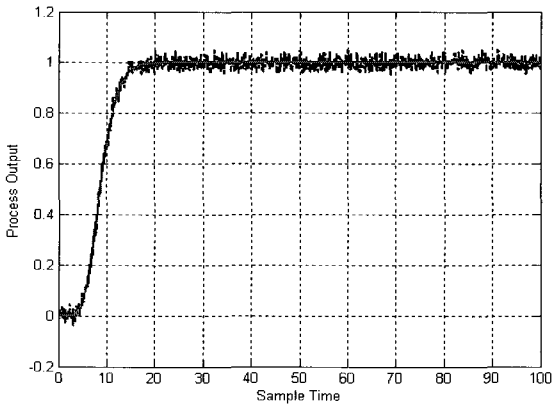
Population number	40
String length	3
Generation number	100
Crossover probability	0.9
Mutation probability	0.05
Crossover method	Modified convex
Selection method	Tournament
Mutation method	Modified real-number creep

Table 2. Comparison of results from identification methods for FOPDT.

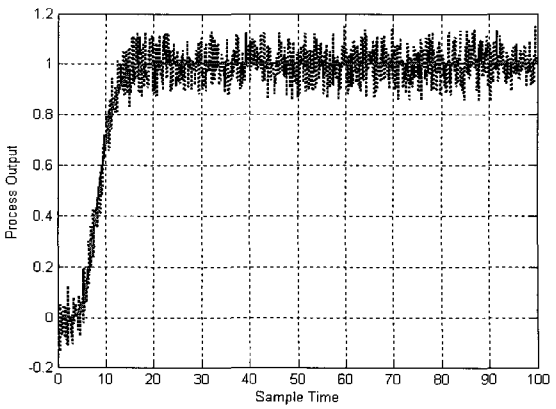
System $G(s)$	Identified system $\hat{G}(s)$	Error ε
$\frac{1-s}{(s+1)^5}$	Area method $\frac{e^{-4s}}{2.11s+1}$	$2.1767e^{-4}$
	Direct method $\frac{1.01e^{-3.73s}}{2.45s+1}$	$1.9714e^{-4}$
	Proposed method $\frac{e^{-3.6s}}{2.5s+1}$	$1.2435e^{-4}$
$\frac{1}{(s+1)^8}$	Area method $\frac{e^{-4.3s}}{4.3s+1}$	$5.1414e^{-4}$
	Direct method $\frac{1.06e^{-4.94s}}{3.81s+1}$	$3.0000e^{-3}$
	Proposed $\frac{0.99e^{-4.7s}}{3.5s+1}$	$4.2383e^{-4}$

To analysis identification characteristic of step response with noise for applying to the actual system, identification validity was examined using system step input with 30% and 10% noise.

Fig. 4(a) depicts the step response with 10% noise



(a) 10% noise.



(b) 30% noise.

Fig. 4. Step input reaction curve with noise.

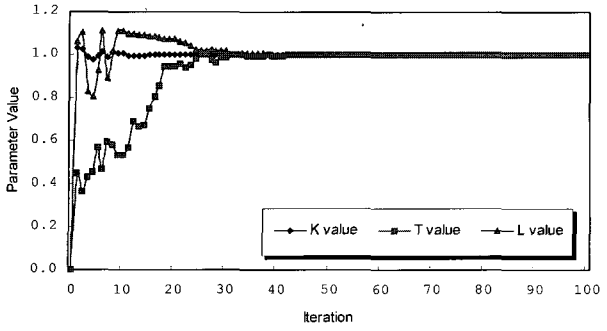


Fig. 5. Identification characteristic of FOPDT system.

about multi-lag system $\frac{1}{(s+1)^8}$, Fig. 4(b) shows the step response with 30% noise.

For the identification result by proposed GA, the estimated transfer function of Fig. 4(a) was determined to $\frac{1.0e^{-5.5s}}{3.55s+1}$ and the error value as $3.9157e^{-4}$ respectively. The estimated transfer function of Fig. 4(b) was acquired as $\frac{0.98e^{-1.5s}}{5.55s+1}$ and the error value as $8.4e^{-3}$ respectively. Fig. 5 shows the convergence characteristics for the parameters of the system model.

4.2. SOPDT system

The input parameters for the identification of the SOPDT system consists of four parameters as seen in(18). Four input parameters as seen in (19) are represented as down-to-four-points of real-number.

$$\frac{K}{T_1 s^2 + T_2 s + 1} e^{-Ls} \tag{18}$$

$$S_i = \{K_i, T_{1i}, T_{2i}, L_i\} \tag{19}$$

To identify a SOPDT system, the detail design parameters of the proposed real-coded GA were represented in Table 1.

Table 3, which is an error comparison table according to identification methods for the two systems, verified that the validity of system identification of proposed GA was better than those of directed identification and graphic identification.

The results of system identification for the two systems indicated that the model parameters converged stably at about 30 generation in Fig. 6 and Fig. 7 respectively. As well, the fitness of the two system models converged into a maximum value of 10 and showed satisfactory characteristics. Table 3, which is an error comparison table according to identification methods for the two systems, verified that the validity of system identification of the proposed GA was better than those of directed identification and graphic identification.

Table 3. Result comparison from identification methods for SOPDT.

System $G(s)$	Identified $\hat{G}(s)$	Error ε
$\frac{1}{(s+1)^5}$	Graphical $\frac{1}{4.71s^2+3.58s+1}e^{-1.35s}$	$1.5237e^{-5}$
	Direct $\frac{1}{4.4s^2+3.45s+1}e^{-1.45s}$	$2.0947e^{-5}$
	Proposed $\frac{1}{4.14s^2+3.54s+1}e^{-1.35s}$	$9.1784e^{-6}$
$\frac{1.08}{(s+1)^2(2s+1)^3}e^{-10s}$	Graphical $\frac{1.08}{11.76s^2+5.64s+1}e^{-12.17s}$	$4.9955e^{-5}$
	Direct $\frac{1.08}{11.9s^2+5.87s+1}e^{-12.06s}$	$1.9137e^{-5}$
	Proposed $\frac{1.08}{10.83s^2+5.82s+1}e^{-12.16s}$	$1.3197e^{-5}$

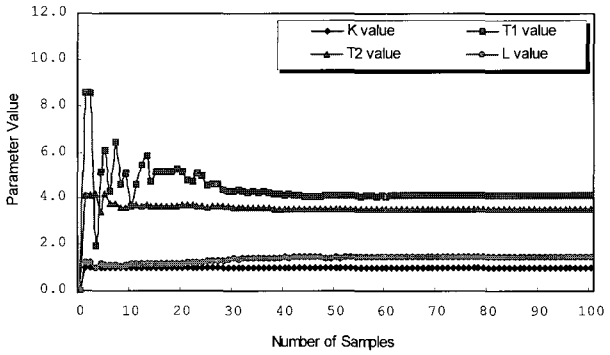


Fig. 6. Convergence curve of model parameters for $\frac{1}{(s+1)^5}$ system.

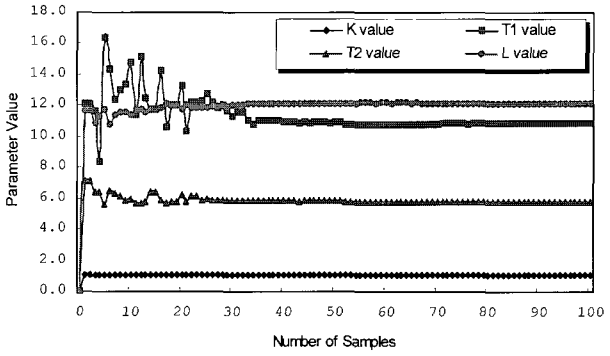


Fig. 7. Convergence curve of model parameters of $\frac{1.08}{(s+1)^2(2s+1)^3}e^{-10s}$ system.

5. CONCLUSIONS

In this paper, a new real-coded genetic algorithm was proposed to identify the parameters of FOPDT and SOPDT systems. To effectively apply the genetic algorithm to the identification, a modified crossover operator was proposed to select optimal parameters of the system. By varying the multiplier to extend the search space of the existing crossover operator, the new method makes a gene with a larger fitness have a greater chance of crossover. Comparing the proposed operator with the conventional crossover operator, excellent crossover characteristics were found from numerical experiments.

To verify the proposed algorithm numerically, this study compared the output responses obtained when step inputs were put into several systems. The proposed identification shows better results in error rates of step responses in comparison among area-based identification, directed identification, and proposed identification.

REFERENCES

[1] W. Zhang, Y. Sun, and X. Xu, "Two degree-of freedom smith predictor for processes with time delay," *Automatica*, vol. 34, no. 10, pp. 1279-1282, 1998.

[2] S. Majhi and D. P. Atherton, "Modified smith predictor and controller for processes with time delay," *IEE Proc. Control Theory Appl.*, vol. 146, no. 5, pp. 359-366, 1999.

[3] R. Gorecki and J. Jekielek, "Simplifying controller for process control of systems with large dead time," *ISA Transactions*, vol. 38, pp. 37-42, 1999.

[4] D. K. Lee, M. Y. Lee, S. W. Sung, and I. B. Lee, "Robust PID tuning for smith predictor in the presence of model uncertainty," *J. Process Control*, vol. 9, pp. 79-85, 1999.

[5] K. J. Astrom and T. Hagglund, *PID controllers: Theory, Design, and Tuning*, Instrument Society of America, 1995.

[6] A. Ingimundarson and T. Hagglund, "Robust tuning procedures of dead-time compensating controllers," *Control Engineering Practice*, vol. 9, pp. 1195-1208, 2001.

[7] Q. Bi, W. J. Cai, E. L. Lee, Q. G. Wang, C. C. Hang, and Y. Zhang, "Robust identification of first-order plus dead-time model from step response," *Control Engineering Practice*, vol. 7, pp. 71-77, 1999.

[8] Q. G. Wang and Y. Zhang, "Robust identification of continuous systems with dead-time from step responses," *Automatica*, vol. 37, pp. 377-390, 2001.

[9] Q. G. Wang, X. Guo, and Y. Zhang, "Direct identification of continuous time delay systems from step responses," *J. Process Control*, vol. 11, pp. 531-542, 2001.

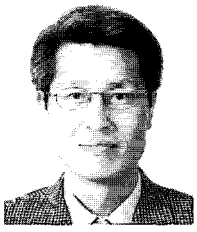
[10] K. H. Kim, *Multicriteria Structural Optimization by Genetic Algorithm*, Seoul National University, 1994.

[11] Y. Y. Kim, *Determination of Coagulant Dosage Rate for the Water Treatment Plant Using Intelligent Algorithms*, Chungnam National University, 2004.

[12] G. G. Jin, *Genetic Algorithms and Their Applications*, Kyowoo, 2000.

Gang-Wook Shin graduated with the B.S. degree in Electronic Engineering from Dongkuck University, Korea in 1987. He received the M.S. degree in Electronic Engineering in 1993 and the Ph.D. degree in Electrical Engineering from Hongik University, Korea in 2005. Currently, he is a Principal Researcher in Korea Institute of Water and Environment, Korea. His research interests include the Plant control system, the Intelligence control.

Young-Joo Song graduated with the B.S. degree in Electrical Engineering from Hongik University, Korea in 1993. He received the M.S degree in 1995 and the Ph.D. degree in Electrical Engineering from Hongik University, Korea in 2004. Currently, he is a Principal Research Engineer in the Institute of Electrical Facility, Hongik University. His research interests include the Electrical and Fire Facility system, the Intelligence control.



Tae-Bong Lee received the B.S., M.S., and Ph.D. degrees in Electronics Engineering from Hongik University, Seoul, Korea in 1986, 1994, and 1999, respectively. Currently, he is a Professor in the Dept. of Electronics and Informations, Kyungwon College. His research interests include variable structure system control, robust control and process control.

Hong-Kyoo Choi graduated with the B.S. degree in Electrical Engineering from Hongik University, Korea in 1975. He received the M.S. degree in Electrical Engineering from Yonsei University, Korea in 1977. He received the Ph.D. degree in Electrical Engineering from Hongik University, Korea in 1992. Currently, he is a Professor of Hongik University, Korea. His research interests include electrical and fire facility system, intelligence control.