

유전자알고리즘을 이용한 시지연 공정 식별

(The Identification of Time-Delay Process Using Genetic Algorithm)

최흥규·전광호·송영주·신강욱*

(Hong-Kyoo Choi · Kwang-Ho Jun · Young-joo Song · Gang-Wook Shin*)

홍익대학교 전기설비연구소·한국수자원공사

Abstract

In this paper, an identification method for a first order dead time process is proposed. This method used the genetic algorithm for parameter identification of process. The proposed method gives a better identification result than the existing methods under step testing. The effectiveness of the identification method has been demonstrated through a number of simulation examples.

1. 서론

공정산업에 있어서 대부분의 제어문제는 공정의 긴 지연시간에 따라 제어성능을 만족하기 어려운 부분이다. 이러한 지연 시간 문제를 해결하기 위하여 여러 가지 시간지연 보상기가 제안되었으며, 그 대표적인 보상기로는 1957년에 제안된 스미스예측기가 있다[1]. 스미스예측기는 공정과 모델의 일치에 따라 지연항이 포함되지 않은 모델식을 이용하여 미래의 제어입력을 예측하도록 함으로써 공정의 지연시간 문제를 해결할 수 있었다. 그러나, 스미스예측기는 공정모델의 정확성에 따라 많은 성능변화를 나타낸다. 이러한 모델불일치에 따른 문제에 대하여 다양한 연구가 진행되었으며, 대표적인 연구 결과는 모델 불일치에 따른 에러를 보상하기 위한 제어기를 추가하는 방법을 제시된 이중 스미스제어기가 제안되었다[2].

따라서, 긴 지연시간 문제를 해결하기 위해서는 기본적으로 공정의 정확한 모델을 구하기 위한 공정 파라메타 식별이 아주 중요하다. 기계시스템과 같은 동작을 요구하는 공정에 있어서는 일차 지연시간모델로는 정확한 식별이 힘들고 최소 이차 지연시간 모델 이상으로 모델링 되어야 한다. 그러나, 식별방법에 대한 지금까지의 여러 가지 연구들이 여러방면으로 진행되어 왔으나, 그 중에서도 산업공정에 가장 많이 활용되고 있는 모델로서 일차 지연시간모델에 대하여 고려되어야 할 필요가 있다.

일차 시간지연모델로 식별된 방법으로는 대표적으로 그래픽방법이라 할 수 있는 식별방법[3]이 있으며, 이는 계단입력에 대한 공정의 출력에 따른 그래프를 이용하

여 정상상태 값으로부터 공정의 이득을 구하고, 시정수와 지연시간항을 구하는 방법이다. 또한, 정상상태값의 28%와 63%에 도달하는 각각의 시간값을 이용하여 시정수와 지연시간항을 구할 수 있다. 그리고, 모델식별을 위한 선형회귀방정식의 변경으로 직접 모델 파라메타를 구하는 방법이 제시되었다[4].

본 논문에서는 유전자알고리즘을 적용하여 일차 시지연 공정에 대한 변수 추정알고리즘을 제시하였으며, 제시된 방법에 대한 서로 다른 네가지 유형의 공정을 선정하여 그 유용성을 시뮬레이션을 통하여 확인하고자 한다.

2. 배경

2.1. 대상공정

대부분의 화학공정과 선형 단조증가 공정을 나타내는 모델로 공정에 대한 근사식으로 일차 지연시간 모델(FOPDT Model : First-Order Plus Dead-Time Model)을 대부분 사용하고 있다[5]. 그 표현식은 다음 식 (1)과 같다.

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

여기서, K , T , L 은 각각 공정의 이득, 시정수, 그리고 지연시간을 말한다.

산업현장에서 이용되고 있는 모델식별 방법에 있어서 공정에 인가되는 입력 신호가 모델식별에 가장 중요한

영향을 미친다. 일반적으로 시험 입력신호로 사용되는 것은 펄스, 계단, 램프, 그리고 정현파함수 등이다. 그러나, 이들 가운데 계단입력시험이 가장 단순하며 장비가 거의 필요하지 않고 수동으로 이루어질 수 있는 장점을 가지고 있다. 또한, 계단입력 시험은 대부분의 PLC (Programmable Logic Controller) 혹은 DCS(Distribute Control System) 내부에 표준 모듈로 구성되어 있기 때문에 쉽게 구현 가능하다.

2.2. 면적식별방법

면적식별법에서 공정 이득 K 는 공정 입·출력의 정상상태로부터 구할 수 있다. 평균점유시간 T_{ar} 은 그림 1에서와 같이 면적 A_0 로부터 계산된다.

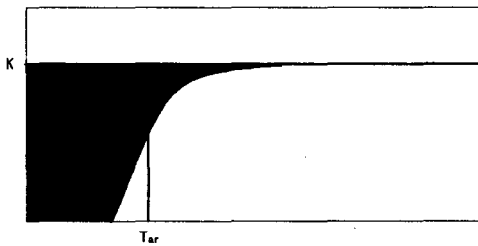


그림 1 계단응답에 대한 면적식별
Fig.1. Area identification for a step response

따라서, 평균점유시간은 다음 식 (2)와 같이 표현할 수 있다.

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

시간 T_{ar} 까지의 계단응답 면적을 A_1 으로 설정하여 시정수와 지연시간을 다음 식과 같이 각각 구할 수 있다.

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

$$L = T_{ar} - T = \frac{A_0}{K} - \frac{eA_1}{K} \quad (4)$$

이 방법은 식별 정확도가 주로 면적 A_0 에 의해 결정되며, 정확도를 높이기 위해서는 공정의 정상상태가 완전히 이루어질 때 까지의 충분한 시험 범위를 가지고 시험하여야 한다.

2.3. 적점식별방법

공정입력으로 $t=0$ 에서 h 크기의 계단입력 이전의 공정 초기상태는 영이라 가정한다. 새로운 정상상태에 공정이 도달할 때 까지 입력과 출력값을 저장한다. $t=L$ 이후의 식(1)의 공정에 대한 출력 과도응답은 식 (5)와 같이 나타낼 수 있다.

$$y(t) = hK(1 - e^{-(t-L)/T}) + w(t) \quad (5)$$

여기서, $w(t)$ 는 $y(t)$ 의 측정잡음으로의 백색잡음을 나타낸다.

식 (5)를 다시 표현하면, $t \geq L$ 에 대해

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

$t=0$ 부터 $t=\tau$ ($\tau \geq L$)까지 식 (5)에서의 출력을 적분하면 식 (7)을 얻는다.

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

그리고, $A(\tau)$ 와 $\delta(\tau)$ 를 다음과 같이 설정하고

$$A(\tau) = \int_0^{\tau} y(t) dt$$

$$\delta(\tau) = [Tw(t)] \Big|_L^{\tau} + \int_0^{\tau} w(t) dt$$

식 (7)을 다시 표현하면,

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

혹은

$$[h(\tau) - h - y(\tau)] \begin{bmatrix} K \\ LK \\ T \end{bmatrix} = A(\tau) - \delta(\tau) \quad (8)$$

$\tau \geq L$ 이상의 모든 샘플된 $y(\tau)$ 에 대한 식 (8)의 합

에 의해 선형방정식은 다음과 같이 얻어진다.

$\tau \geq L$ 에 대해

$$\Psi \theta = \Gamma + \Delta \quad (9)$$

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

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

$$\Gamma = \begin{bmatrix} A[mT_s] \\ A[(m+1)T_s] \\ \vdots \\ A[(m+n)T_s] \end{bmatrix},$$

$$\Delta = \begin{bmatrix} -\delta[mT_s] \\ -\delta[(m+1)T_s] \\ \vdots \\ -\delta[(m+n)T_s] \end{bmatrix}$$

T_s 는 샘플링 간격이며 $mT_s \geq L$ 이다. 식 (9)에서 θ 의 추정값 $\hat{\theta}$ 는 최소자승법을 이용하여 다음과 같이 얻을 수 있다.

$$\hat{\theta} = (\Psi^T \Psi)^{-1} \Psi^T \Gamma \quad (10)$$

따라서, 식 (1)의 K, T, L은 $\hat{\theta}$ 로부터 계산되어진다.

3. 유전 알고리즘

유전 알고리즘은 자연 선택의 원리와 자연계의 생물 유전학에 기본 이론을 둔 병렬적이고 전역적인 탐색 알고리즘으로서, 모든 생물은 주어진 다양한 환경 속에 적응함으로써 살아 남는다는 다윈의 적자생존(survival of the fittest)의 이론을 기본 개념으로 한다. 유전 알고리즘에서는 생물학적 유전인자인 염색체(chromosome)에 해당하는 문자열을 가지고 생물 유전을 모방한 연산자인 재생산(reproduction), 교배(crossover), 돌연변이(mutation)를 거쳐 다음세대의 새로운 자손(offspring)을 인공적으로 만들어 내고, 이를 반복적으로 적용하여 적합한 해를 탐색한다. 재생산 연산자는 각각의 개체가

가지는 적합도에 따라 그 개체를 복제하는 과정이며, 이 때 적합도 함수는 사용자가 최대화하기를 원하는 어떠한 형태의 함수도 가능하다. 이 과정은 적합도가 높은 개체일수록 다음 세대에 더 많은 자손을 가질 확률이 높음을 의미한다.

이러한 재생산 과정은 탐색 방향을 적합도가 높은 방향으로 향하게는 하지만 새로운 개체를 생성하지는 못한다. 자연계에서 자손은 부모 세대로부터 유전자를 이어받으면서도 부모와는 전혀 다른 형질을 가지듯이 유전 알고리즘에서 이 과정을 모방한 것이 교배 과정이다. 교배과정에서의 교배점과 교배율은 확률적인 과정에 의해 구함에도 불구하고 재생산과 함께 쓰여 개체간의 정보를 교환하고 스트링의 좋은 부분을 결합하는 효과적인 수단이 된다.

돌연변이 과정은 현재 집단에 존재하지 않는 새로운 정보를 제공해 준다. 이는 아주 낮은 돌연변이율에 의해 스티링의 한 비트를 0에서 1 혹은 1에서 0으로 우연히 바꿈으로써 구현된다. 이 과정에 의해 탐색의 방향이 지역적 극값으로 향할 경우 여기서 벗어나도록 한다.

한편, 적합도 함수는 각 개체의 적합도를 평가하는 기반이다. 이러한 적합도를 이용해서 다음 세대에서 선택되어질 확률이 결정된다. 즉 적합도가 높은 개체일수록 다음세대에서 선택되어질 확률이 높아지며 적합도가 낮을수록 다음 세대에서 선택되어질 확률이 낮아진다. 일반적으로 적합도 함수를 선정하는 방법은 각 개체를 이용해서 최적화 하고자 하는 대상에 적용하고 그 개체에 대한 결과와 원하는 기준입력에 대한 차를 이용해서 적합도 함수를 선정한다.

단계 1 : 부호화 및 초기화

공정식별을 위한 기본 파라미터는 유전자 알고리즘의 염색체로 설정하도록 부호화과정이 요구된다. 또한, 각 염색체에 대한 초기값을 각각 설정하여 최초의 염색체 군을 생성한다. 각 염색체는 2진수로 설정될 수 있으나, 공정이득과 시정수, 그리고 지연시간 각각에 대하여 2진수가 아닌 10진수 다섯자리가 되도록 선정하였다. 소숫점 세자리를 포함한 다섯자리 십진수로 설정하여 염색체 하나당 총 15자리가 되도록 구성한다.

단계 2 : 적합도

적합도 함수는 초기화된 염색체에 대하여 각각의 적합도를 비용함수 평가를 통하여 얻을 수 있다. 그러나, 본 연구에서는 먼저, 초기화된 염색체를 계단입력에 대한 실제공정의 응답특성과의 비교를 통하여 변화특성에 대하여 가중치를 갖도록 함으로써 전체적인 적합도를 구하였다.

단계 3 : 재생산

재생산은 개체집단 중에서 우수한 염색체를 선택하여 다음 세대를 생산하는 것이다. 재생산을 위해서 여러 가지 방법들이 존재하는데, 본 연구에서는 엘리트 방법을 사용하였다. 엘리트 방법은 적합도가 큰 순서로 염색체를 정렬한 후, 높은 적합도의 염색체가 새로운 세대를 생산하는데 사용된다. 특히 본 연구에서는 개체집단의 개체 수를 100개로 정하여 상위 50개를 엘리트로 선정하여 유전자 알고리즘을 수행하였다.

단계 4 : 교배

교배는 새로운 개체를 생성하는 것으로 선택된 모든 개체가 교배하는 것이 아니라 정해진 확률에 따라 교배하게 된다. 두 개의 부모개체가 새로운 두 개의 자식 염색체를 생성하는 것으로 교배 연산자는 일점교배와 다점교배, 그리고 유니폼 교배가 있다. 본 연구에서는 발생한 랜덤수가 교배확률보다 작을 경우 교배가 이루어지는 일점교배 방식을 적용하였다.

단계 5 : 돌연변이

탐색의 방향이 지역적 극값으로 향할 경우 여기서 벗어나도록 하기 위하여 돌연변이 과정은 현재 집단에 존재하지 않는 새로운 정보를 제공해 준다. 이는 탐색의 방향이 지역적 극값으로 향할 경우 여기서 벗어나도록 한다.

이상과 같은 유전 알고리즘에 의한 탐색에 있어서 일반적으로 교배율과 돌연변이율의 값이 크면 새로운 영역을 탐색하고자 하는 경향이 커져 진화의 초기에 적합도가 높은 탐색공간을 찾는다는 유리하나 탐색 결과를 이용하려는 경향이 작아져 최적해로 수렴해가는데 있어서 수렴속도를 저해한다. 또한 개체군의 크기가 작으면 시간을 절약할 수 있으나 개체간의 다양한 정보교환이 이루어지지 않아 최적해를 구하기 전에 수렴할 위험성을 내포한다. 반면에 개체군의 크기가 크면 최적해에 도달할 확률은 높으나 많은 기억 용량과 계산시간을 필요로 하는 단점이 있다. 따라서 주어진 대상물에 대해 적절한 유전 알고리즘의 파라미터 선정이 필요하다.

4. 사례 연구

유전 알고리즘을 이용한 공정식별법의 유용성을 확인하기 위하여 네가지 유형의 공정에 대해 면적식별법과 직접식별법에 대해 각각 비교 시험을 실시하였다. 본 사례연구에서 사용한 유전자 알고리즘의 표 1과 같이 개체수 및 세대수를 100으로 하고, 교배확률을 0.9로 선정하여 공정의 파라메타값을 최적화하였다.

표 1. 유전자 알고리즘 파라메타

Table 1. Parameters of Genetic Algorithm

개체의 수	100
스트링의 길이	15
세대의 수	100
오차수준	0.01
교배확률	0.9
돌연변이 확률	0.05

본 사례연구에서 고려된 시스템은 일차 지연시간 공정, 2차 지연시간 공정, 비최소 위상 공정, 그리고 다중 지연 공정으로 선정하였으며, 공정에 대한 수식은 다음과 같다.

$$G(s) = \frac{e^{-s}}{s+1}$$

$$G(s) = \frac{e^{-4s}}{(10s+1)(2s+1)}$$

$$G(s) = \frac{1-s}{(s+1)^5}$$

$$G(s) = \frac{1}{(s+1)^8}$$

각 공정의 계단입력에 대한 시간영역 공정식별의 오차를 측정하기 위하여 다음과 같이 실제공정의 출력과 추정값과의 평균 오차를 구하는 식을 이용하였다.

$$\epsilon = \frac{1}{n+1} \sum_{k=m}^{m+n} [y(kT_s) - \hat{y}(kT_s)]^2 \quad (11)$$

여기서, $y(kT_s)$ 는 계단입력 변화에 대한 실제 공정의 출력을 나타내고, $\hat{y}(kT_s)$ 는 추정된 공정의 응답을 나타낸다.

위 시뮬레이션을 통하여 각 유형의 공정에 대한 오차값을 살펴보면, 면적식별법과 직접식별법을 통한 응답특성은 비슷한 결과를 나타내지만, 유전자 알고리즘을 이용한 공정식별을 통한 오차값은 다른 식별법보다 양호한 결과를 얻었다.

네가지 공정중 일차 지연시간 공정에 대한 각 공정식별법의 오차값은 비슷한 결과를 나타내고, 2차 지연시간 공정을 비롯한, 비최소 위상 공정과 다중 지연 공정에 대해서는 비교적 오차값이 커지는 것을 알 수 있다.

5. 결론

참고 문헌

본 연구에서는 유전 알고리즘을 이용하여 일차 시지연 공정에 대한 새로운 식별방법을 제시하였다. 기존의 직접식별법은 면적식별법보다 빠른 식별이 가능하지만, 시뮬레이션을 통하여 실제 공정과의 오차는 면적식별법보다 약간 큰 편인 것을 알 수 있었다. 그러나, 유전자 알고리즘을 이용한 식별법을 통해서 알 수 있는 실제 공정과의 오차는 세가지 식별방법 중 가장 양호한 것으로 판단할 수 있다.

향후 유전자 알고리즘을 이용한 공정식별의 실제 산업 현장에 적용하기 위해서는 적합도 및 교배에 따른 계산량의 요구를 해결하고, 실시간 공정식별이 가능하도록 연구되어야 한다.

- (1) Smith, O. J. M. "Closed control of loops with dead time", Chemical Engineering Progress, 53, May, pp.217~219, 1957.
- (2) Tian, Y. C., Cao, F. "Double-controller scheme for control of processes with dominant delay", IEE Proc. Control Theory Application, Vol. 145, No. 5, September, pp.479~484, 1998.
- (3) Astrom, K.J., Hagglund, T., "PID controllers: theory, design, and tuning (2nd ed.)", Instrument Society of America, 1995.
- (4) Qiang Bi, Wen-jian Cai, "Robust Identification of First-Order Plus Dead-Time Model from Step Response", Control Engineering Practice, No. 7, pp.71~77, 1999.
- (5) Seborg, D.E., Edgar, T.F., Mellichamp, D.A., "Process dynamics and control", John Wiley & Sons, Inc., 1989.

표 2. 식별방법에 따른 결과 비교

Table 2. Comparison of results from identification methods

대상 공정 $G(s)$	식별 공정 전달함수 $\hat{G}(s)$	오차값
$\frac{e^{-s}}{s+1}$	면적식별 $\frac{e^{-1.01s}}{0.99s+1}$	$2.0828 e^{-7}$
	직접식별 $\frac{1.00 e^{-1.00s}}{0.997s+1}$	$2.2534 e^{-8}$
	유전자식별 $\frac{1.00 e^{-1.00s}}{0.999s+1}$	$2.5012 e^{-9}$
$\frac{e^{-4s}}{(10s+1)(2s+1)}$	면적식별 $\frac{1.00 e^{-5.83s}}{10.19s+1}$	$4.2598 e^{-5}$
	직접식별 $\frac{1.03 e^{-5.47s}}{11.41s+1}$	$4.5042 e^{-4}$
	유전자식별 $\frac{1.00 e^{-5.67s}}{10.52s+1}$	$3.4717 e^{-5}$
$\frac{1-s}{(s+1)^5}$	면적식별 $\frac{1.00 e^{-4.00s}}{2.11s+1}$	$2.1767 e^{-4}$
	직접식별 $\frac{1.01 e^{-3.73s}}{2.45s+1}$	$1.9714 e^{-4}$
	유전자식별 $\frac{1.00 e^{-3.6s}}{2.5s+1}$	$1.2435 e^{-4}$
$\frac{1}{(s+1)^8}$	면적식별 $\frac{1.00 e^{-4.3s}}{4.3s+1}$	$5.1414 e^{-4}$
	직접식별 $\frac{1.06 e^{-4.94s}}{3.81s+1}$	$3.0000 e^{-3}$
	유전자식별 $\frac{0.99 e^{-4.7s}}{3.5s+1}$	$4.2383 e^{-4}$