

유전자 알고리즘 기반 다항식 뉴럴네트워크를 이용한 비선형 질소제거 SBR 공정의 모델링

Modeling of Nonlinear SBR Process for Nitrogen Removal via GA-based Polynomial Neural Network

김 동 원, 박 장 현, 이 호 식, 박 영 환, 박 귀 태*

(Dong-Won Kim, Jang-Hyun Park, Ho-Sik Lee, Young-Whan Park, and Gwi-Tae Park)

Abstract : This paper is concerned with the modeling and identification of sequencing batch reactor (SBR) via genetic algorithm based polynomial neural network (GA-based PNN). The model describes a biological SBR used in the wastewater treatment process for nitrogen removal. A conventional polynomial neural network (PNN) is applied to construct a predictive model of SBR process for nitrogen removal before. But the performances of PNN depend strongly on the number of input variables available to the model, the number of input variables and type (order) of the polynomials to each node. They must be fixed by the designer in advance before the architecture is constructed. So the trial and error method must go with heavy computation burden and low efficiency. To alleviate these problems, we propose GA-based PNN. The order of the polynomial, the number of input variables, and the optimum input variables are encoded as a chromosome and fitness of each chromosome is computed. Simulation results have shown that the complex SBR process can be modeled reasonably well by the present scheme with a much simpler structure compared with the conventional PNN model.

Keywords : sequencing batch reactor (SBR) process, nitrogen removal, GA-based polynomial neural network

I. 서론

활성오니는 공해물질을 정화하기 위해 폐수가 미생물의 부유물과 혼합된 생물학적 공정이다. Sequencing batch reactor (SBR) 공정은 시설 및 운전이 간단하고, 질소 및 인의 고도 처리가 가능하며 읍면 단위의 소규모 시설로써 적당하다 [1]. 하나의 반응기에서 운전시간에 따라 오수를 공급하는 단계, 산소를 공급하는 단계, 침전기, 배수기, 휴지기와 같은 서로 다른 상태를 가진 여러 반응이 SBR에서 나타나며, 장치는 간단하나 반응기내에서 일어나는 각각의 반응은 연속식 공정보다 매우 복잡하고 다양한 메커니즘을 가지고 있다 [2]. 그러므로 다항식 뉴럴네트워크 (polynomial neural network: PNN)와 같은 지능기술을 이용하여 복잡한 메커니즘을 모델링한 연구결과가 있다[6]. [6]에서는 질소제거를 위한 DO, ORP, pH의 실시간 변화를 관찰하여 질소제거특성을 검토하였으며 이를 통해 얻은 실제의 입출력데이터를 가지고 SBR 반응조의 질소제거 공정을 모델링 하였다. 그러나 [6]에서 사용된 PNN의 성능은 노드를 구성하는 다항식의 입력변수와 차수에 상당히 의존적이다. 이러한 요소들은 설계자의 배경 지식과 시행착오에 의해 결정되어 모델설계가 이루어진다. 그러므로 과중한 연산 시간과 저효율의 부담을 안고 설계자의 시행착오에 의해 모델링 된다. 따라서 설계자의 경험적인 방법으로 생성된 최종적인 PNN 모델은 비선형 시스템에 대

한 최상의 모델을 보장할 수 없다.

본 논문에서는 이러한 문제점을 해결하기 위해 유전자 알고리즘을 적용하여 다항식 뉴럴네트워크를 최적화 하였으며, 이를 이용하여 SBR 반응조의 질소제거 공정을 모델링 하였다. 유전자 알고리즘 기반 다항식 뉴럴네트워크 (GA-based PNN)는 이미 비선형성과 불확실성이 강해 제어가 어려운 box-jenkin의 가스로 공정과 중화점 근처에서 매우 민감하게 반응하여 다루기 힘든 pH 중화공정에 적용되어 우수한 성능과 구조의 간결함이 기존 모델링 방법과 비교하여 입증되었다 [12]. GA-based PNN의 성능비교를 위해 기존 PNN에서 사용된 질소제거 SBR공정을 모델링하여 비교 설명하고, 제안된 모델의 성능이 보다 우수함을 보인다. 따라서 GA-based PNN을 이용하여 효율적인 자동제어 공정을 구현하여 질소제거의 효율을 증가시키며, 유지관리가 용이하고 비용을 절감할 수 있는 질소제거 SBR공정을 기대할 수 있다.

II. 질소제거를 위한 SBR공정

폐수에서 질소나 인을 제거하기 위해 생물학적인 SBR 공정이 주로 사용된다. 활성오니에 있는 박테리아는 유해한 성분을 물이나 질소가스등의 무해한 성분으로 바꾸며 반응조에서 폐수는 활성오니와 섞여진다. 그러므로 활성오니는 폐수와 섞인 상태로 유지되기 위해 혼합작업이 필요하다. SBR 공정에서 폐수처리 작업은 반응조에 폐수를 공급하면서 혼합하는 단계와 산소를 공급하는 단계, 혼합과 산소공급을 중단하는 단계, 그리고 마지막으로 정화된 물을 반응조에서 배출하는 단계로 구분할 수 있다.

전체 시스템의 구성은 [6]와 같으며, [6]의 내용을 간략히 정리하면 다음과 같다. 반응조는 아크릴로 제작하였으며 반응조의 용적과 실용적은 각각 15L 와 10L 이며, 처리용량은

* 책임저자(Corresponding Author)

논문접수 : 2003. 7. 30., 채택확정 : 2003. 10. 30.

김동원, 박귀태: 고려대학교 전기공학부

(dwkim@elec.korea.ac.kr; gtpark@korea.ac.kr)

이호식 : 충주대학교 환경공학과(hlee@chungju.ac.kr)

박영환 : 충주대학교 정보제어공학과(pyh@chungju.ac.kr)

박장현 : 목포대학교 제어시스템 공학과(jhpark72@mokpo.ac.kr)

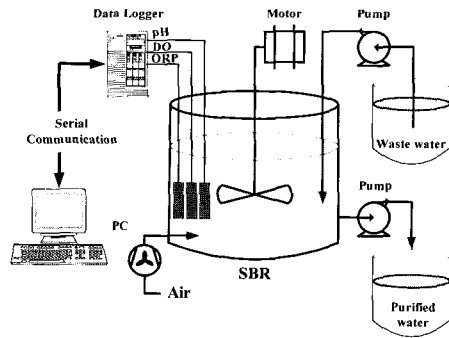


그림 1. 질소제거를 위한 반응조의 실험장치.
Fig. 1. Experimental setup of SBR for nitrogen removal.

표 1. SBR의 운전조건.

Table 1. Working condition of SBR.

Parameter	Average
volume	10 l
SRT	30 d
MLSS	2500 mg/l
MLVSS/MLSS	0.701
Temperature	23 ± 2 °C

3L이다. 미생물의 혼합은 반응조내에 설치된 60rpm의 믹서를 이용하였고 동시에 반응조 바닥에 설치된 산소주입 포기 장치에 의해 혼합이 이루어 진다. 유출수는 솔레노이드 밸브를 달아 자동으로 조절되고 유입수량은 masterflux tubing pump (cole parmer model 7553-80)를 이용하여 조정되었으며 이들은 타이머에 의해 제어되었다. ORP, pH, DO 센서를 반응조의 벽면에 일정하게 고정시켜 연속모니터링 하였고 2주에 한번씩 보정하였다. 전체동작은 6시간을 한 주기 (1cycle)로 해서 운전이 이루어진다. 무산소, 호기조건은 운전시간이 고정되어 있지 않고 반응조의 ORP나 DO 센서의 실시간 모니터링 자료에 의해 변화 운전되었다. 아울러 침전과 배수, 휴지기 시간은 각각 0.33hr, 0.17hr로 고정하였다. 반응조의 실험 장치와 운전조건은 그림 1과 표 1과 같다.

그림 1에서 반응조인 SBR은 원통형으로 센서부와 구동부의 기기들이 장치되어 있다. 센서부는 항상 폐수에 잠겨있으며, ORP, DO, pH값을 주기적으로 읽어 들인다. 데이터 수집기는 센서부에서 측정된 값을 직렬통신으로 PC로 전달한다. SBR 상단에는 폐수를 공급하는 펌프가 있으며, 하단에는 정화수를 배출하는 펌프와 산소를 공급하는 펌프가 각각 연결되어 있다. 믹서는 모터와 연결되어 폐수를 저어서 화학반응을 촉진시키는 역할을 한다. 각각의 펌프와 모터는 전기계전기와 연결되어 있으며 PC에서 내보내는 제어신호에 의해 on/off 된다.

실험에 사용한 시료는 충주 처리장에 유입되는 생활하수로서 시료의 종류와 상태는 표 2와 같다. 수질분석 항목으로는 BOD, COD_{cr}, TSS/VSS, TKN, NH₄-N, NO₂-N, NO₃-N, TP 등으로 모든 분석방법은 수질오염공정시험법과 standard methods 및 일부 기기분석법에 의해 수행되었다. ORP, DO, pH의 실시간 측정은 ALMEMO사의 센서 (pH: FY8PHK, ORP: FYA8RXEN, DO:FYA64002)를 사용하여 data logger (ALMEMO

표 2. 시료의 종류와 상태.

Table 2. Constituent and state of material.

Constituent	Range	Average	Standard deviation
pH	6.7~7.5	7.1	0.12
BOD ₅	50~85	63	14.4
COD _{cr}	112~152	131	18.8
TSS	31~52	45	6.9
VSS	23~46	36	7.4
NH ₄ -N	18~28	22	3.5
TP	2.18~2.22	2.3	0.05

20090-8)에 연결하여 측정하였다. 실험주기는 유입수는 일주일에 1회, 유출수는 2일에 1회, 반응조 MLSS분석은 3일에 1회 분석하였고, SRT는 30일로 유지하였다.

III. 다항식 뉴럴네트워크

다항식 뉴럴네트워크 (polynomial neural network: PNN)[9]는 GMDH [10]알고리즘에 기반한 것으로 유용한 근사방법의 하나이다. PNN의 구조는 다항식 노드를 가진 전방향 뉴럴네트워크와 유사하며, 각 노드의 출력은 1차 다항식 (linear), 2차 다항식 (quadratic), 변형된 2차 다항식 (modified quadratic)을 이용하여 얻는다. 이러한 다항식을 일컬어 부분표현식 (partial description: PD)이라 한다. PNN 알고리즘의 설계순서를 간단히 기술하면 다음과 같다.

단계 1- 입력변수를 설정하는 단계로, 출력변수 y에 관련된 입력의 N개의 입력변수를 설정한다.

단계 2- 시스템을 통해 얻은 데이터를 학습용 데이터와 평가용 데이터로 분할하는 단계로, 학습용 데이터는 모델동정을 위해 사용하고, 평가용 데이터는 동정된 모델을 평가하기 위해 사용된다.

단계 3- PNN 구조를 선택하는 단계로, PNN 구조는 첫 번째 층과 그 이후 층의 부분표현식을 구성하는 방법에 따라 그 구조가 정해진다. 첫 번째 층과 그 이후 층의 각 노드의 입력변수 수가 동일한 경우 (기본 PNN)와 그렇지 않은 경우 (변형된 PNN)의 두 구조로 분류되며, 또 각 구조는 노드의 부분표현식의 차수에 따라 각각 Case 1과 Case 2로 다시 세분화된다. Case 1은 모든 층에서 노드의 PD가 동일 차수를 가지는 구조이며, Case 2는 첫 번째 층과 그 이후 층의 노드의 PD가 서로 다른 차수를 가지는 구조이다.

단계 4- 노드의 차수 및 입력변수 수를 선택하는 단계로, N개의 입력변수 중에서 각 노드에 사용할 입력변수 수와 차수를 선택한다. N에서 r개의 변수를 각 노드의 입력으로 선택하면 필요한 총 노드의 개수는 $\frac{N!}{r!(N-r)!}$ 가 되며, N은 전

체 입력변수의 수이고, r은 설계자가 택한 입력변수의 수이다. 예를 들어, 노드의 입력 수로 2 입력을 선택했을 때 PD로 나타날 수 있는 형태는 다음과 같다.

$$\text{Type 1} = c_0 + c_1x_1 + c_2x_2$$

$$\text{Type 2} = c_0 + c_1x_1 + c_2x_2 + c_3x_1^2 + c_4x_2^2 + c_5x_1x_2$$

$$\text{Type 3} = c_0 + c_1x_1 + c_2x_2 + c_3x_1x_2$$

표 3. 다항식 뉴럴네트워크의 진화론적 설계를 위한 유전 알고리즘의 매개변수.

Table 3. Parameters for evolutionary design of PNN.

parameters	1st layer	2nd layer	3rd layer
Max. generation	40	60	80
Population size	20(15)	60(50)	80
Chromosome length	9	20	55
Crossover rate	0.85		
Mutation rate	0.05		
Type	1 ~ 3		

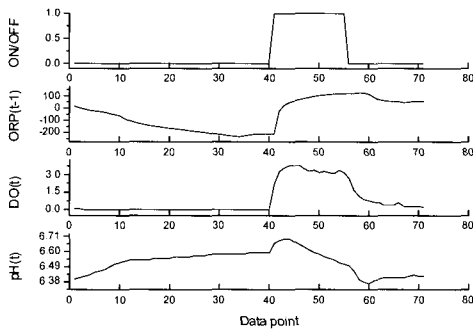


그림 4. 입력변수로 사용된 데이터 벡터.

Fig. 4. Used data vector for input variables.

관찰하여 실제의 입출력데이터를 얻었으며, 이것을 이용하여 모델링을 수행한다.

시스템의 입력변수로는 반응조에 딸린 각 펌프를 on/off 시키는 스위치와 이에 따라 변화하는 ORP, DO, pH값이 사용되었으며, 출력변수로는 비선형성과 불확실성이 강한 ORP값이 사용되었다. 입력변수로 사용된 ORP는 출력변수로 사용된 것보다 한 샘플링이 지연된다. 총 데이터 개수는 71개가 사용되었으며, 모두 학습용 데이터로 이용하였다. 입력변수로 사용된 입력 데이터벡터를 그림 4에서 나타내었다.

모델의 정밀도와 성능을 평가하기 위한 성능지수 (PI)는 평균자승오차 (mean squared error: MSE)를 사용하였으며, 다음과 같이 계산된다.

$$PI = \frac{1}{71} \sum_{i=1}^{71} (y_i - \hat{y}_i)^2 \quad (2)$$

여기서, y_i 는 실제측정치이며, \hat{y}_i 는 모델의 출력값이다.

유전자 알고리즘에 기반한 다항식 뉴럴네트워크의 설계는 다음과 같이 이루어진다. 1, 2, 3층에서 각각 20, 60, 80개의 염색체를 생성한다. 생성된 염색체는 각 층에서 40, 60, 80세대 동안 진화과정을 수행한다. 1층에서 20개의 염색체는 유전자알고리즘에 의해서 정의된 입력변수와 차수를 가지고 각각의 PD를 구성한다. 즉 전체 4 시스템입력 중 노드의 입력변수와 차수의 형태를 염색체로 자동 선택하며 이에 따라 여러 가지 형태의 PD를 구성한다. 각 PD는 정해진 세대수까지 진화하면서 최적의 입력변수와 차수를 탐색한다.

염색체의 진화를 통해 탐색된 최적의 정보로 구성된 PD는 학습데이터를 이용하여 성능지수 (동정오차)가 구해지며 (2)

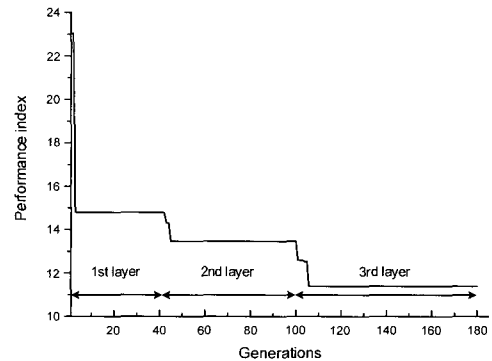


그림 5. 유전자 알고리즘에 의한 성능지수의 수렴과정.

Fig. 5. Convergence process of performance index by GA.

표 4. 기존 다항식 뉴럴네트워크 모델과 동정 오차비교.

Table 4. Comparison of identification error with conventional polynomial neural network.

Structure	layer				PI
	Inputs		Order		
	1	2-3	1	2-3	
Basic -	2	2	Type 2	Type 2	27.37
Case 1	3	3	Type 3	Type 3	22.80
Basic -	2	2	Type 3	Type 2	30.62
Case 2	3	3	Type 3	Type 2	13.37
[6] Modified	2	3	Type 2	Type 2	15.17
-Case 1	3	2	Type 2	Type 2	15.72
Modified	2	3	Type 2	Type 3	26.64
-Case 2	3	2	Type 3	Type 2	21.01
Our model	자동탐색				11.39

로 정의된다. 각 PD는 성능지수뿐만 아니라 목적함수에 의해 평가되며, 각각의 적합도 값에 따라 순위가 정해진다. 적합도는 개체군내의 각 개체에 대하여 풀고자 하는 문제에 대하여 어느 정도 적합한지를 평가하는 함수로 (3)과 같이 정의한다.

$$fitness = \frac{1}{1 + PI} \quad (3)$$

여기서 우수한 적합도 값을 가진 PD는 선택되어 다음 층까지 유지되며 그렇지 못한 것은 제거된다. 각 층에서 선택되는 PD의 수는 1, 2, 3층에 따라 각각 15, 50, 80개가 된다. 선택된 PD의 출력값은 다음 층의 입력값이 되어 새로운 입출력 데이터를 구성한다. 이제 새로운 층의 PD정보가 다시 탐색되며 위의 과정이 반복된다. 이 과정은 종료판정 기준인 3층까지 반복 진행되며 3층에서의 가장 좋은 성능지수가 최종 출력값이 된다. 최적의 성능을 얻기 위한 유전 알고리즘의 매개변수와 설정치는 표 3과 같다.

각 층에서 GA의 연속된 세대수에 대한 성능지수의 수렴과정이 그림 5에서 나타내었다. 성능지수는 PI의 형태로 구성되었으며 이에 대한 적합도 함수값은 그림 6과 같다. 그림 7에서는 실제의 출력값과 유전자 알고리즘에 기반한 다항식 뉴럴네트워크 모델의 출력값을 비교하여 나타내었다. 이때 제한된 모델의 출력값은 3층에서 11.398을 얻었다. 그림 7에서

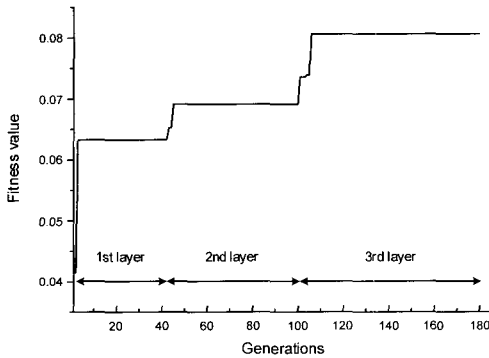


그림 6. 유전자 알고리즘에 의한 적합도의 수렴과정.
Fig. 6. Convergence process of fitness value by GA.

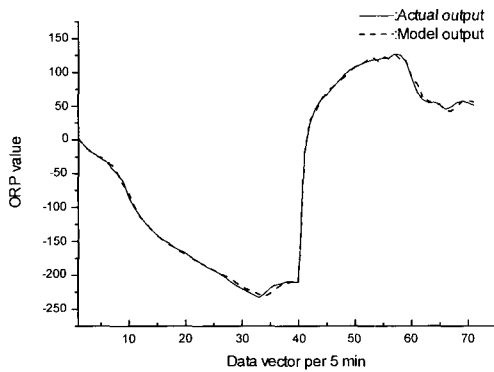


그림 7. 실제데이터와 유전자 알고리즘 기반 다항식 뉴럴네트워드의 결과값 비교.
Fig. 7. Comparison results of actual data and GA-based PNN.

보인 data vector는 위에서 언급한 시스템의 입력변수로, 센서 부에서 측정된 값을 데이터 수집기를 통해 5분당 실측된 값으로 반응조에 딸린 각 펌프를 on/off 시키는 스위치와 이에 따라 변화하는 ORP, DO, pH값이다.

기존의 PNN [6]과 비교하기 위해 표 4에서 PNN의 성능지수를 나타내었다. 표에서 보듯이 PNN은 입력변수와 차수에 따라 다양한 결과값이 생성되므로 적절한 입력변수와 차수를 결정하기 위해 시행착오를 통한 다양한 경우의 결과값을 확인해야 한다. 그러나 제안된 유전자 알고리즘 기반 PNN은 유전자 알고리즘에 의하여 입력변수와 차수의 형태를 자동적으로 검색하여 최적의 해를 얻었으며 3층에서의 결과값을 비교할 때 제안된 모델의 결과값이 더욱 우수함을 알 수 있다.

VI. 결론 및 추후과제

본 논문에서는 질소제거를 위한 SBR 공정에서 유입수질에 따라 매우 민감하게 반응하는 ORP를 모델링 하기 위해 유전자 알고리즘 기반 다항식 뉴럴네트워크 알고리즘 (GA-based PNN)이 사용되었다. 기존 PNN 알고리즘은 입력변수와 차수에 따라 다양한 형태의 구조와 이에 따른 성능의 차이가 있으므로 최적의 모델을 찾기 어려웠다. 그러나 제안된 GA-based

PNN 알고리즘은 각 노드에 인가되는 입력변수와 차수의 형태를 자동적으로 검색하여 최적의 해를 얻었으며 시뮬레이션 결과에서 보듯이 본 논문에서 제안된 모델이 기존의 PNN보다 우수한 성능을 보임을 알았고, 이를 통하여 비선형 시스템을 합리적이고 효율적으로 모델링 할 수 있었다. 앞으로의 과제는 시스템에 인가되는 잡음을 효율적으로 처리할 수 있는 새로운 지능기법을 이용하여 모델링을 수행하고 이의 결과를 비교 분석하는 것이다.

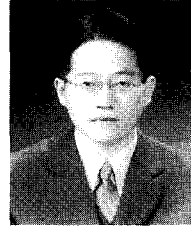
참고문헌

- [1] 이재근, 김영균, 이광호, "SBR공정에서 혐기시간 및 HRT 변화에 따른 탈질율이 질소제거에 미치는 영향," 대한환경공학회 2003 춘계학술연구발표회 논문집, pp. 151-155, 2003.
- [2] 심문용, 부경민, 박지영, 전병희, 김창원, 우혜진, "On-line sensor (ORP, DO, pH)를 이용한 SBR(Sequencing Batch Reactor)공정진단에 대한 기초평가," 대한환경공학회 2003 춘계학술연구발표회 논문집, pp. 160-170, 2003.
- [3] J. Moreno, "Optimal time control of sequencing batch reactors for industrial wastewaters treatment," *Proceedings of the 36th Conference on Decision & Control*, pp. 826-827, 1997.
- [4] V. V. Lira, P. R. Barros and J. S. Neto, "Automation of an anaerobic-aerobic wastewater treatment process," *IEEE Instrumentation and Measurement Technology Conference*, pp. 1177-1182, 2001.
- [5] 김창원, "활성슬러지 공정 연속호흡율 측정기 및 MLSS meter와 이를 이용한 활성슬러지 공정제어," Eco-2 상하수도 정보화 세미나, 2001.
- [6] 김동원, 박영환, 박귀태, "다항식 뉴럴네트워크를 이용한 질소제거 SBR 공정의 모델링," 대한전기학회 논문지 (심사중)
- [7] H. Kim, and O. J. Hao, "pH and oxidation-reduction potential control strategy for optimization of nitrogen removal in an alternating aerobic anoxic system," *Water Environ Res*, vol. 73, pp. 95-102, 2001.
- [8] 이호식, "TKN/COD비가 높은 소규모지역 오수의 생물학적 영양소 제거,"충주대학교 논문집, 2001.
- [9] 오성권, 김동원, 박병준, "다항식 뉴럴네트워크 구조의 최적 설계에 관한 연구," *Trans. KIEE*, vol. 49D, no. 3, pp. 145-156, 2000.
- [10] A. G. Ivakhnenko, "Polynomial theory of complex systems", *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-1, no. 1, pp. 364-378, 1971.
- [11] J.H. Holland, "Adaptation in natural and artificial systems", The University of Michigan Press, Ann Arbor, M.I., 1975.
- [12] 김동원, 박귀태, "다항식 뉴럴네트워크의 최적화: 진화론적 방법," *Trans. KIEE*, vol. 52D, no. 7, pp. 424-433, 2003.



김 동 원

1974년 3월 15일생. 2002년~현재 고려대학교 전기공학과 박사수료. 관심분야는 소프트웨어컴퓨팅기반 뉴로-퍼지 시스템, 시스템 모델링 및 지능제어, GMDH-type 알고리즘.



박 장 현

1972년 1월 24일생. 2003년~현재 목포대학교 제어시스템 공학과 전임강사. 관심분야는 비선형제어, 지능제어 이론, 퍼지추론 시스템, 신경망.



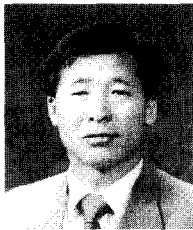
이 호 식

1963년 10월 10일생. 2002년~현재 충주대학교 환경공학과 부교수. 관심분야는 소규모 마을 하수도관리, 자동제어를 이용한 질소, 인, 폐수처리.



박 영 환

1961년 1월 28일생. 1998년~현재 충주대학교 정보제어공학과 조교수. 관심분야는 적응제어, 비선형제어, 지능제어 이론.



박 귀 태

1947년 10월 25일생. 1981년~현재 고려대학교 전기공학과 교수. 2000~현재 (사) IBS KOREA 회장. 관심분야는 퍼지 및 신경망, 이동로봇, 컴퓨터 비전, 지능제어, 인텔리전트 빌딩시스템 및 정보통신.