

유전알고리즘과 퍼지추론시스템의 합성을 이용한 정수처리공정의 약품주입률 결정

Determination of Dosing Rate for Water Treatment
Using Fusion of Genetic Algorithms and Fuzzy Inference System

김용열, 강이석**

*한국수자원공사 정보관리실 (Tel: (042) 629-2590; Fax: (042)623-0963)

**충남대학교 기계설계공학과 (Tel: (042) 821-6647; Fax: (042)822-7366)

Abstract It is difficult to determine the feeding rate of coagulant in water treatment process, due to nonlinearity, multivariables and slow response characteristics etc. To deal with this difficulty, the fusion of genetic algorithms and fuzzy inference system was used in determining of feeding rate of coagulant. The genetic algorithms are excellently robust in complex optimization problems, since it uses randomized operators and searches for the best chromosome without auxiliary information from a population consists of codings of parameter set. To apply this algorithms, we made the look up table and membership function from the actual operation data of water treatment process. We determined optimum dosages of coagulant (PAC, LAS etc.) by the fuzzy operation, and compared it with the feeding rate of the actual operation data.

Keywords Fuzzy inference system, Genetic algorithms, Water treatment, Optimum dosages, Feeding rate

1. 서 론

정수처리공정에서 원수의 수질에 따라서 주입되는 약품주입률은 원수탁도, pH, 알칼리도, 수온등 여러 가지 복합인자가 합수가 되는 반면에 약품주입지점인 혼화지 유입구와 탁질제거효과를 판단할 수 있는 침전지 유출구간에 체류시간이 존재하여 실시간 피드백 제어를 할 수 없다. 이러한 문제점은 일종의 모의실험장치인 자-테스트를 실시하여 약품주입률을 결정하고 있는데 실제 운영단계에서는 실험에 소요되는 시간, 조건의 변화, 처리시간의 차이, 수동시험의 불편등으로 어려움이 존재하고 있다.

약품주입률결정에 어려움을 개선하고자하는 연구는 크게 두 가지로 분류되고 있다. 첫째는 약품주입직후의 상태를 SC D(stream current detector)나 제타 전위(zeta potential)측정과 같은 센서를 이용하여 체류시간 후의 결과를 예측하여 실시간 피드백제어를 실현하고자하는 방안이다[3]. 둘째로는 통계적인 방법, 실험에 의한 계산식 산출 및 학습에 의한 최적주입률 결정등과 같은 피드 퍼워드에 의한 방법이 연구되고 있다. 운전실적자료를 토대로 한 수학적 모델식을 구하는 통계적인 방법은 수학적 모델식이 기후, 시각, 계절에 따르는 원유량, 수질의 큰 변화에 정확하고 효과적으로 약품주입공정을 기술하지 못한다. Jar-test를 이용한 실험적방법에 의한 수식을 유도한 방법[9], 퍼지추론에 의한 방법[7], 그리고 학습에 의한 방법으로 유전알고리즘과 뉴로퍼지에 의한 방법[10]이 연구되고 있다. 본 논문에서는 유전알고리즘과 퍼지추론시스템의 합성을 이용하여 정수장 운전자료를 입력자료로 학습을 실시하였고, 만들어진 참음표(look up table)와 멤버쉽 함수(membership function)를 이용하여 학습에 사용하지 않은 100일간의 원수수질자료를 퍼지

연산하여 용집계주입률을 계산하였으며, 실제 운전 약품주입률과 비교검증을 실시하였다.

2. 정수처리공정

정수처리 약품주입공정은 크게 용집제 및 알칼리제 주입과 전염소, 후염소처리등으로 구분되고, 그 특징으로는 1)원수수질 측정(탁도, 알칼리도, pH, 수온, 전기전도도등)의 복잡성과 다양성 2)극심한 부하변화(기후, 시각, 및 계절에 따른 원수 수량과 수질의 큰 변화), 3)비선형 다변수 시스템, 4)제어량과 조작량의 선택, 조합의 불명확성을 들 수 있다[8]. 수질은 하절기에 고탁도, 고수온이고, 동절기에는 저탁도, 저수온인 뚜렷한 특징을 가지고 있고, 그 변화정도가 매우 커서 용집처리를 어렵게하는 원인이 된다.

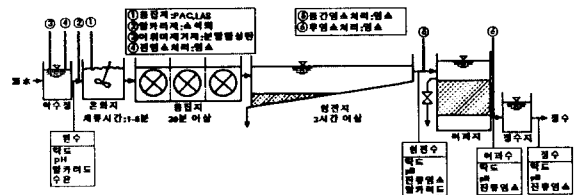


그림 1. 정수처리공정도.

Fig. 1. Schematics of water treatment process.

3. 유전알고리즘과 퍼지추론시스템의 합성

3.1 유전알고리즘

유전알고리즘은 진화의 법칙인 적자생존과 자연도태의 원리를 기반으로 정립된 최적화 알고리즘이다. 재생, 교배, 돌연변이 연산자에 의하여 탐색을 실시하는데 탐색영역에 다수의 탐색점을 분포시켜 목적함수값과 제한 조건의 오차정도에 따라 각 탐색점에 적합도를 부여한다. 재생은 문자열 선정방법으로 각 문자열은 정수처리공정의 약품주입률에 대한 적합도가 평가되며, 적합도가 클수록 다음 단계인 교배와 돌연변이의 과정에 참여할 확률을 크게하여 적합성이 좋은 유전자를 가진 탐색점들이 다음 단계에 보다 많이 형성되어 계산이 진행될수록 전체적인 탐색은 좋은 방향으로 진행된다. 재생에서는 적합도가 우수한 문자열이 선택되고, 적합도가 낮은 문자열은 도태되었지만 더 좋은 방향으로 진전되지는 못한다. 교배연산자에서는 선정된 한 쌍의 문자열에 대한 교배가 이루어져서 각 문자열에 대한 적합도의 변화가 이루어진다. 비록 재생과 교배가 효과적으로 탐색할 수 있지만 그들만으로는 유용한 문자열을 잃어버리거나 얻을 수 없기 때문에 돌연변이가 필요하다. 본질적으로 돌연변이는 문자열 공간을 통한 랜덤작업으로 재생 및 교배와 함께 사용될 때 중요한 유전인자의 조기 손실을 막아주거나 얻을 수 있다[2],[1].

3.2 퍼지추론시스템

비선형적이고 다변수적인 시스템을 대상으로한 퍼지 제어기와 퍼지 모델링 기법의 유용성은 이미 알려진 사실이며 이들은 퍼지 추론 시스템에 기초하고 있다. 퍼지 추론 시스템은 크게 퍼지 규칙과 추론엔진으로 구성되어 있다. 퍼지 규칙으로는 IF THEN에 의하여 입력변수 4개와 출력변수 1개로 된 총 625개의 규칙으로 구성되었다. 추론엔진으로는 최소-최대 무게중심(min-max gravity)법과 결과에서는 차이가 적으나 계산이 간단한 내적-합 무게중심(product-sum gravity)법을 사용하였다. 퍼지추론에 사용되는 멤버십 함수는 삼각형으로 수질변동폭을 균등 분할한 후 학습을 통하여 점차 입력변수특성에 적합하게 탐색하여 사용하였다[4],[6].

3.3 유전알고리즘과 퍼지추론시스템의 합성

퍼지추론시스템에서 가장 어려운 점은 찾을표와 멤버십 함수를 구성하는 일이다. 고도의 비선형 다변수특성을 갖는 시스템의 규칙을 단순히 전문가의 경험적 지식이나 실험에 의해 구성하는 일은 상당히 주관적이고 경험에 의존하기 때문에 최적의 찾을표를 결정하기란 쉽지 않다. 유전알고리즘은 정점 근방까지는 빨리 수렴하지만 최종 수렴치를 보증하지는 못한다. 이와 같은 단점을 서로 보완하고자 퍼지추론시스템에서의 찾을표(Look up table)와 멤버십 함수를 유전알고리즘의 문자열로 변환하여 학습을 통한 최적의 퍼지시스템구성을 달성하고, 퍼지추론에 의하여 최종목표인 응집제 주입률을 결정하였다.

최적의 응집제 주입률 결정을 위하여는 궁극적으로 식(1)의 목적함수값 E를 최소화시키는 찾을표와 멤버십함수를 구성하는 것이다. 반면에 유전알고리즘에서는 식(2)의 적합도 함수를 최대화시키도록 진화가 이루어지므로 최소화 문제를 최대화문제로 변경하여야 한다. 변환 방법에는 여러 가지가 있지만 본 논문에서는 식(2)와 같이 목적함수의 역수를 적합도 함수(fitness function)로 사용하였다.

$$E = \frac{1}{N} \sum_{i=1}^N \frac{|RLAS-CLAS|}{RLAS} \quad (1)$$

$$f = 1/E \quad (2)$$

여기서 RLAS는 실제 주입된 응집제 주입률, CLAS는 퍼지 추론에 의하여 계산된 응집제 주입률 그리고 N은 자료의 갯수이다.

복수 파라메타로 구성된 최적화 문제를 푸는데 성공적으로 사용된 문자열의 코딩 방식은 다수의 단일 파라메타를 직렬로 연결하는 것이다. 찾을표의 경우에 입력변수 4가지에 대하여 퍼지 변수의 수를 5로 중복순열에 의하여 $5 \prod 4 = 625$ 가지의 퍼지규칙이 필요하고, 멤버십 함수의 경우에 퍼지변수 5개에 멤버십 함수의 기울기 변수 1개와 좌우 이동 변수 1개가 필요하여 입력변수가 4가지이므로 40개의 문자로 구성하였다.[5]

유전알고리즘은 여러 방향에서 탐색을 통하여 최적값으로 수렴하지만, 본 논문에 적용한 결과 원수수질이 급격하게 변하는 자료에 대하여는 주입률 계산을 포기하고, 나머지 자료에 대해서 정확한 계산을 하므로써 전체적인 적합도를 향상시키는 방향으로 탐색이 진행되었다. 이러한 부분적인 최적화를 방지하기 위하여 적합도함수의 계산에서 각 원수수질자료의 오차가 한계치를 초과하면 도태시키는 방법을 적용하였다.

찾을표(look up table) 최적화에서는 입력변수의 조건에 따른 출력변수의 상태값으로 구성된 문자열을 유전연산하여 적합도가 가장 우수한 찾을표를 선정한다. 이렇게 선정된 찾을표는 멤버십 함수를 최적화하거나, 응집제 주입률을 결정하기 위하여 퍼지추론할 때 이용된다.

찾을표의 최적화에 부가하여 멤버십함수(membership function)를 최적화시키면 좀더 빠르고, 정확한 계산이 가능해진다. 멤버십 함수는 초기에 수질항목별로 변화범위를 균등 분할하여 초기화하였는데 이를 다시 조절하여 수질항목특성에 적합한 형태로 변형시켜 좀더 정확하고 빠르게 수렴하도록 하기 위하여 기울기 변경과 좌우 이동을 하게 된다.

식(3)의 δ_i 는 퍼지 멤버십 함수의 좌우 이동량을 나타내는 문자이고, η_i 는 이동변 삼각형의 기울기를 조절하는 변수를 나타내는 문자이다.

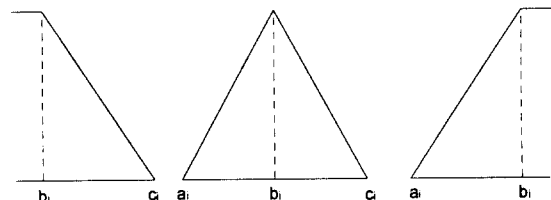


그림 2. 퍼지 멤버십 함수.

Fig. 2. Fuzzy membership function.

$$\begin{aligned} a_i &= (a_i + \delta_i) - \eta_i \\ b_i &= (b_i + \delta_i) \\ c_i &= (c_i + \delta_i) + \eta_i \end{aligned} \quad (3)$$

유전알고리즘에서는 이들 조절량을 변환해서 최적의 멤버십 함수를 탐색한다.

4. 전산모사

4.1 유전알고리즘과 퍼지추론시스템의 모델링

정수처리공정의 응집제 주입제어의 모델링은 Fig. 3과 같다. 최근의 운전실적자료를 이용하여 유전알고리즘과 퍼지추론시스

템의 합성에서 충분한 탐색을 통하여 최적의 찾음표(look up table)와 멤버십 함수가 결정되면, 퍼지추론에 의하여 온라인으로 응집제 주입제어를 실시하고, 정기적으로 과거운전실적자료를 이용하여 찾음표와 멤버십 함수를 충분히 탐색한 후 갱신하게 된다.

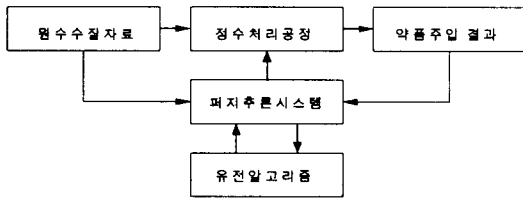


그림 3. 정수처리공정의 약품주입률 모델링.
Fig. 3. Block diagram of dosages control for water treatment process.

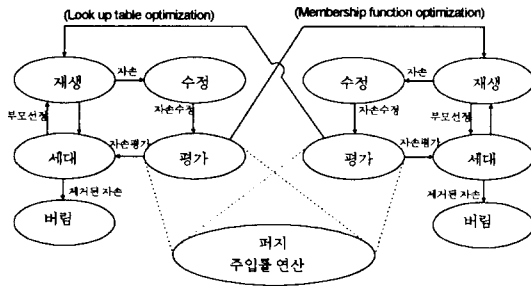


그림 4. 찾음표와 멤버십함수의 최적화 탐색과정.
Fig. 4. The cycle of optimization with genetic algorithms for look up table and membership function.

유전알고리즘과 퍼지추론시스템의 합성을 이용한 찾음표와 멤버십 함수의 최적화 탐색과정은 Fig. 4와 같다. 탐색과정은 찾음표를 최적화하는 부분과 멤버십 함수를 최적화하는 부분으로 구분되어 있다. 찾음표를 최적화하는 부분에서는 유전알고리즘에 의하여 최적의 찾음표를 탐색하다가 미리 설정한 일정반복회수나 적합도한계를 초과하면 첫 번째 문자열 즉 찾음표를 멤버십 함수를 최적화하는 부분으로 인계한다.

멤버십 함수를 최적화하는 부분에서는 찾음표 최적화 부분에서 인계된 찾음표와 초기 상태에서 주어진 멤버십 함수를 시작으로 멤버십 함수를 최적화하기 위하여 탐색을 실시한다. 찾음표 최적화 부분으로 인계되는 것은 유전알고리즘에 사용하는 문자열이 아니고, 이 문자열에 의하여 조정되는 멤버십 함수인 a_i , b_i , c_i 를 인계하게 된다. 초기의 a_i , b_i , c_i 는 3년간 원수수질변화를 고려하여 Table 1과 같이 정해지는데, 이 멤버십 함수 변수로 입력변수 1개당 13개의 값을 가진다. 입력변수가 4개이므로 총 52개의 멤버십 함수가 실제 퍼지추론에 사용된다.

이와 같은 과정을 반복하여 멤버십 함수 최적화 부분에서 설정한 반복회수나 설정한 적합도 한계치를 초과하면 가장 우수한 멤버십 함수를 찾음표를 최적화하는 부분으로 인계되는데, 이때 목표 적합도나 반복회수를 초과되면 반복을 정지한다. 이렇게 작성된 찾음표와 멤버십 함수를 가지고 자동 측정되는 원수수질을 입력자료로 응집제주입률을 결정하게 된다.

표 1. 입출력 변수의 초기설정치.

Table 1. Threshold values of input and output variables.

항목	NB	NS	ZE	PS	PB
탁도	5.0	25.0	45.0	65.0	85.0
pH	6.5	7.0	7.5	8.0	8.5
알칼리도	20.0	35.0	50.0	65.0	80.0
수온	0.0	10.0	20.0	30.0	40.0
LAS주입률	20.0	35.0	50.0	65.0	80.0

4.2 전산모사

탐색과 검증에 사용된 자료는 1992년도에서 1994년도까지 3년간 한국수자원공사 금강수도사무소의 일간 운전실적자료를 사용하였다. 탐색에서의 반복횟수는 3만번 이상으로 충분한 탐색을 실시하였으며 탐색에 소요되는 시간은 3000번 반복시 80486-66MHz PC로 53분 15초가 소요되고, 이를 10회이상 실시하였다.

Table 2와 같이 초기설정을 하여 탐색과 성능평가를 실시하였으며 전산모사에 사용한 유전알고리즘과 퍼지추론시스템의 합성은 C로 프로그래밍하였다.

표 2. 초기 설정값.

Table 2. Value of initial setting.

세대수	8
교배율	0.7
돌연변이율	0.05
교배방법	Uniform
부모문자열 선정방법	Roulette wheel

5. 결과 및 고찰

5.1 찾음표와 멤버십 함수 최적화

찾음표와 함께 멤버십 함수를 최적화하여 최적의 퍼지추론시스템을 구성하였다. Fig. 6과 Fig. 7은 퍼지추론시스템에서 비교 검증하기 위한 원수수질자료의 변화를 나타내고 있다. 이 변화에 따라 실제 주입된 주입률과 유전알고리즘과 퍼지추론시스템의 합성을 이용한 전산모사에서 계산된 주입률과의 비교를 Fig. 5에 나타내었다.

실제 주입된 주입률의 단위가 5ppm임을 감안할 때, 실제 주입된 주입률보다는 원수수질의 변화에 민감하게 작용하고 있음을 알 수 있다. Fig. 8은 탐색과정에 발생하는 적합도의 변화상태를 나타내고 있으며, Fig. 9은 학습에 의한 결과로 나타난 탁도의 멤버십함수를 보여주고 있다

5.2 고찰

비선형 다변수적인 특성과 수질변화가 심한 하천에서 취수하는 경우, PAC에 비하여 효과가 적은 LAS를 이용하는 등 악조건에 적용하여도 찾음표와 멤버십 함수를 동시에 최적화하는 경우는 만족하게 응집제주입률을 결정함을 확인하였다.

본 유전알고리즘과 퍼지추론시스템의 합성을 이용한 정수처리공정의 응집제주입률 결정은 발생하는 오차의 원인이 운전자료의 오차인지 퍼지추론에서의 오차인지를 확인하여 약품주입의 자동화를 유도해낼 수 있을 것이다.

6. 결론

비선형이고 다변수적인 특성을 가지면서 극히 열악한 조건에 있는 정수처리장의 응집제주입률 결정에 유전알고리즘과 퍼지추론시스템의 합성을 적용하여 수질변화에 만족하게 추적하고 있음을 확인하였고, 다음과 같은 결론을 얻었다.

- 1) 퍼지추론시스템을 구성하는데 가장 어려운 문제인 찾을표와 멤버쉽 함수를 자동으로 최적의 시스템을 구성할 수 있었다.
- 2) 각 정수장에 적용할 경우 실험이나 계산이 필요없고, 모든 종류의 약품에 적용이 가능하며, 한정적이거나, 한시적이지 않은 시스템으로 과거운전실적자료만 있으면 구성이 가능하다.
- 3) 실증시험을 통한 오차의 원인규명과 최적주입률이 확인되면 응집제등의 약품주입 자동화가 가능할 것으로 판단된다. 향후과제로는 오차를 좀더 줄이는 방안과 오차의 원인을 규명하는 연구가 있어야 할 것이다.

참고문헌

- [1] Lawrence Davis, *Handbook of Genetic Algorithms*, Van Nostrand Reinhold Newyork, 1991.
- [2] David E. Goldberg, *Genetic algorithms in search, optimization, and machine learning*, Addison-wesley, 1989.
- [3] V. S. Gupta, S. K. Bhattacharjya, B. K. Dutta, "Zeta-Potential Control for Alum Coagulation", *American Water Work Association*, pp. 21-22, 1975. 1.
- [4] M. Mizumoto, "Fuzzy Controls under Product-Sum-Gravity Methods and New Fuzzy Control Methods", *Fuzzy Control System*, CRC Pres. Inc, 1994.
- [5] Daihee Park, Abraham Kandel, and Gideon Langholz, IEEE "Genetic-Based New Fuzzy Reasoning Models with Application to Fuzzy Control". *IEEE Trans on System, Man, and Cybernetics*. Vol. 24, No, 1 1994.
- [6] Timothy J. Ross, *Fuzzy logic with engineering application*, MCGraw hill, 1995.
- [7] Tomohiro Takagi, Michio Sugeno "Fuzzy Identification of Systems and Its Applications to Modeling and Control", *IEEE Trans on System, Man and Cybernetics*, Vol. SMC-15, NO. 1, 1985.
- [8] 建設部, "上水道施設基準", 1992. 12.
- [9] 김홍식, 김성현, "Jar-Test를 이용한 凝集劑 注入率 決定에 관한 實驗研究", *대한상하수도학회지*, 제2호, pp. 39-46, 1993. 6.
- [10] 하대원, 고택범, 황희수, 우광방, "상수처리 수질제어를 위한 약품주입 자동연산", *한국자동제어학술회의 논문집*, pp. 772-777, 1993.
- [11] 황희수, 오성권, 우광방, "유전알고리즘과 퍼지 추론 시스템의 합성", *전기학회논문집*, 41권, 9호, pp. 1095-1103, 1992.

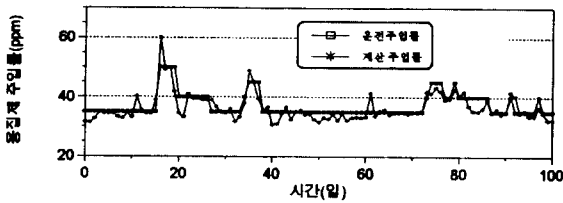


그림 5. 응집제 주입률의 비교.
Fig. 5. Comparison of feeding rate.

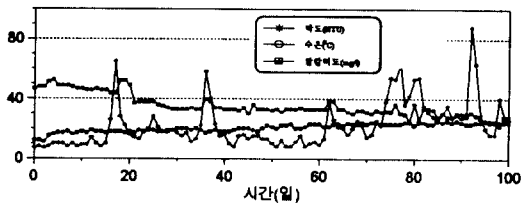


그림 6. 원수의 탁도, 수온, 알칼리도 변화
Fig. 6. Variation of turbidity, alkalinity, water temperature.

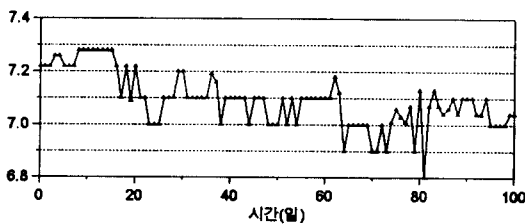


그림 7. 원수의 pH변화.
Fig. 7. Variation of pH.

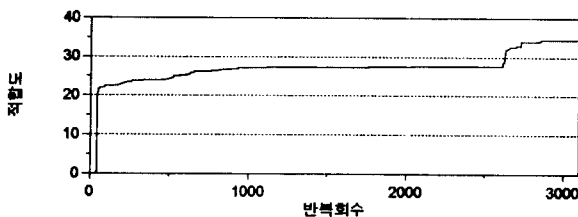


그림 8. 적합도의 수렴.
Fig. 8. Convergency of fitness value.

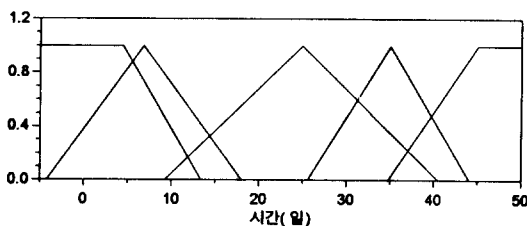


그림 9. 탁도 멤버쉽 함수.
Fig. 9 Membership function of turbidity.