

# 유전자 알고리즘의 관개배수 분야 활용

## Application of Genetic Algorithm to Irrigation and Drainage

윤 광 식\*  
Yoon, Kwang-sik

### 1. 머리말

유전자 알고리즘(Genetic algorithm)은 자연 세계의 진화과정을 컴퓨터 상에서 시뮬레이션 함으로써 복잡한 실세계의 문제를 해결하고자 하는 계산 모델이다. 유전자 알고리즘은 구조가 간단하고 방법이 일반적이어서 응용범위가 매우 넓으며, 특히 적응적 탐색과 학습 및 최적화를 통한 공학적인 문제의 해결에 많이 이용되고 있다. 또한, 최근에는 신경망, 퍼지로지과의 결합으로 그 응용범위는 점점 늘어나고 있는 추세이다.

세계적으로 여러 분야에서 유전자 알고리즘을 이용하여 복잡한 문제의 최적해를 얻기 위한 연구와 적용이 활발하게 이루어지고 있으며, 수자원 및 관개배수 문제에 적용한 여러 연구 결과들이 발표되고 있다. 그러나 우리 나라의 경우 수자원 분야 또는 관개배수 분야에 이 새로운 기법을 이용해야 한다는 논의가 있으나 이의 이용은 미미한 편이다. 본 소고에서는 유전자 알고리즘에 대한 기본적인 내용을 소개하고, 관개배수 분야에의 활용방안을 살펴보고자 한다.

### 2. 유전자 알고리즘의 개요

진화 알고리즘 또는 유전자 알고리즘은 적

자생존(Survival of fittest)과 자연도태(Natural selection)의 자연의 진화법칙에 기초한 탐색 알고리즘(Search algorithm)이다. 자연진화의 법칙에 의하면 일정한 환경에 무리를 지어 살고 있는 생명체들은 그 환경에 적합한 형질을 가진 개체는 생존할 확률을 크게 가지고 교배와 변이의 과정을 통하여 계속적으로 더 좋은 방향으로 진화해간다. 유전자 알고리즘은 이와 같이 생물의 진화과정을 인공적으로 모델링한 알고리즘이다.

1950년대 후반 컴퓨터 공학자와 생물학자가 초기 컴퓨터에 유전자의 기작을 모델링하기 위하여 시작된 이래, 1970년대에는 미시건대학교의 John Holland와 그의 연구진들이 기존의 유전자 알고리즘에 대한 연구를 종합하여 체계화 시켰으며, 동료와 제자들에 의해 개발되어 왔다.

#### 가. 유전자 알고리즘의 기본 용어

자연계에 있는 생물의 진화과정에 있어서, 어떤 세대(Generation)를 형성하는 개체(Individual) 들의 집합, 즉 개체군(Population) 중에서 환경에 대한 적합도(Fitness)가 높은 개체가 높은 확률로 살아남아 재생(Reproduction)할 수 있게 되며, 이때 교배(Crossover) 및 돌연변

\* 전남대학교 농공학과

이(Mutation)로서 다음 세대의 개체군을 형성하게 된다.

유전자 알고리즘에서 개체의 수를 개체군의 크기(Populationsize)라고 한다. 각각의 개체는 염색체(Chromosome)를 가지고 있으며 염색체는 복수개의 유전자(Gene)의 집합으로 구성된다. 유전자의 위치를 유전자좌(Locus)라 하고 유전자가 취하게 되는 유전자의 후보를 대립 유전자(형질, Allele)라고 한다. 생물의 경우 염색체는 어떤 개체의 특징을 상세하게 결정하게 되는데 예를 들어 머리가 검은 것은 염색체 중에 이러한 특징을 나타내도록 하는 유전자의 조합이 존재하기 때문이다. 이와 같이 유전자에 의해 결정되는 개체의 형질을 표현형(Phenotype)이라고 하고, 이에 대응되는 염색체의 구조를 유전형(Genotype)이라 한다. 여기에서 표현형이 여러 개의 유전자좌의 영향을 받아 복잡한 형태가 결정되는데 이것을 에피스타시스(Epistasis)라고 한다. 또한 표현형을 유전형으로 바꾸는 것을 코드화(Coding) 그 역을 디코드화(Decoding)라고 한다. <표 - 1>은 생물학과 유전자 알고리즘간의 용어를 나타낸 것이다.

<표 - 1> 생물학과 유전자 알고리즘의 용어 비교

생물학	유전자 알고리즘
염색체(Chromosome)	문자열(String)
유전자(Gene)	특성(Feature), 형질(Character)
대립 유전자(Allele)	특성치 (Feature value)
유전자좌(Locus)	문자열의 위치(String position)
유전형(Genotype)	구조체(Structure)
표현형(Phenotype)	파라미터 집합(Parameter set) 대체해(Alternative solution), 디코드화를 위한 구조체 (A Decoded structure)
에피스타시스(epistasis)	비선형성(Nonlinearity)

나. 유전자 알고리즘의 연산과정

유전자 알고리즘의 연산과정은 첫번째 집단

를 만드는 것으로 시작하며 하나의 집단이 만들어진 후에는 여러 가지 과정을 거친 다음 새로운 세대를 만들어야 한다. 이때 만들어진 새로운 세대는 앞 세대보다 환경에 좀더 적응하는 유전자를 가진다.

유전 알고리즘에서 새로운 세대를 만들기 위해서 평가, 선택, 도태, 교배, 변이의 방법을 사용한다. 이러한 방법에 의해 해에 대한 오차항을 만족할 때까지 반복하는 과정으로 진행된다.

(1) 초기집단

유전자 알고리즘에서 한 집단은 유전자들의 집합으로 구성되어 있다. 그렇기 때문에 유전 알고리즘을 시작하기 위해서는 하나의 세대를 구성하여야한다. 각 변수를 2진수로 코딩(Coding)하여 하나의 유전자를 구성하게 된다.

(2) 평가

평가단계는 현세대의 유전자가 환경에 대해 갖는 적응 정도를 조사하는 부분으로 유전자 알고리즘에서는 환경에 대한 적응정도를 적합성(Fitness)이라는 값으로 평가하게 된다. 적합성이 좋은 유전인자 일수록 다음 세대에 살아남아야 하고, 적합성이 나쁜 인자 일수록 다음 세대로 유전되어서는 안 된다. 적합성을 조사하는 방법은 문제에 따라 다르다.

(3) 선택

각 유전자는 적합성의 크기에 따라 추후 교배와 변이의 과정에서 선택될 확률을 부여받게 된다. 즉 적합성이 큰 유전자는 많이 선택되도록 하고 적합성이 나쁜 유전자는 적게 선택되도록 한다.

(4) 교배

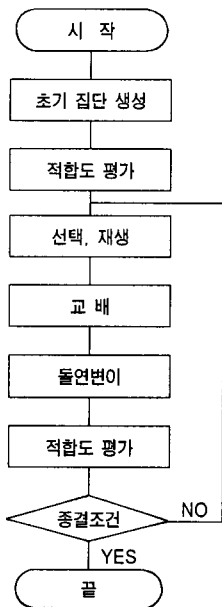
선택의 과정을 통하여 좋은 개체들은 유전되고 나쁜 개체들은 소멸되었지만 더 좋은 방향으로 진화하지는 못했다. 따라서 개체들의 적합성 향상을 위해 생명체의 유전인자 교환과 같은 교배과정을 거친다.

(5) 변이

교배에 의해 개체들이 너무 한쪽 방향으로만 치우치는 것을 보완하기 위하여 아주 작은 확률(일반적으로 0.001)로 각 유전자의 한 문자를 아무런 이유 없이 변이시킨다.

유전자 알고리즘은 이상 설명한바와 같이 전체 전개과정은 비교적 간단하지만 다수의 세대 집단과 도태 교배 그리고 변이의 과정이 합쳐져서 만들어낸 결과는 매우 강력하다고 할 수 있다.

기본적인 단순 유전자 알고리즘(Simple Genetic Algorithm : SGA) 및 흐름도는 다음과 같다.



<그림 - 1> 단순 유전자 알고리즘 및 흐름도

3. 유전자 알고리즘의 활용분야

진화 알고리즘은 <표 - 2> 에서와 같이 다양한 분야에 적용되고 있는데, 이와 같은 많은 응용의 배경에는 몇 가지 이유가 있다. 그 한 가지는 알고리즘의 단순성과 일반성에서 찾을

수 있다. 아무리 복잡한 문제라도 일단 염색체 형태로 표현이 되면, 염색체의 복제와 재결합 및 적합도의 평가 등과 같은 비교적 단순한 연산과정의 반복을 통해 계산이 수행된다.

<표 - 2> 유전자 알고리즘의 응용 분야

응용분야	응용 사례
최적화	수학적 함수의 최적화, 가스 파이프라인의 최적화, 전력 송전망의 최적화, 컴퓨터 자판의 최적배정문제, 항공기 승무원 배정 문제
설 계	VLSI회로 설계, 비행기 날개의 공기 역학적 설계, 엔진 노즐의 설계, 컴퓨터 통신망의 최적 설계, 심장박동기의 설계
인공지능	LISP 프로그램의 자동생성, 문제해결 규칙의 자동 습득, 신경망의 합성 및 학습, 패턴 인식, 자연언어 처리, 멀티 에이전트 시스템
시스템분석 및 예측	시스템 동정, 케이오틱 시계열의 예측, 환율 변화 예측, 단백질의 구조 분석, 재정 및 경제 분야에서의 예측 및 분석 문제
제어 및 로보틱스	도립진자의 제어, truck backer upper problem, 이동 로봇의 경로 계획, 자율 이동 로봇군의 군 제어, 신경망 및 퍼지로직과 유전 알고리즘의 결합에 의한 제어

또 하나의 이유는 기존의 문제 해결방법과 결합하여 사용하기가 쉽다는 것이다. 이것은 진화 알고리즘이 문제해결에 특수한 정보를 많이 사용하지 않고 또한 문제에 대한 배경 지식이 있으면 이를 쉽게 수용할 수 있기 때문이다. 예를 들어 문제 P에 대한 기존의 경험적 알고리즘 A가 존재한다면, 먼저 A를 사용하여 여러 개의 가능한 해들을 생성한 다음 이를 초기의 개체군으로 생각하고 그 위에 진화 알고리즘을 적용함으로써 적어도 기존의 알고리즘 A의 결과보다는 더 우수한 결과를 얻을 수 있다.

진화 알고리즘이 흥미를 끄는 또 하나의 이유는 계산 모델이 자연 현상에 기반을 두고 있다는 것이다. 인간이 오랫동안 풀려고 했던 많은 실세계의 문제들을 자연은 실제로 유전

하게 해결하고 있다. 현재 추세로 간다면 진화 알고리즘의 응용범위는 점점 더 늘어날 것이며 이론적인 발전도 가속화되리라 예상된다. 특히 지능 제어와 관련하여 신경망과 퍼지 로직과의 결합으로 적응 학습 및 최적화의 문제에 새로운 해결법으로 많은 새로운 결과들이 나올 것으로 예상된다

일반적으로 다음과 같은 특성을 갖는 문제에 대해서는 유전자 알고리즘 계산방법이 효과적으로 적용될 수 있다는 것이 많은 응용결과에 의하여 확인되었다.

- 여러 개의 국부적인 해가 존재하는 문제 (multi-modal)
- 어느 정도의 규칙성을 가지고 있는 문제 (regularity)
- 문제영역의 규칙성이 어느 정도 염색체로 표현 가능한 문제
- 부분적인 해들 간에 상대적인 우위관계가 존재하는 문제(Epistasis)

#### 4. 관개배수 분야에의 활용

수자원 및 관개배수 분야의 대표적인 유전자 알고리즘의 이용은 수자원 계획의 최적화, 수학적 모형의 최적 매개변수 추정, 관개조직 설계의 최적화 등을 들 수 있겠다. 유전자 알고리즘이 다른 탐색이나 최적화 방법과 다른 점은 다음과 같다.

- 파라미터를 코딩한 것을 직접 이용한다. 유전 알고리즘은 변수로 실제 값을 직접 사용하는 대신, 2진수 코드 형태의 문자열을 사용한다. 이러한 문자열 형태의 변수의 사용은 염색체에 유전인가 있는 것과 같은 형태를 지니고 있어, 교배와 변이의 변환 과정을 수행하기가 매우 단순하고 편리하다.
- 점(Point)이 아닌 다점(Multi points, : 군

(Population)) 탐색 방법으로 다수의 조건을 집단적으로 탐색할 수 있다. 여러 개의 개체들이 개체군(Population)을 이루어 동시에 탐색을 행하기 때문에 보다 넓은 영역에 대한 정보를 활용한다.

- 직접탐색방법이다. 탐색에 비용 정보(fitness function)를 이용하며, Blind search를 한다. 유전자 알고리즘은 목적 함수와 제한 조건의 값만을 사용하고 미분 값이나 그 외의 다른 정보를 필요하지 않은 직접 탐색방법이므로 복잡하고 다양한 문제의 최적화 문제에 적합하며 또한 기본 모델의 변경으로 인한 수정이 용이하다.
- 결정론적인 규칙이 없고 확률적 연산자를 사용하여 수행된다.

이와 같은 특징으로 인해 다른 탐색 또는 최적화 방법 중 하나인 계산에 의존한 방법(Calculus-based method)에 비하여 전역적 해를 구할 가능성이 높으며 다른 여러 탐색 방법에 비하여 효율적이다. 다음으로 유전자 알고리즘의 실제 적용 예를 살펴보도록 한다.

#### 가. 수학적 모형 매개변수 추정 예

많은 수문, 수리, 관개배수 모형은 정확한 예측을 위해 현장 실측 결과를 이용 매개변수들(Parameters)을 보정(Calibration)한 후에 사용되고 있다. 매개변수 추정을 위해 대개 시행착오 방법을 이용하고 있으나, 최적의 매개변수 값을 찾는 것은 용이한 일이 아니다. 모형의 최적의 매개변수 값을 찾는데 유전자 알고리즘이 유용하다는 연구결과들이 발표된 바 있다(Wang, 1991; Liong, 1995).

Liong(1995) 등은 SWMM(Storm Water Management Model)을 이용 6.1km<sup>2</sup> 유역의 유출해석시 유전자 알고리즘을 이용하여 매개

변수를 보정하였다. SWMM 모형의 매개변수 중 8개 매개변수에 대해 최적의 조합을 찾으려 하였다. 각 매개변수 값은 유전자 값으로 표현하고 8개의 유전자를 가진 염색체를 생성하였다. 각 유전자의 초기 값으로 즉 매개변수의 초기 시도 값으로 0.5부터 1.5까지 값을 배정하고 실측 침투유량과 모의발생 값을 비교하면서 새로운 염색체를 생성하였다. 이때 교배비율을 0.6, 변이확률을 0.001로 하였다. 대략 50번 정도의 시도 후 선정된 염색체내의 유전자 값들을 매개변수 값으로 이용시 모의 발생 값은 실측 침투유량과 10% 내외의 오차 기준을 만족하는 것으로 밝히고 있다.

A	B	C	D	E	F	G	H	유전자 위치
0.50	0.89	0.77	1.02	1.48	0.99	0.69	0.74	유전자 값
1	2	3	4	5	6	7	8	유전자 위치
0.60	0.92	0.64	1.44	0.88	0.79	1.02	0.88	유전자 값
A	B	C	D	5	6	7	8	새 염색체 1
1	2	3	4	E	F	G	H	새 염색체 2

〈그림 - 2〉 교배를 통한 새로운 염색체 생성 예

나. 관망해석 예

최근 관망 해석에 유전자 알고리즘을 적용한 사례 (Murphy 등, 1993 ; Simpson 등, 1994 ; Dandy 등, 1996)가 늘고 있으며 기존의 Linear Programming, Non-Linear Programming, Dynamic Programming에 비해 더 우수한 기법이라는 연구 결과(Simpson 등, 1994)가 발표되기도 하였다. Dandy 등(1996)은 미국 뉴욕시 수도 터널 문제에 개선된 유전자 알고리즘을 적용하였으며, 기존의 최적화 방법은 하나의 답만을 내는데 비해 유전자 알고리즘은 선택 가능한 여러 대안 값을 찾아 주는 것을 장점으로 꼽았다. 유전자 알고리즘의 수자원 및 관개배수분야

에의 이용은 미국과 같은 선진국의 경우에만 한정된 것은 아니다. 공산주의 경제체제에서 자본주의 경제체제로 전환의 변화를 겪고 있는 동구권 국가들은 사유재산권 확대에 따른 좀더 생산성 있는 농업기반시설 요구에 직면하고 있으며, 그 예로 슬로바키아 공화국의 경우 주 관개시설이 스프링클러인데 대부분의 시설이 노후하며 시설용량이 현재의 요구에 미치지 않아 관개시설의 개보수 및 신설이 절실히 요구되고 있다. 과거의 대부분의 관개조직은 분기형(Branched network) 관개 조직이었는데, 새롭게 회로형 관망(Looped network) 조직 설계의 필요성에 따라 이의 최적 설계를 위해 유전자 알고리즘을 이용한바 있다(Cisty 등, 1999).

단순 유전자 알고리즘을 이용한 관망해석 최적화를 위한 과정은 Simpson 등(1994)에 의해 잘 기술되어 있다. 그 내용을 요약하면 다음과 같다.

(1) 초기집단(Initial Population)

유전자 알고리즘은 임의로 코드화된 문자열(Coded string)로 된 N개의 초기집단을 생성하는데 각각은 시행 관망 제원 값을 나타낸다. 통상적으로 초기 집단 수는 100에서 1,000개 정도로 한다. 문자열은 1과 0으로 조합되며, 문자열은 가능한 파이프 크기의 조합을 나타낸다. 예를 들어 21개의 파이프로 이루어진 관망을 나타내기 위해서는 각각 4bit인 21개의 부분 문자열(Substring) 조합을 통한 84bit의 2진수 문자열이 생성된다. 이 경우 각각 4bit인 부분 문자열은 각 파이프에 선택 가능한 16개 파이프 관경을 나타낸다.

(2) 관망 소요경비 산출

유전자 알고리즘은 집단 안의 N개의 문자열을 차례로 해석한다. 문자열의 부분 문자열(Substing)을 해독하여 파이프 크기와 전체 재료와 건설비용을 계산한다. 유전자 알고리즘은

현재 집단 안에 있는 모든 시행(Trial) 관망설계의 비용을 결정한다.

### (3) 각 관망의 수리학적 해석

현재 집단에 포함된 모든 N개의 관망설계에 대해 정상류 관망해석 프로그램으로 압력과 유량을 계산한다. 분지점(Nodal point)의 압력수두는 최소 허용압력 수두와 비교되고, 이상유무는 기록된다.

### (4) 벌칙비용 계산

만일, 관망설계가 최소압력 제한을 만족시키지 못하면 유전자 알고리즘은 벌칙비용(Penalty cost)를 계산한다.

### (5) 관망의 총비용 계산

현재 세대에 포함되어 있는 모든 관망에 대해 총 비용을 계산한다. 총 비용은 2단계에서 구한 관망비용과 4단계에서 구한 벌칙비용의 합이다.

### (6) 적합성 계산

현재 세대에 있는 모든 관망의 경우에 대해 각 관망의 총비용의 역수 값으로 적합성을 판정한다.

### (7) 선택

적합성 판정결과에 따라 선택연산자 (Selection operator)를 이용 새로운 집단을 탄생시킨다.

### (8) 교배연산자

새로운 집단을 교배 연산자(Crossover operator)를 통해 새로운 문자열을 만들어낸다.

### (9) 변이연산자

교배과정을 거친 문자열을 변이연산자(Mutation operator)를 이용 일정확률에 따라 문자열안의 bit의 값을 반대의 값으로 (0 을 1로 또는 1 을 0으로) 변이시킨다

### (10) 연속적인 세대 생성

2번에서 9번까지의 과정을 반복하여 새로운 세대를 생성한다. 최소 비용을 만족하는 문자열은 기록되고 더 적은 비용을 만족시키는 대

안이 나타나면 순위를 바꾼다. 이렇게 해서 최적의 관망 설계안을 얻는다.

## 5. 맺는 말

유전자 알고리즘 이용이 수자원 및 관개배수 문제해결에 유용하다는 외국의 연구 결과가 많이 나오고 있지만, 아직까지 생소한 기법이며 이를 현실문제에 적용하기에는 어려움이 있다. 이의 해결을 위해서는 국내의 사례에 적용하는 연구들이 많이 이루어져서 그 유용성이 입증되어야 하겠고, 자세한 적용방법이 소개되어야 하겠다. 또한 유전자 알고리즘을 이용한 기 개발된 프로그램들에 대한 소개와 이용 평가가 필요하며 국내 관개배수 문제 해결에 적용될 수 있는 범용 프로그램 개발이 뒤따라야 할 것이다. 수학적 모형 매개변수 추정, 관수로 수리해석, 관관계획, 수리시설물 노선 설정, 집중 물관리 등에 유용하게 쓰일 수 있을 것으로 여겨진다. 새로운 관개배수 분야 연구로 지금까지의 설계방법을 개선하기 위해 유전자 알고리즘을 적용하는 시도를 기대해본다.

### 참고문헌

1. 장병탁, 1995. 유전 알고리즘의 이론 및 응용, 대한전자공학회지, 22(1), pp.1290~1300.
2. 장병탁, 1995. 인공 진화에 의한 학습 및 최적화, 제어·자동화·시스템공학회지, 1(3), pp.52~61.
3. Bäck, T. , H.P. Schwefel, 1993. An overview of evolutionary algorithms for parameter optimization, Evolutionary Computation, 1(1), pp.1~23.
4. Cisty Milan, D.A. Savic, and G.A. Walters, 1999. Rehabilitation of pressurised pipe

- networks using genetic algorithms, paper presented in the 1999 ICID Conference.
5. Dandy, G.C., A.R. Simpson, and L.J. Murphy, 1996. An improved genetic algorithm for pipe network optimization, *Water Resources Research*, 32(2), pp.449~458.
  6. Goldberg, D.E., 1989. *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison Wesley.
  7. Kinnear, jr K.E., 1994. *Advanced in Genetic Programming*, MIT Press.
  8. Koza, J.R., 1992. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, MIT Press.
  9. Koza, J.R., 1994. *Genetic Programming II: Automatic Discovery of Reusable Programs*, MIT Press.
  10. Liong, S.Y., W.T. Chan, and J. ShreeRam, 1995. Peak-Flow Forecasting with Genetic Algorithm and SWMM, *J. of Hydraulic Engineering*, ASCE, pp.613~617.
  11. Michalewicz, Z., 1995. *Genetic Algorithms + Data Structures = Evolution Programs*, Springer-Verlag, pp.160~168.
  12. Murphy, L.J., A.R. Simpson and G.C. Dandy, 1993. Design of a network using genetic algorithm, *Water*, 20, pp.40~42.
  13. Simpson, A.R., G.C. Dandy, and L.J. Murphy, 1994. Genetic algorithms compared to other techniques for pipe optimization, *J. Water Resour. Plann. Manage.* ASCE 120(4), pp.423~443.
  14. Wang, Q. J., 1991. The genetic algorithm and its application to calibrating conceptual rainfallrunoff models, *Water Resour. Res.*, 27(9), pp.2467~2471.
  15. 坂和正敏, 田中雅傳, 1995. 遺傳的アルゴリズム, 朝倉書店.
  16. 北野宏明 外, 1993. 遺傳的アルゴリズム, 産業圖書.
  17. 伊庭齊志, 1994. 遺傳的アルゴリズムの基礎, オーム社.
  18. 上田完次, 1995. 人工生命の方法, 工業調査會.
  19. 玉置 久, 喜多 一, 岩本 貴司, 三宮 信夫, 1995. "遺傳アルゴリズム-I -GAの基礎," *システム/制御/情報*, 39(6), pp.295~302.
  20. 玉置 久, 喜多 一, 1995. 遺傳アルゴリズム-II -GAによる最適化計算1, *システム/制御/情報*, 39(8), pp.400~407.
  21. 玉置 久, 喜多 一, 1995. 遺傳アルゴリズム-III -GAによる最適化計算2, *システム/制御/情報*, 39(10), pp.431~438.
  22. 岩本 貴司, 1995. "遺傳アルゴリズム-IV -GAの理論," *システム/制御/情報*, 39(12), pp.507~514.
  23. 玉置 久, 喜多 一, 岩本 貴司, 1996. 遺傳アルゴリズム-V -GAの擴張, *システム/制御/情報*, 40(2), pp.69~76.
  24. 玉置 久, 喜多 一, 岩本 貴司, 遺傳アルゴリズム-VI -進化型計算の動向, *システム/制御/情報*, 40(4), pp.170~177.