

# Genetic 프로그래밍을 이용한 방사성폐기물 발생량 예측

이동규 · 김영민 · 오원진

한국원자력연구소

**방**사성 폐기물은 오랜 기간 동안 환경에 심각한 영향을 줄 수 있다.

그러므로 방사성 폐기물은 안전하고 신뢰할 수 있는 방법으로 관리되어야 한다.

원자력발전소에서 발생하는 방사성 폐기물의 양은 환경 친화성, 방사능적 안전성, 폐기물 처분의 경제성 측면에서 매우 중요하다.

원자력발전소에서 발생하는 방사성 폐기물의 양은 발전소 증지 횟수나 작업자의 방사능 피폭 선량 등과 같은 발전소 운전 변수들에 따라 변한다.

방사성 폐기물 발생량에 주로 영향을 미치는 주요 운전 변수들을 선정하여 이들과 발생량 사이의 상관 관계를 분석하였다.

본 논문에서는 방사성 폐기물 발생량 예측을 위한 새로운 방법으로 Genetic Programming(GP) 모델을 제안하였다.

GP는 주어진 문제에 적합한 상관

식을 직접 찾아내는 방법이다.

GP는 가능해 영역 전체를 동시에 탐색하면서 최적해를 찾을 수 있도록 개발된 Genetic Algorithms(GAs)의 원리를 이용하여 주어진 문제의 상관식을 찾아낸다.

기존의 회귀 분석법은 독립 변수들과 종속 변수 사이의 상관 관계를 일정한 수식으로 가정하여야 하기 때문에 비선형 문제를 해결하기가 상당히 어렵다.

그러나 GP는 상관 관계에 대한 수식의 가정 없이 수식 자체를 찾아내기 때문에 비선형 문제에도 쉽게 적용을 할 수 있다.

모형을 도출하였다.

〈표 1〉은 국내 원자력발전소에서 연간 발생하는 방사성 고체 폐기물의 양을 나타내고 있다.

국내 각 발전소에 대한 연간 발생량은 상업 운전 개시 후부터 95년까지의 자료이다.

총 39개의 자료중 29개는 Genetic Programming 모델을 설정하는 입력 자료로 사용하였으며, 나머지 10개는 개발된 GP 모델을 검증하는데 사용하였다.

## 2. 변수 분석

원자력발전소에서 발생하는 방사성 폐기물량을 예측하기 위한 변수로 설계/운전 변수 4가지를 선정하였다.

이들은 방사성 폐기물 발생량에 많은 영향을 미치는 것으로 알려져 있다.

폐기물 발생량 예측에 대한 미국 전력연구소(EPRI)의 연구 결과를 고려하여 발전소 용량, 정지 횟수, 부지 내 작업자의 수, 연간 총작업자 피폭

### 방사성 폐기물 발생량

#### 1. 국내 발전소 자료

본 연구에서는 원자력발전소에서 발생하는 방사성 폐기물의 총량을 예측하기 위해 기존의 발전소 자료를 이용하여 예측 모형을 도출하였다.

우선 고체 폐기물에 대해서만 예측

선량 등 4가지를 주요 변수로 선정하였다.

국내 원자력발전소에 대한 이들의 연간 자료를 <표 2>에 나타내었다.

이들 외에 기술적 요소(즉 감량 방법)에 따라 폐기물 발생량이 상당히 많이 변화할 수 있는데, 국내의 경우 95년까지는 감량 방법이 크게 변하지 않았고, 96년 이후부터 고감량 방법이 사용되었기 때문에 본 연구에서는 기술적 요소를 배제하였다.

그러나 향후에는 이를 고려하여 발생량 예측 모형을 개발해야 할 것이다.

선정된 각 변수들에 따른 폐기물 발생량의 변화를 <그림 1>에 나타내었다.

<그림 1>에 나타난 바와 같이 각 변수들의 영향은 상당한 비선형적 요소를 가지고 있다.

그러므로 각 변수들을 종합적으로 고려하여 폐기물 발생량을 예측하기 위해서는 비선형 해석을 하여야 한다.

**방사성폐기물 발생량 예측을 위한 GP 모델**

**1. GP의 기본 원리**

GP는 주어진 자료에 대해 독립 변수들과 종속 변수 사이의 관계를 거의 정확하게 혹은 최적의 상태로 모사할 수 있는 수식(비록 간단하지는 않지만)을 찾아낸다.

즉 GP는 주어진 자료에 적합한 모델을 도출한다.

기존의 회귀 분석법들은 가정에 의해 설정된 수식의 계수만을 결정하는 반면 GP는 모형에 대한 수식과 계수 둘 다 결정한다.

**가. 표 현**

Genetic Programming에서 수식은 'parse tree'라는 수목의 형태로 표시된다.

Parse tree는 함수·변수·상수 등을

하나의 노드로 하여 각 노드들이 수목의 질차나 함수를 표시한다.

목의 가지와 같이 연결되어 있다. 가지의 맨 끝부분은 변수와 상수로 한 노드에 연결된 가지는 그 노드 구성되어 있다.

(표 1) 국내 NPP에서의 연간 고체 방사성 폐기물 발생량

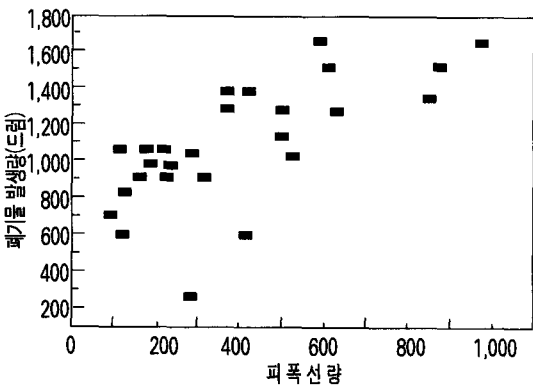
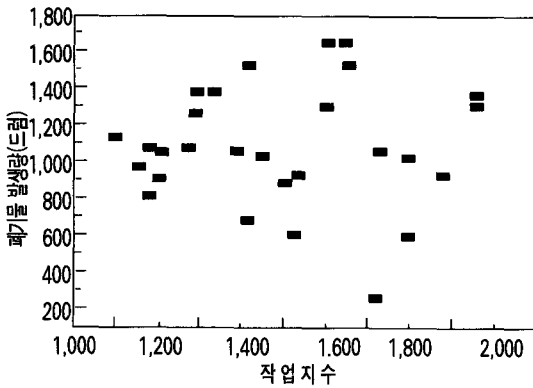
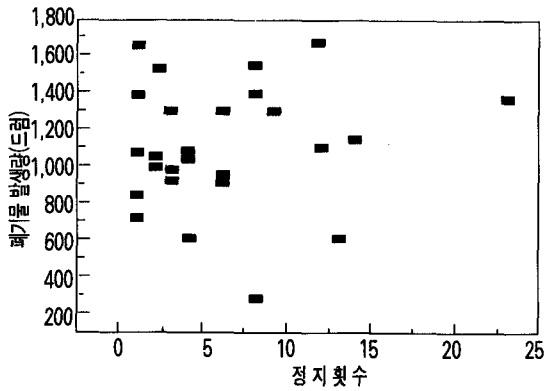
단위 : 드럼

연 도	고리 1&2	고리 3&4	영광 1&2	울진 1&2
83	1,149			
84	1,657			
85	1,363			
86	1,537	615		
87	1,396	1,305	285	
88	1,656	1,037	613	
89	1,288	1,392	921	716
90	1,537	954	1,296	924
91	1,091	1,038	990	1,081
92	1,121	1,078	1,076	1,077
93	1,047	948	888	988
94	497	739	836	1,018
95	1,109	854	883	1,028

(표 2) 국내 원전의 운전 변수들

연 도	고리 1&2			고리 3&4			영광 1&2			울진 1&2		
	Shu	Per	Exp	Shu	Per	Exp	Shu	Per	Exp	Shu	Per	Exp
83	14	1,086	497.5									
84	12	1,571	589.4									
85	23	1,872	854.1									
86	8	1,376	873.6	13	1,727	122.5						
87	8	1,298	405.7	9	1,873	488.8	8	1,659	281.9			
88	1	1,549	974.8	4	1,405	520.8	4	1,472	408.9			
89	6	1,260	616.6	1	1,277	368.8	3	1,460	315.5	1	1,384	100.4
90	2	1,602	603.5	6	1,793	232.0	3	1,553	373.6	6	1,178	162.8
91	12	1,251	180.2	2	1,721	276.4	3	1,265	191.0	4	1,176	115.1
92	5	1,304	429.4	4	1,658	213.7	4	1,360	173.6	1	1,193	179.3
93	3	1,293	454.6	6	1,485	223.1	2	1,398	180.9	2	1,137	239.3
94	2	1,365	316.4	1	1,414	190.5	1	1,158	124.4	1	1,258	180.8
95	3	1,148	138.1	1	1,613	403.1	2	1,507	382.9	2	1,262	137.0

주 : Shu : 정지 횟수  
 Per : 부지 내 작업자의 수  
 Exp : 개인 피폭 선량 합계



〈그림 1〉 폐기물 발생량에 대한 각 변수들의 영향

즉 그들은 수목의 나뭇잎과 같이 맨 끝에 위치한다. 컴퓨터 프로그램에 있어서 모든 수

모든 terminal들의 조합을 'terminal set' 이라 하고, 모든 function들의 조합을 'function set'

식들은 parse tree로 표현될 수 있다.

예를 들어 〈그림 2〉(a)에 나타난 수식은 parse tree에서 〈그림 2〉(b)와 같이 표현된다.

GP에서 사용되는 parse tree의 노드로 사용되는 모든 가능한 함수·변수·상수들을 정의할 수 있다.

변수·상수 그리고 독립 변수들을 가지고 있지 않은 함수 등은 parse tree의 맨 끝에 위치하므로 'terminals' 라고 불린다.

독립 변수들을 가지고 있는 함수들은 parse tree의 가지 부분이 되는데 이를 'functions' 라고 불린다.

이라 한다.

나. Genetic 연산자

GA와 마찬가지로 2개의 주요 연산자들은 'mutation' 과 'crossover' 이다.

그러나 GP는 tree-based 표현이기 때문에, 이 연산자들은 GA에서와는 다르게 작동된다.

Mutation은 말 그대로 parse tree에서 임의의 한 노드를 전혀 다른 노드로 변환시키는 것이다.

Mutation은 다음과 같은 방법으로 새로운 tree를 생성시킨다.

- ① Mutation을 시킬 노드를 tree에서 무작위로(randomly) 선정한다.
- ② 새로운 tree를 생성시킨다.
- ③ 선정된 mutation 노드와 새로 생성된 tree를 교체한다.
- ④ Mutation 시킨 tree의 깊이(depth)가 최대 허용깊이보다 작거나 같으면 새로운 tree를 사용한다.

그러나 만약에 생성된 tree의 깊이가 최대 허용 깊이보다 크면 위의 과정을 다시 반복하여 새로운 tree를 생성한다.

Mutation의 과정을 〈그림 3〉에 나타내었다.

'Crossover' 는 2개의 tree에서 임의의 노드를 서로 바꿈으로써 새로운 tree를 생성하는 방법이다.

Crossover의 적용 절차는 다음과 같다.

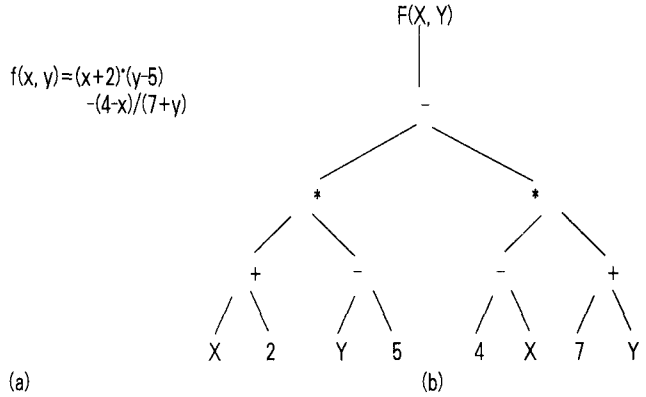
- ① Crossover point로 각 tree(2개의 tree)에서 임의의 노드를 무작위로(randomly) 선정한다.
- ② 각 tree의 crossover point 이하의 subtree를 서로 바꾼다.
- ③ 새로운 tree들의 깊이가 최대 허용 깊이보다 작거나 같으면 사용하고 그렇지 않으면 위의 과정을 다시 반복하여 새로운 tree를 생성한다.

Crossover의 과정을 <그림 4>에 나타내었다.

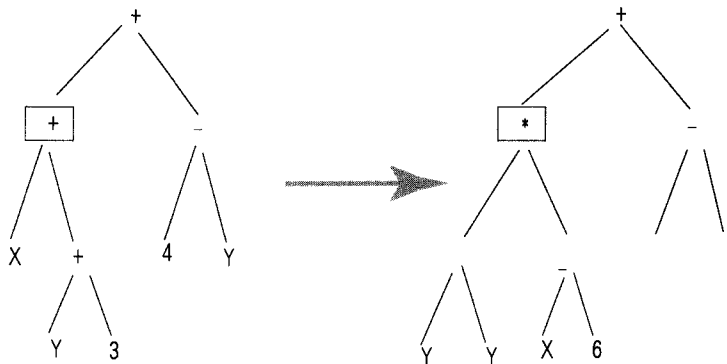
다. 실행 절차

GP는 다음의 실행 절차에 따라 문제 해결을 위한 컴퓨터 프로그램(단순하지는 않지만 수학적 연산자들로 구성된 수식)을 생성한다.

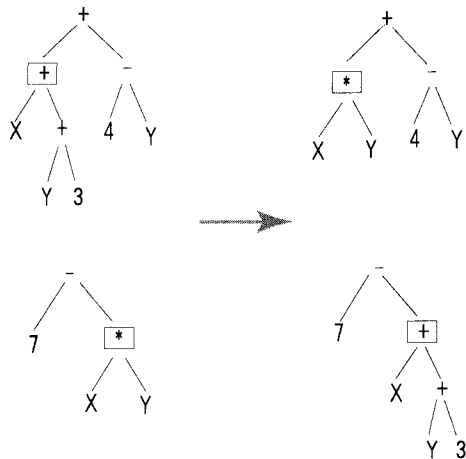
- ① 주어진 'function' 과 'terminal' 들을 임의로 조합하여 초기 population(parse tree)을 생성한다.
- ② 종료 기준에 부합될 때까지 다음의 과정을 반복 실행한다.
- ③ 생성된 tree가 주어진 문제에 얼마나 잘 맞는지 그 정도(fitness)를 알아본다.
- ④ 다음 과정을 거쳐 새로운 population을 생성한다. 실행은 문제 해결 능력(fitness)에 따라 선정된 population에 대해 이루어진다.
- ⑤ 기존의 parse tree들을 새로운 population에 복사한다.



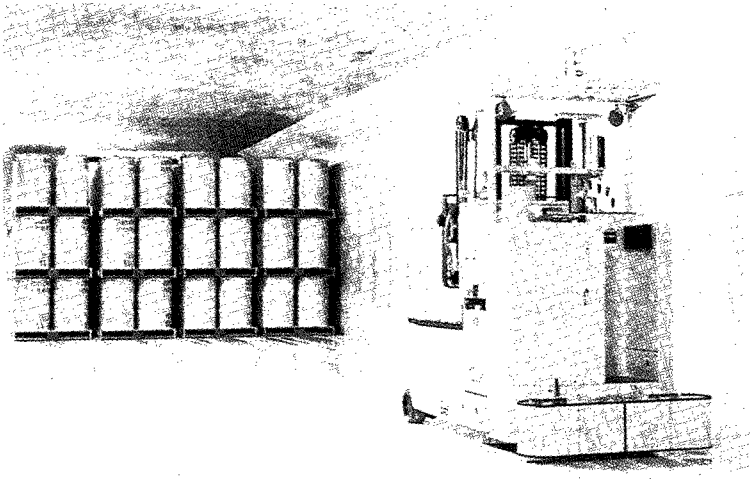
<그림 2> A Subroutine and an Equivalent Parse Tree



<그림 3> Mutation for Genetic Programming



<그림 4> Crossover for Genetic Programming



중·저준위 방사성 폐기물 저장고

① genetic 연산자들을 이용하여 새로운 tree를 생성한다.

③ 여러 번의 반복 과정을 거치면서 주어진 문제에 가장 적합한 컴퓨터 프로그램(tree)을 GP의 결과로 선정한다.

## 2. GP 모델의 구성

Function set은 +, -, \*, %와 같은 수학적 연산자들(입의의 함수를 포함한 모든 수학함수 포함)로 구성된다.

+, -, \* 연산자는 각각 일반적인 더하기, 빼기, 곱하기를 표시하며, %는 'protected division'으로 0으로 나눌 경우 에러 발생을 방지하기 위한 나누기이다(즉 분모가 0이면 0의 값을 준다).

폐기물 발생량 예측을 위해 사용된

function set은 다음과 같다.

$$F = \{+, -, *, \%\}$$

Terminal set은 앞서 선정된 인자들, 즉 원자력발전소에서 폐기물 발생량에 영향을 미치는 설계 및 운전 변수들로 구성되었다.

사용된 terminal set은 다음과 같다.

$$T = \{\text{Cap, Shu, Per, Exp}\}$$

여기서 Cap는 발전소 용량(MWe), Shu는 발전소 정지 횟수, Per은 부지 내 작업 종사자들의 수, 그리고 Exp는 개인 피폭 선량 합계를 의미한다.

### 결과 및 고찰

Genetic Programming은 무작위(random) 탐색이기 때문에 여러 번

을 독립적으로 시행해야 좋은 결과를 얻을 수 있다.

폐기물 발생량을 예측할 수 있는 적절한 모델을 선정하기 위해서 주어진 function set과 terminal set을 가지고 15회의 독립 시행을 실시하였다.

본 연구에서는 GP 모델의 성능을 평가하기 위해서 다음의 퍼센트 오차를 사용하였다.

$$\% \text{ 오차} = \frac{|\text{예측량} - \text{실제량}|}{\text{실제량}}$$

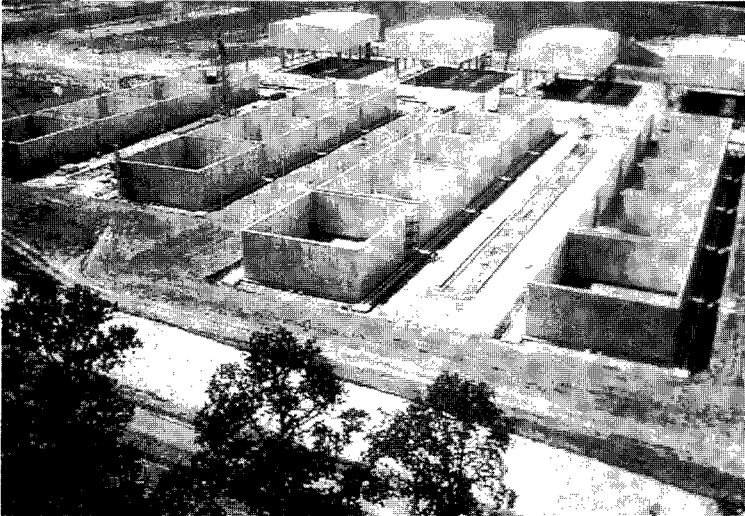
15번의 독립 시행 가운데 가장 잘 맞는 GP 모델의 결과를 <표 3>에 나타내었다.

검증 기간 동안 평균 % 오차는 17.87%이었고, 이때 표준 편차는 18.67%이었다.

그러나 다른 자료에 비해 패턴이 많이 틀린 자료 1개를 제외하면 평균 오차는 12.55%이고, 표준 편차는 8.61%로 비교적 잘 예측됨을 알 수 있다.

<표 3> GP 모델의 결과

실제 발생량	예측된 발생량	% 오차
1,121	927.83075	17.23187
1,047	928.56708	11.31164
497	502.53876	1.11444
739	858.99890	16.23801
854	1062.63892	24.43079
888	848.23169	4.47841
883	1088.03198	23.21993
1,018	994.85925	2.27316
1,028	897.63684	12.68124
1,109	380.09564	65.72627



프랑스 L'Aube 방사성 폐기물 처분 시설

(그림 5)는 GP 모델을 사용하여 예측한 폐기물 발생량과 실제 폐기물 발생량을 비교하여 나타내었다.

그러나 수식이 매우 복잡하여 각 변수들이 폐기물 발생량에 미치는 영

향의 정도를 알아보는 매우 어렵다. 앞으로 이들의 영향력을 알아볼 수 있는 물리적으로 의미가 있는 상관관계식을 찾기 위해서는 최대 tree 깊이를 줄여서 GP를 시행해야 할 것이다.

최대 깊이를 줄일 경우 예측의 정확도는 떨어질 수 있으나 폐기물의 경우 발생량을 정확하게 예측하는 것 못지 않게 발생량을 줄이는 방안도 매우 중요하다.

폐기물 발생량을 효과적으로 줄이기 위해서는 발생량에 가장 많은 영향을 미치는 인자를 찾아서 그 인자를 변경시켜야 한다.

그러기 위해서는 발생량과 변수들과의 상관 관계식에서 영향력이 가장 큰 인자를 찾아야 하므로 상관 관계식이 해석을 할 수 있도록 복잡하지 않아야 한다.

예측도 잘하고 상관 관계도 해석할 수 있는 GP 모델을 찾기 위해서는 향후 더 많은 연구가 수행되어야 한다.

### 결론

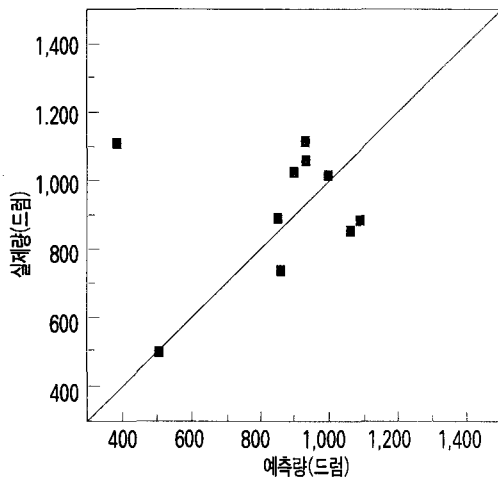
Genetic Programming 방법은 원자력발전소에서 발생하는 방사성 폐기물의 양을 예측하는 도구로 유용하게 이용될 수 있다.

본 논문에서 제시된 시험 결과를 바탕으로 GP 모델은 원자력발전소의 방사성 폐기물 발생량 예측을 위해 사용될 수 있다고 결론지을 수 있다.

GP 모델은 독립 변수들과 종속 변수 사이의 상관 관계식을 생성하기 때문에 문제 해결을 위한 수식적 가정이 필요 없으며, 결과를 다른 곳에 이용하기가 매우 용이하다.

향후 보다 나은 GP 모델을 개발하기 위해서는 기술적 요소를 포함하는 Terminal Set과 여러 가지 Function Set들을 고려해야 할 것이다.

특히 발전소별 특성을 고려하는 방안을 모색하여야 하며, 정지 횟수를 일반 정지와 불시 정지 등 폐기물 발생량이 차이가 나는 정지를 구분하여 고려하는 것이 바람직할 것으로 생각된다. ☻



(그림 5) 실제 발생량과 예측된 발생량과의 비교