

## 개선된 유전적 프로그래밍 기법을 이용한 선박 입출항 의사결정 지원 시스템

이경호

한국해양연구원 해양시스템안전연구소  
(whlee@kriso.re.kr)

연운석

대진대학교 기계설계공학과  
(yysuh@sol.nuri.net)

이 옥

대진대학교 전기공학과  
(theewook@road.daejin.ac.kr)

본 연구에서 대상으로 하고 있는 LG정유 광양항 제2부두는 7선석(Berth)에 재화중량(DWT) 300톤에서 48000톤의 선박까지 다양한 선박이 이용하고 있으며, 해상의 기상상태에 따른 선박 입출항 통제 지침 설정이 어렵고, 현재 사용하고 있는 지침의 근거가 명확하지 않아 현재의 부두 운영이 비효율적이거나 안전성이 결여되어 있다고 할 수 있다. 따라서 이를 개선하기 위한 합리적인 부두운영 제한조건 개발이 절실히 요구되었다.

본 논문에서는 대상 부두의 특성, 대상 선박의 특성, 하중상태, 선박 운항자의 특성 등을 고려하여 해상/기상 상황(바람, 조류 및 파랑)에 따른 부두 입출항 가능 여부를 정량적으로 판단하고, 안전성 향상 방안을 제시할 수 있는 의사결정 시스템을 개발하고 5번, 7번 선석을 대상으로 이를 검증하였다.

여기서는 입출항 여부를 정량적으로 판단하여 결과를 제시하기 위해서 유전적 프로그래밍(Genetic Programming)을 이용한 기계학습 방법을 이용하였으며, GP의 방대한 계산량을 줄이기 위한 가중 선형 연상 기억(Weighted Linear Associative Memory:WLAM) 방법의 도입 및 전역 최적점을 쉽게 찾기 위한 Group of Additive Genetic Programming Trees(GAGPT)를 도입함으로써 학습 성능을 개선하였다.

### 1. 서론

최근 들어 유조선 등의 선박 사고로 인한 해양 오염이 심각한 사회문제로 대두되고 있고, 이러한 유조선의 사고는 수년 전 남해안의 씨프린스호 사고에서도 보았듯이 단순한 해양 오염의 차원을 넘어서 재앙이라고 할만큼 심각한 피해를 초래하고 있다. 유조선이 하루에도 수 십 척씩 드나드는 정유회사의 부두는 지난 20여년 동안

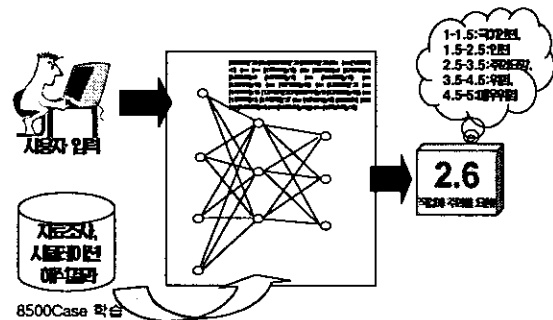
운항자의 부주의, 부두 운영 제한 조건 미비 등으로 인한 부두 접촉사고 및 선박간의 접촉사고의 사례가 많은 것으로 알려져 있다(LG-Caltex 정유 해상출하팀, 2000). 현재 LG정유 제2부두는 7선석(Berth)에 재화중량 (Deadweight : DWT) 300톤 ~ 50,000톤의 선박까지 다양한 선박들이 이용하고 있고, 각 부두의 해역적 특성 및 해상/기상 상황이 고려된 명확한 입출항 통제 기준이 마련되지 않아, 안전한 경우에도 선박의

입출항 및 이접안이 통제되거나 혹은 반대로 매우 위험한 경우에도 입출항 및 이접안 작업을 강행하는 경우가 발생하고 있어 사고 위험요소가 많이 존재하고 있다.

따라서 부두의 특성과 입출항 당시의 해상/기상 상황과 해역 특성이 고려된 명확하고 합리적인 부두 운영 제한 조건의 도출이 절실히 요망된다.

이러한 요구에 부응하여 본 연구에서는 LG정유 제품부두의 지형적/해역적 특성과 선박의 입출항 및 이접안 당시의 해상/기상 상황, 그리고 입출항 선박의 특성 등을 종합적으로 고려하여, 입출항 및 이접안 가능여부를 정량적으로 제시하여 주는 의사결정 시스템을 개발하여 LG정유 제품부두 운영 관리자가 손쉽게 이용할 수 있도록 지원함으로써 제품부두 입출항 작업의 안전성과 효율성을 향상시키는데 그 목적이 있다.

연속 데이터 학습에 뛰어난 성능을 나타내는 유전적 프로그래밍(Genetic Programming : 이하 GP) 기법을 이용하였다. 학습 시스템의 도입은 정량적인 기준을 제시할 수 있다는 것 외에 지속적인 입출항 데이터의 축적을 통한 시스템의 신뢰성을 확보할 수 있다는 긍정적인 측면을 갖고 있다. 본 의사결정 시스템의 개념은 <그림 1>과 같이 매우 단순하다.



<그림 1> 입출항 의사결정 시스템 구성 개념도

## 2. 시스템 구성

본 연구에서 수행한 입출항 제한 의사결정 지원 시스템은 부두와 선박, 그 외의 해상/기상상태에 따른 선박의 입출항 및 이접안 가능 여부 파악 및 통제 기준안을 마련하기 위한 것으로서 입출항 가능여부를 정량적으로 제시한다. 그러나 이러한 의사결정 시스템의 개발은 그 특성 상 현장에서의 의사결정에 있어 체계화된 규칙이 없고 정형화하기 매우 어려우며, 수치화된 결과를 요구하고 있기 때문에 지식기반 시스템의 도입보다는 현장의 입출항 데이터를 수집하고 이들 데이터의 학습(Learning)을 통하여 새로운 입출항 경우의 위험도(난이도)를 찾아내는 학습기반 시스템을 도입하게 되었으며, 이를 위해서 비선형, 불

그러나 고려해야 할 입출항 위험도(난이도)에 영향을 미칠 수 있는 인자들이 많이 있고, 이에 비해 현재까지 체계적으로 축적된 입출항 데이터가 전무한 상태이어서 시스템 개발에 많은 어려움이 있다.

위험도에 영향을 미칠 수 있는 인자로는 선박의 재화중량(DWT), 입출항 구분(입항, 출항), 하중상태(Ballast:0~Full:100), 부두 접안 자세(우현 접안, 좌현접안), 조류(약창조, 중창조, 강창조, 약낙조, 중낙조, 강낙조, 저조정조, 고조정조), 풍속(0~35 노트), 풍향, 예선(Tug boat) 용량 등이 있다. 이 외에도 정량화 하기 힘든 입력인자, 즉 날씨 및 시정, 도선사(Pilot)의 자질 등에 따른 위험도(난이도)는 나름대로의 규칙을 정하여 조정

하였다.

본 시스템의 출력으로는 주어진 입출항 상황에서 난이도를 정량적으로 표시하고, 이 값의 범위를 다음과 같이 5단계로 나누어 입출항 관리자에게 현재의 상황을 표시해 준다.

난이도

- 1.0 ~ 1.5 : 극히 안전
- 1.5 ~ 2.5 : 안전
- 2.5 ~ 3.5 : 주의 요망
- 3.5 ~ 4.5 : 위험
- 4.5 ~ 5.0 : 매우 위험 (입출항 불가)

학습에 이용할 데이터의 추출을 위해서 본 논문에서는 수 십년 동안 본 연구소에서 개발되어 그 성능이 검증된 선박 운항 시뮬레이션 프로그램을 이용하여, 주어진 입력 인자의 값들을 변화시키면서 Fast Time Simulation(FTS)을 수행하였다. 본 연구에서는 5번, 7번 부두를 대상으로 하였으며, 이들 부두의 입력인자를 변화시키면서 5번 부두 3968개, 7번 부두 6144개의 입출항 사례를 생성하였다. FTS를 통해 학습을 위한 데이터 셋을 얻었다 하더라도 이들 각 경우에 대한 위험도(난이도)를 부여하여야 한다. 이를 위하여 과제가 진행되는 동안 부두에서 추출한 입출항 자료를 바탕으로 각 사례의 난이도를 추정하였다.

이렇게 생성된 학습 데이터를 바탕으로 유전적 프로그래밍에 의한 학습 시스템을 구성하였다.

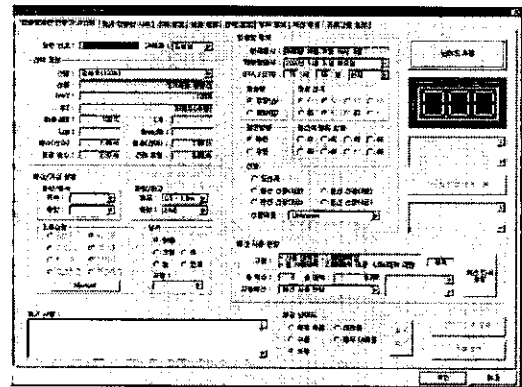
<그림 2>는 구현된 의사결정 시스템의 메인 화면이다. 여기서 입출항 선박에 대한 정보는 이 제품부두에 입출항 하는 모든 선박의 정보를 데이터 베이스화 하여 선박에 대한 제원을 쉽게 알 수 있고, 이를 입출항 난이도를 추정하는데 이용

할 수 있다.

해상/기상 데이터는 부두에 설치된 풍향/풍속 계와의 인터페이스를 통해 실시간으로 정보를 수집한다.

조류 데이터는 제품부두 주위의 조류상태를 측정하여 이를 데이터 베이스화 하였으며, 주어진 날짜에 따른 조류상태가 자동으로 입력될 수 있다.

이러한 데이터를 바탕으로 본 의사결정 시스템은 학습된 데이터를 바탕으로 평가 난이도를 제시하게 되며, 시스템의 추가 학습을 통한 신뢰성 향상을 위하여 부두 운영 관리자가 느낀 체감 난이도를 매번 적어 넣도록 하여 이를 데이터베이스에 저장하고, 추가 학습에 사용하도록 하였다.



<그림 2> 선박 입출항 의사결정 지원 시스템 구현 화면

### 3. 개선된 유전적 프로그래밍

유전적 프로그래밍은 유전적 알고리즘의 확장으로써 그 개체(Individual)가 트리(Tree) 형태의 컴퓨터 프로그램이 된다(이경호 외, 1998). 본 논문

문에서 컴퓨터 프로그램은 터미널 집합(Terminal set)과 함수 집합(Function set)의 조합으로 생성된 문법적으로 올바른 GP tree를 뜻한다. 진화과정(Evolving process)을 통하여 GP tree는 적합도(Fitness)를 최적화하기 위해서 그 구조 자체가 동적으로 변화하는데, 적합도 계산을 위해서 트리의 학습오차(Learning error)를 계산할 수 있는 함수가 사용된다. 기저함수 바탕의 근사화 기법은 그 함수의 형태가 이미 결정되어 있는 반면, 유전적 프로그래밍에서는 함수 즉 GP tree의 구조 자체가 적합도를 최적화하기 위하여 변화된다. 이러한 특징을 고려할 때, 유전적 프로그래밍은 함수 근사화의 유용한 도구로 활용될 가능성이 크다. 그러나 이것을 위해서는 다음과 같은 두 가지 근본적인 문제의 해결이 선행되어야 한다.

첫째, 유전적 프로그래밍에서는 다양한 공학적 응용이나, 함수 근사화에 필수적인 사항인 수치적 파라미터(Parameter) 혹은 가중치(Weight)를 추정하기 위한 방법이 결여되어 있다. 최근에 가중치 추정을 위한 많은 연구들이 진행되고 있는데, 이 모든 연구들은 비선형 최적화 기법에 의존하고 있다. 이러한 접근 방법의 명백한 단점은 방대한 계산량을 요구한다는 점이다. 경우에 따라서는 한 개체군(Population) 내의 모든 개체에 대한 적합도를 계산하기 위해서 몇 대의 워크스테이션을 병렬로 사용해도 하루 이상의 시간이 소요될 수 있다. 따라서 본 연구에서는 비선형 최적화 기법의 사용을 지양하고, 가중 선형 연상 기억(Weighted linear associative memory, 이하 WLAM)의 사용을 모색하였다(Yeun et al. 2001; Lin and Durand, 1996). 만일 연상행렬(Association matrix)의 크기가 커지면, 이 또한 많은 계산량이 요구되므로, 몇 개의 크기가 작은 연상

행렬들로 구성된 하나의 집합을 구성하여 사용하는 방안을 고안하였다. 이 연상 행렬은 GP tree의 전체 가중치 및 학습 집합의 일부만을 사용하여 구축된다. 이 접근 방식의 장점은 계산량의 대폭적인 감축과 아울러 비교적 정확한 가중치의 값을 산정할 수 있다는 점이다.

둘째, 유전적 프로그래밍의 알고리즘은 종종 지역최적점(Local minimum point)에서 벗어나지 못하기 때문에 만족할 만한 성능을 갖는 GP tree의 생성이 어려운 경우가 있다. 유전적 알고리즘은 전역적 최적화의 기법으로 매우 많은 영역에서 활용되고 있는데, 지역최적점을 탈출하기 위하여 돌연변이 연산자(Mutation operator)가 사용된다. 그 반면 유전적 프로그래밍도 유사한 돌연변이 연산자가 존재하지만 일반적으로 유전적 알고리즘과는 달리 그 유용성이 크지 않다고 인식되고 있다(Koza, 1992; 1994). 이런 문제점을 해결하기 위해서, 본 연구에서는 Group of additive genetic programming trees(GAGPT)를 제시하였다. GAGPT는 하나의 주 트리(Primary tree)와 보 트리(Auxiliary tree)들의 집합으로 구성된다. 주 트리는 학습 집합을 사용하여 입력-출력 관계를 근사 시키는 일반적인 GP tree이다. 이 주 트리를 생성한 후 보 트리들이 첨가되는데, 보 트리는 원하는 출력과 주 트리 및 GAGPT에 이미 존재하는 보 트리들의 모든 출력을 합친 값에 기초한 적합도 함수를 최적화하는 방향으로 진화가 진행된다. 따라서 보 트리들의 첨가들 통하여 GAGPT에 대한 학습성능(Learning capability) 및 일반화 성능(Generalization capability)의 향상을 꾀할 수 있다(연운석, 1998).

### 3.1 가중 선형 연상기억

최근에 선형연상기억(Linear associative memory)을(Kohonen, 1988) 시스템 파라미터 식별(System parameter identification)에 적용하려는 연구가 수행되었는데(Kalaba et al. 1992; Kalaba et al. 1993), 만일 시스템 파라미터를 GP tree의 가중치와 대응하여 생각한다면 선형 연상기억은 GP tree의 가중치 산정에 충분히 사용될 수 있음을 짐작할 수 있다.

선형 연상기억을 이용하여 GP tree의 가중치를 산정하는 과정을 요약하면 다음과 같다.

학습집합  $L((X_1, y_1), (X_2, y_2), \dots, (X_p, y_p))$ 과 GP tree의 노드에 대하여 서로 다른  $m$ 개의 가중치 벡터들의 집합  $W(w_1, w_2, \dots, w_m)$ 가 주어질 경우, 이들을 사용하여 GP tree의 출력 벡터들의 집합  $Z(z_1, z_2, \dots, z_m)$ 를 구성할 수 있다. 여기서  $X_i$ 는 입력 벡터이고,  $y_i$ 는 GP tree에 요구되는 출력 값이다. 또한  $W_i$ 와  $Z_i$ 는 각각  $n$  차원,  $p$  차원의 벡터이다.  $Z_i$ 는 입력 벡터들의 집합  $X(X_1, X_2, \dots, X_p)$ 에 대하여 가중치 벡터가  $w_i$ 일 때 GP tree의 출력 값들로 구성된 벡터이다. 이 가중치 벡터와 출력 벡터를 사용하여 행렬  $W(n \times m)$ 와  $Z(p \times m)$ 를 구성할 수 있고, 식(1)을 사용하여 연상기억 행렬(Association matrix)  $M$ 을 구한다.

$$W=MZ \quad (1)$$

여기서  $M(n \times p)$ 은 Singular value decomposition 기법을 사용하여 구할 수 있다(Gruder, 1990). 이때, 구하고자 하는 가중치 벡터  $w^*$ 는 식(2)에 의하여 계산할 수 있다.

$$w^*=My \quad (2)$$

여기서  $y$ 는 학습집합  $L$ 에 포함된 요구되는 출력 벡터  $y(y_1, y_2, \dots, y_p)$ 이다.

$w^*$ 와  $y$ 의 실제 관계는 비선형인데, 식(2)에서는 이 관계를 근사적으로 선형화 한 것으로 볼 수 있다. 따라서 정확한 가중치의 산정이 어렵다고 판단할 수 있지만, 선형 연상기억을 사용하여 매우 정확한 시스템 파라미터를 계산하는 것이 가능하다는 연구 결과가 보고되어 있다(Kalaba et al. 1992; Kalaba et al. 1993).

보다 정확한 가중치의 추정을 도모하기 위해서, 가중함수(Weight function)를 도입하여  $M$ 을 변경시키고 이로부터 다시 가중치를 추정하는 방식이 있는데, 이러한 경우가 WLAM에 해당한다. 본 논문에서는 WLAM에 관련된 사항 중 GP 알고리즘에 사용될 것만 간략히 기술하고자 하며, 자세한 내용은 참고문헌(Lin and Durand, 1996)을 참조하기 바란다.

$M(n \times p)$ 의 크기가 가중치 추정시 필요한 계산량에 가장 중요한 영향을 미친다. 만일 GP tree의 노드 개수( $n$ )가 많고, 학습집합의 양( $p$ )이 많다면  $M$ 의 크기가 증가하고 이에 따라서  $M$ 을 구축하는데 필요한 계산량이 대폭 증가한다. 따라서 이러한 문제점을 해결하기 위한 방안으로 크기가 작은 몇 개의  $M$ 을 구축하여 크기가 큰 원래의  $M$ 대신에 사용하는 방법을 다음과 같이 모색하였다.

첫째, GP tree의 모든 가중치를 한꺼번에 처리하는 대신, 그 일부만을 선정하여 값을 추정하고, 모든 가중치의 값이 추정될 때까지 이 과정을 반복한다.

둘째, 전체 학습집합을 몇 개의 작은 집합으로 나누어 사용한다. 본 연구 수행의 경험에 의하면 보통 10-20개 정도의 원소들을 갖는 집합들로 구성하는 것이 적절하다고 판단된다.

위의 방안을 기초로 유전적 프로그래밍 알고리즘 하에서 WLAM을 이용한 가중치 산정 과정을 기술하면 다음과 같다.

- 1) 터미널 집합과 함수 집합을 정의하고, WLAM에 필요한 상수들( $D_0, \omega$ )의 초기 값을 설정한다.
- 2) GP tree를 무작위로 생성하여 초기 개체군을 생성한다. GP tree의 모든 가중치는 1로 초기화한다.
- 3) 개체군 내의 모든 GP tree에 대하여 다음 작업을 수행한다.
  - i. 입력 벡터들  $X(X_1, X_2, \dots, X_p)$ 를 사용하여 GP tree의 출력을 계산하고, 이로부터 출력 벡터  $Z$ 를 만든다. 이때,  $Z$ 와 실제 출력 벡터 사이의 학습오류(여기서는 Squared-error)를 계산한다.
  - ii. GP tree의 전체 가중치 중 그 일부  $\bar{\omega}$ 를 무작위 또는 정해진 방법에 따라서 선정하여, 가중치 벡터  $\bar{\omega}(\bar{\omega}_1, \bar{\omega}_2, \dots, \bar{\omega}_d)$ 를 구성한다. 여기서  $d$ 는  $n$ 보다 작거나 같아야 한다.
  - iii. 학습집합  $L$ 에서 일부의 데이터를 무작위 또는 정해진 방법에 따라서 추출하여, 크기가 작은 학습집합  $L((\bar{X}_1, \bar{y}_1), (\bar{X}_2, \bar{y}_2), \dots, (\bar{X}_g, \bar{y}_g))$ 을 구성한다. 여기서  $g$ 는  $p$ 보다 작거나 같아야 한다.
  - iv.  $\bar{\omega}$ 에 난수를 더하는 방법을 사용하여,  $h$ 개의 서로 다른 가중치 벡터를 구하고 이로부터 가중치 행렬  $\bar{W}(dxh)$ 를 구성한다.  $\bar{W}$ 에 대응하는 GP tree의 출력 행렬  $\bar{Z}(gxh)$ 를 입력 벡터  $\bar{X}(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_g)$ 를 사용하여 구한다.  $\bar{W}$ 와  $\bar{Z}$ 를 사용하여, 연상기억 행렬  $\bar{M}(dxg)$ 를 구한다.

- v. 추정된 가중치 벡터  $\bar{\omega}$ 를 계산한다. 이때, GP tree의  $\bar{W}$  대신에  $\bar{\omega}$ 를 사용하여,  $i$  과정에서와 같이 GP tree의 학습오류를 계산한다. 만일 계산된 학습오류가 전의 것 보다 작으면  $\bar{\omega}$ 를  $\omega'$ 에 복사한다.
  - vi.  $\bar{\omega}$ 에 대한 GP tree의 새로운 출력 벡터  $\bar{z}$ 를  $\bar{X}$ 를 사용하여 구한다.  $\bar{M}$ ,  $\bar{\omega}$  그리고  $\bar{z}$ 를 사용하여 새로운 연상기억 행렬  $\bar{M}$ 을 계산한다
  - vii. v-vi과정을 원하는 정밀도의  $\bar{\omega}$ 가 얻어지거나 최대 반복 횟수에 도달할 때까지 반복한다.  $\omega'$ 를  $\bar{\omega}$ 에 복사한다.
  - viii. iii-vii과정을 학습집합의 모든 데이터가 가중치 추정에 사용 때까지 반복한다.
  - ix. ii-viii과정을 모든 가중치가 추정될 때까지 반복한다.
- 4) 유전적 연산자들을 개체군에 적용하여 다음 세대의 개체군을 생성한다.
- 추정된 가중치는 지역 최적해일 가능성이 매우 크지만, GP tree 구조가 자체가 그 학습능력을 최적화하기 위해서 동적으로 변하기 때문에 추정된 가중치가 전역 최적해가 되어야 할 필요는 없다. 더욱이, 전역 최적해로 찾아진 가중치를 소유한 GP tree의 거동은 종종 Overfitting 현상을 보일 가능성이 높다. GP tree의 가중치 산정시 최적화 기법을 적용할 경우와 WLAM을 사용한 경우 차이점을 보이게 된다. 그러나 세대가 지나감에 따라서 그 성능에 있어서 큰 차이점을 보이지 않게 됨을 관찰할 수 있는데, 이것은 유전적 연산자들에 의해서 적합도 함수를 최적화하기 위해서 트리의 구조 자체가 변화되기 때문이라고 판단된다. 이때, 각기 다른 방법에

의거하여 추정된 가중치를 갖는 최적 트리들이 비록 동일한 성능을 보이는 경우라도 이들의 구조와 크기는 서로 차이가 있게 되는 것이 일반적인 현상이다.

### 3.2 Group of Additive Genetic Programming Trees (GAGPT)

유전적 프로그래밍 알고리즘이 지역 최적점에 도달할 때, 이 지역을 탈출하지 못하는 경우를 자주 관찰 할 수 있는데, 이런 경우에 많은 세대가 지나더라도 좋은 성능을 갖는 GP tree가 생성되기 어렵다. 따라서, 본 논문에서는 이런 문제를 극복하기 위하여 GAGPT를 도입하였다.

GAGPT의 기본 개념은 주 트리(Tp)에 첨가될 보 트리(Ta)들을 도입한 것인데, 식(3)에 기술되어 있다.

$$GAGPT = T_p + \sum_{i=1}^q T_a^i \quad (3)$$

여기서, q는 보 트리의 개수이고,  $T_a^i$ 는 i번째 보 트리를 의미한다.

즉 GAGPT의 출력은 주 트리와 모든 보 트리들의 출력을 누적함으로써 얻어진다.  $T_p$ 와  $T_a^k$ 의 적합도 함수는 식(4)와 (5)에 나타나 있다.

$$f_{T_p} = \sum_{i=1}^p |y_i - z(X_i)|^2 \quad (4)$$

$$f_{T_a^k} = \sum_{i=1}^p \left| y_i - z(X_i) - \sum_{j=1}^k z_a^j(X_i) \right|^2 \quad (5)$$

여기서,  $z_a^i(X_i)$ 는 입력이  $X_i$ 일 때, i번째 보 트리의 출력이다.

GAGPT를 구축하는 과정을 요약하면 다음과 같다.

1)  $f_{T_p}$ 를 적합도 함수로 사용하여 유전적 프로그래밍의 알고리즘을 통하여 최적의 주 트리  $T_p$ 를 먼저 생성한다.

2)  $f_{T_a^k}^k$ 를 적합도 함수로 사용하여 최적의 GP tree를 생성하는데, 이것이 보 트리  $T_a^k$ 가 된다.

3) 만일 GAGPT의 학습 오류를 좀더 줄이기 위해서는 또 다른 보 트리를 2)과정을 통하여 생성한다. 이 과정을 원하는 GAGPT의 성능이 얻어질 때까지 또는 어떤 종료 조건이 만족될 때까지 반복한다.

보 트리의 개수가 증가함에 따라서 GAGPT의 학습 성능은 꾸준히 개선되지만 보다 중요한 일반화 성능은 다른 경향을 보인다. 즉 어떤 특정한 보 트리의 숫자에서 최대의 일반화 성능을 보인 후, 또 다른 보 트리를 추가함에 따라서 일반화 성능은 개선되지 않거나 저하되는 현상을 관찰할 수 있다. 보 트리의 증가에 따른 GAGPT의 성능은 GP tree의 크기, 문제 영역의 특성에 크게 지배를 받는다. 그러나 무엇보다도 중요한 요인은 보 트리의 적합도 함수의 특징에 기인한다. 식(6)에서 알 수 있듯이 보 트리의 적합도 함수는 주 트리와 GAGPT에 포함된 모든 보 트리 해의 합과 요구되는 출력 값과의 차이에 의존하고 있다.

$$\Delta_i = y_i - \left\{ z(X_i) + \sum_{j=1}^{k-1} z_a^j(X_i) \right\} \quad (6)$$

현재의 보 트리는  $\Delta_i$ 를 기반으로 한 적합도 함수를 최적화하려는 방향으로 진화한다. 이때,  $\Delta_i$ 가 매우 복잡한 패턴을 보여 준다면 좋은 성능을 소유한 보 트리를 생성하기 어렵고, 따라서 보

트리의 추가에 따른 GAGPT의 성능 향상을 기대하기 어렵다.

#### 4. 유전적 프로그래밍을 이용한 입출향 난이도 추정

앞서 2절에서 언급하였듯이 선박의 입출항 의사결정 지원 시스템을 구성하기 위하여 먼저 학습에 이용될 사례들을 FTS 기법을 이용하여 추출하였고, 이렇게 추출된 데이터를 바탕으로 3절에서 구현된 GP의 학습을 통해 난이도를 추정하였다.

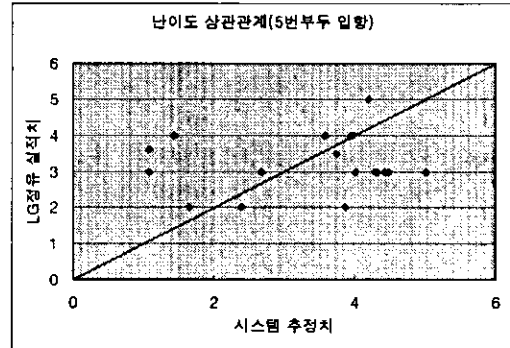
제품부두 현지에서 수집된 부두 관리자가 작성한 데이터를 바탕으로 각각의 상황을 구현된 시스템에 그대로 입력하여 시스템이 추정하는 난이도와 비교를 수행하였다.

<그림 3> ~ <그림 6>은 5번 부두 입항과 출항, 7번 부두 입항과 출항에 대해 그 결과를 각각 도식화하였다.

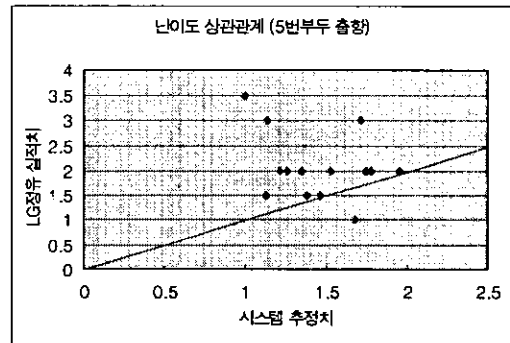
참고로, 각각의 점들이 대각선 상에 놓이게 되면 가장 이상적인 경우라고 할 수 있다.

그림에서 볼 수 있듯이 5번 부두에 대한 결과는 대각선을 중심으로 비교적 고르게 분포되어 있으며, 많이 편중된 점들은 실제 정량화하기 어려워 입력 변수에서 제외한 인자의 값이 비정상적으로 크거나 작아서 부두 운영자가 느끼는 체감 난이도가 다르게 느껴지는 것으로 판단되며, 7번 부두의 경우는 현장에서 수집된 데이터가 매우 적어서 이를 바탕으로 한 FTS에서 분석된 데이터의 난이도 근사 과정이 문제가 있는 것으로 보여진다. 이것은 향후 현지 데이터의 수집이 충분히 확보하여 이를 바탕으로 학습 데이터의 난이도를 근사하게 되면 이 문제가 해결될 수 있을

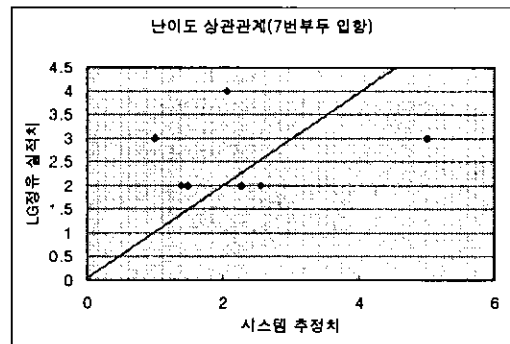
것으로 예상된다.



<그림 3> 5번 부두 입항에 대한 난이도 상관관계

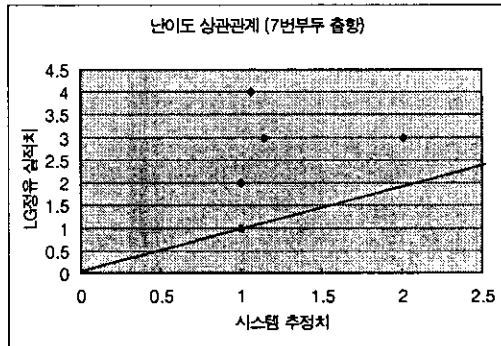


<그림 4> 5번 부두 출항에 대한 난이도 상관관계



<그림 5> 7번 부두 입항에 대한 난이도 상관관계





<그림 6> 7번 부두 출항에 대한 난이도 상관관계

## 5. 결론 및 향후계획

본 연구에서는 선박의 입출항 의사결정 시스템을 구현하기 위하여 GP를 통한 학습 기법을 도입하였다.

우선, GP의 성능을 개선하는데 있어서 방대한 계산량을 줄이기 위한 가중 선형 연상기억(WLAM) 방법을 도입하였고, GAGPT의 도입을 통해 GP의 학습 능력과 일반화 성능을 향상시킬 수 있는 학습 시스템을 구현하였다.

이렇게 구현된 학습 시스템을 학습시키는데 이용될 입출항 데이터를 추출하기 위하여 FTS(Fast Time Simulation)을 수행하였으며, 시스템의 검증에 위하여 제품부두 현장에서 수집된 실제 입출항 사례를 구현된 시스템에 입력하여 그 결과를 비교 분석하였다.

현재의 학습시 이용되는 입력 파라미터들이 제한적이며, 이들의 영향도에 대해서 정확히 아는 것이 불가능하기 때문에 실제 현장에서 축적되는 데이터를 바탕으로 추후에 dominant한 파라미터들을 찾아내고, 이들을 바탕으로 재학습을

수행할 것이다.

또한 5번, 7번 부두에 대한 추가 학습을 추진할 계획이며, 이를 위하여 5번 및 7번 부두에 입출항하는 다양한 크기의 선박에 대한 실제 운영 자료가 더 필요하다.

현재 현장에서 1차 버전을 이용하여 수행되는 모든 경우의 데이터를 DB에 축적하여 적당한 시점에 그 데이터를 확보하고 이를 바탕으로 지속적인 시스템 Upgrade를 수행할 방침이다.

## Acknowledgement

본 논문은 LG정유㈜의 지원을 받아 1년간 당 연구소에서 수행한 "부두 입출항 제한조건 개발 및 적정예선용량 추정" 과제 결과의 일부임을 밝혀둔다.

## 참고 문헌

- [1] LG-Caltex 정유 해상출하팀, 아차 사고 사례집, 2000.
- [2] 연윤석, "가중 선형 연상기억을 채용한 유전적 프로그래밍과 그 공학적 응용", 한국 CAD/CAM학회논문집, 제3권 제1호(1998), 57~67.
- [3] 이경호, 연윤석, "다중 인공신경망과 유전적 프로그래밍의 복합적 접근에 의한 공학설계 시스템의 개발", 한국전문가시스템학회지, 제4권 제1호(1998), 25~40.
- [4] Gruber, M.H.J., Regression Estimator: A Comparative Study, Academic Press, New York, 1990.
- [5] Kalaba, R.E., Z. Lichtenstein, T. Simchony

- and L. Tesfatsion, "Linear and Nonlinear Associative Memories for Parameter Estimation", *Information Sciences*, Vol. 61 (1992), 45~66.
- [6] Kalaba, R.E. and F.E. Udwarda, "Associative Memory Approach to the Identification of Structural and Mechanical Systems", *J. of Optimization Theory and Applications*, Vol. 76, No.2(1993), 207~223.
- [7] Kohonen, T., *Self-Organization and Associative Memory*, Springer-Verlag, New York, 1988.
- [8] Koza, J.R., *Genetic Programming : On the Programming of Computers by Means of Natural Selection*, The MIT Press, 1992.
- [9] Koza, J.R., *Genetic Programming II : Automatic Discovery of Reusable Programs*, The MIT Press, 1994.
- [10] Lin, J.C. and D.M. Durand, "Weight Linear Associative Memory Approach to Non-linear Parameter Estimation", *J. of Optimization Theory and Applications*, Vol. 90, No.1(1996), 139~159.
- [11] Yeun, Y. S., K. H. Lee, S. M. Han and Y. S. Yang, "Smooth fitting with a method for determining the regularization parameter under the genetic programming algorithm", *Information Sciences*, Vol. 133, No.3(2001), 175~194.

Abstract

## Decision Support System for Arrival/Departure of Ships in Port by using Enhanced Genetic Programming

Kyung-ho Lee\*  
Yun-seog Yeun\*\*  
Wook Rhee\*\*\*

The main object of this research is directed to LG Oil company harbor in kwangyang-hang, where various ships ranging from 300 ton to 48000 ton DWT use seven berths in the harbor. This harbor suffered from inefficient and unsafe management procedures since it is difficult to set guidelines for arrival and departure of ships according to the weather conditions and the current guidelines dose not offer clear basis of its implications. Therefore, it has long been suggested that these guidelines should be improved.

This paper proposes a decision-support system, which can quantitatively decide the possibility of entry or departure on a harbor by analyzing the weather conditions (wind, current, and wave) and taking account of factors such as harbor characteristics, ship characteristics, weight condition, and operator characteristics. This system has been verified using 5<sup>th</sup> and 7<sup>th</sup> berth in Kwangyang-hang harbor.

Machine learning technique using genetic programming (GP) is introduced to the system to quantitatively decide and produce results about the possibility of entry or arrival, and weighted linear associative memory (WLAM) method is also used to reduce the amount of calculation the GP has to perform. Group of additive genetic programming trees (GAGPT) is also used to improve learning performance by making it easy to find global optimum.

**Key words:** Genetic Programming, Decision-making System, Machine Learning, Ship Arrival/Departure

---

\* Korea Research Institute of Ships and Ocean Engineering

\*\* Dept. of Mechanical Design Engineering, Daejin University

\*\*\* Dept. of Electronic Engineering, Daejin University