

論文2003-40CI-6-10

## 강화와 다양화의 조화를 통한 협력 에이전트 성능 개선에 관한 연구

## (Performance Improvement of Cooperating Agents through Balance between Intensification and Diversification)

李承寬\*, 鄭泰忠\*

(Seung-Gwan Lee and Tae-Choong Chung)

## 요약

휴리스틱 알고리즘 연구에 있어서 중요한 분야 중 하나가 강화(Intensification)와 다양화(Diversification)의 조화를 맞추는 문제이다. 개미 집단 최적화(Ant Colony Optimization, ACO)는 최근에 제안된 조합 최적화 문제를 해결하기 위한 메타 휴리스틱 탐색 방법으로, 그리디 탐색(greedy search)뿐만 아니라 긍정적 반응의 탐색을 사용한 모집단에 근거한 접근법으로 순회 판매원 문제(Traveling Salesman Problem, TSP)를 풀기 위해 처음으로 제안되었다. 본 논문에서는 ACO접근법의 하나인 개미 집단 시스템(Ant Colony System, ACS)에서 강화와 다양화의 조화를 통한 성능 향상 기법에 대해 알아본다. 먼저 에이전트들의 방문 횟수 적용을 통한 상태전이는 탐색 영역을 넓힘으로써 에이전트들이 더욱 다양하게 탐색하게 한다. 그리고, 전역 갱신 규칙에서 전역 최적 경로만 갱신하는 전통적인 ACS알고리즘에서 대하여, 경로 사이클을 구성한 후 각 경로에 대해 긍정적 강화를 받는 엘리트 경로를 구분하는 기준을 정하고, 그 기준에 의해 추가 강화하는 방법을 제안한다. 그리고 여러 조건 하에서 TSP문제를 풀어보고 그 성능에 대해 기존의 ACS 방법과 제안된 방법을 비교 평가해, 해의 질과 문제를 해결하는 속도가 우수하다는 것을 증명한다.

## Abstract

One of the important fields for heuristic algorithm is how to balance between Intensification and Diversification. Ant Colony Optimization(ACO) is a new meta heuristic algorithm to solve hard combinatorial optimization problem. It is a population based approach that uses exploitation of positive feedback as well as greedy search. It was first proposed for tackling the well known Traveling Salesman Problem(TSP). In this paper, we deal with the performance improvement techniques through balance the Intensification and Diversification in Ant Colony System(ACS). First, State Transition considering the number of times that agents visit about each edge makes agents search more variously and widen search area. After setting up criteria which divide elite tour that receive Positive Intensification about each tour, we propose a method to do addition Intensification by the criteria. Implemetation of the algorithm to solve TSP and the performance results under various conditions are conducted, and the comparison between the original ACS and the proposed method is shown. It turns out that our proposed method can compete with the original ACS in terms of solution quality and computation speed to these problem.

**Keywords** : 개미집단최적화(ACO), 개미집단시스템(ACS), 강화(Intensification), 다양화(Diversification), 메타 휴리스틱(Meta Heuristics)

\* 正會員, 慶熙大學校 컴퓨터工學科

(Kyunghee University, Dept. of Computer Engineering)

接受日字:2003年8月9日, 수정완료일:2003年10月29日

## I. 서론

휴리스틱 알고리즘 연구에 있어서 중요한 분야 중 하나가 강화(Intensification)와 다양화(Diversification)의 조화를 맞추는 문제이다. 조합최적화 문제인 순회외판원 문제(Traveling Salesman Problems)<sup>[9]</sup>, 일정 계획 문제(Job-shop scheduling Problems), 이차배정 문제(Quadratic assignment Problems), 차량 경로 문제(vehicle routing Problems), 그래프 착색 문제(graph coloring Problems)등에서 전통적으로 최적의 해를 구하기 위해 사용되는 널리 알려진 메타 휴리스틱(Meta Heuristics)방법으로 유전자 알고리즘(Genetic Algorithm, GA)<sup>[3]</sup>, 타부서치(Tabu Search, TA), 시뮬레이티드 어닐링(Simulated Annealing, SA), Lin-Kernighan(LK) Algorithm<sup>[10]</sup> 등이 있다. 이러한 휴리스틱 탐색에서 강화와 다양화 사이의 조화 문제는 해당 알고리즘의 성능에 결정적인 영향을 미치고 있다.

본 논문에서는 TSP문제를 풀기 위해 Colomi, Dorigo 그리고 Maniezzo<sup>[1, 2]</sup>에 의해 처음 제안된 지역 갱신과 전역 갱신 과정을 거쳐 최적해를 구하는 메타 휴리스틱 방법인 ACS 알고리즘<sup>[6, 7]</sup>에 대해 소개를 하고, 이 ACS 알고리즘에서 에이전트라 불리는 각 개미들이 최적 경로(Best Tour)만을 갱신하는 일반적 전역 갱신 방법과 달리 에이전트들이 구성한 경로 사이클(Tour Cycle)에 대해 긍정적 강화를 받을 엘리트 경로를 구분하고, 구분된 경로 사이클에 대해 방문 횟수를 추가 전역 강화하는 방법과 각각의 에이전트들이 경로 사이클을 이루는 동안 방문한 모든 간선에 대한 방문 횟수를 상태전이 규칙에 적용해 에이전트들이 탐색영역을 더욱 다양하게 검색하게 하는 제안된 방법이 ACS 알고리즘 성능에 어떤 영향을 미치는지 비교 분석해 본다.

## II. Ant Colony Optimization Algorithms

### 1. Ant System

개미 시스템(Ant System, AS)은 실제 개미들이 먹이에서 집까지 가장 짧은 경로를 찾는 능력을 모방한 메타 휴리스틱 탐색<sup>[1, 2, 4, 7]</sup>으로 최근에는 강화학습

(Reinforcement Learning)의 특별한 한 분야로 소개되고 있다<sup>[5]</sup>.

에이전트라 불리는 개미들이 목적지를 향해 나아가는 동안 각 경로에 페로몬을 분비하고, 이후에 지나가는 에이전트들은 그 경로에 쌓여있는 페로몬 정보를 이용해 다음 경로를 선택하는 원리를 휴리스틱 탐색에 적용시킨 시스템이 AS이다<sup>[4, 6, 7]</sup>. 그러나, 이 AS는 에이전트들이 짧은 경로가 있으면 그것만을 선택하고자 하는 성질로 인하여 국부 최적에 빠질 확률이 높아지기 때문에, 이 문제를 쉽게 해결하기 위해 확률분포를 이용해서 다음 노드를 선택하는 과정을 추가한 ACS 알고리즘이라는 방법이 새롭게 연구되었다.

### 2. Ant Colony System

개미 집단 시스템(Ant Colony System, ACS) 알고리즘은 AS의 성능을 향상시키기 위해 Dorigo and Gambardella<sup>[4, 8]</sup>에 의해 소개되었다.

일반적으로 ACS는 다음과 같은 방법으로 수행을 한다. 먼저  $m$ 개의 에이전트들이 초기화 규칙에 따라 무작위로  $n$ 개의 노드(node)를 선택한 다음, 각 에이전트들은 상태전이 규칙(State Transition Rule)에 따라 다음에 방문할 노드를 선택하고 계속해서 탐색과정을 거친다. 이러한 탐색과정을 거치는 동안 에이전트들은 지역 갱신 규칙(Local Updating Rule)에 따라 방문한 각 간선에 페로몬 양을 변경하게 된다. 그리고 일단 모든 에이전트들이 탐색과정을 마치게 되면 전역 갱신 규칙(Global Updating Rule)에 따라 다시 한번 페로몬 양을 변경하게 된다. 결국, AS처럼 각 에이전트들은 짧은 간선을 선택하려는 휴리스틱 정보와 많은 양의 페로몬을 가진 간선을 선택하려는 페로몬 정보에 따라 탐색경로를 완성하게 된다.

계속해서, ACS 알고리즘의 상태전이 규칙(State Transition Rule), 지역 갱신 규칙(Local Updating Rule), 전역 갱신 규칙(Global Updating Rule)에 대해 살펴보고자 한다<sup>[4]</sup>.

#### 2.1 The State Transition Rule

노드( $r$ )에 있는 에이전트가 노드( $s$ )로 이동할 때 식 (1)을 사용해 다음 노드로 이동하며, pseudo random proportional action choice rule<sup>[4]</sup>으로 불린다. 여기서  $\tau(r, u)$ 는 노드( $r$ )과 노드( $u$ )사이 간선의 페로몬의 양,  $\eta = 1/\delta$  은  $\delta(r, u)$  (노드 $r$ 과  $u$ 의 거리)의 역수이고,  $J_k(r)$  은 노드( $r$ )에 있는 에이전트가 방문할 수 있는 남아있

는 노드들의 집합이다. 그리고  $\beta$ 는 페로몬과 간선 길이의 상대적인 중요도를 결정하는 파라미터이다( $\beta > 0$ ).

$$s = \begin{cases} \arg \max_{u \in J_k(r)} \{[\tau(r,u)] \cdot [\eta(r,u)]^\beta\} & \text{if } q \leq q_0 \\ S & \text{otherwise} \end{cases} \quad (1)$$

$q$ 는  $[0,1]$ 사이에서 분포된 무작위 파라미터이고,  $q_0$ 는  $[0,1]$ 사이의 값을 가지는 인자,  $S$ 는 식 (2)에서 주어진 확률분포에 따라 선택된 무작위 파라미터이다. 이것은 무조건 페로몬과 간선 길이의 연산만으로 다음 노드를 선택하는 전통적인 AS와는 달리, 확률분포를 이용해서 다음 노드를 선택하는 과정이 추가됨으로 인해 AS가 가지고 있는 국부 최적에 빠지기 쉬운 한계에서 어느 정도 벗어나고자 하고 있다.

$$p_k(r,s) = \begin{cases} \frac{[\tau(r,s)] \cdot [\eta(r,s)]^\beta}{\sum_{u \in J_k(r)} [\tau(r,u)] \cdot [\eta(r,u)]^\beta} & \text{if } S \in J_k(r) \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

### 2.2 The Local Updating Rule

에이전트들은 TSP 문제를 해결하기 위해 각 간선들을 방문하는 동안 아래의 지역 갱신 규칙을 적용시켜 페로몬의 양을 갱신시킨다.

$$\tau(r,s) \leftarrow (1 - \rho) \cdot \tau(r,s) + \rho \cdot \Delta\tau(r,s) \quad (3)$$

$\rho$  ( $0 < \rho < 1$ )는 페로몬 자연 파라미터이고,  $\Delta\tau(r,s) = \tau_0 = (n \cdot L_m)^{-1}$ 는 초기 페로몬 양으로, 여기서  $L_m$ 은 nearest neighbor heuristic에 의해 생성된 경로길이,  $n$ 은 노드수이다. 일반적으로  $\Delta\tau(r,s)$  값을 결정할 때는 다음의 세가지 전략이 있다. 첫번째 전략은 Q-학습에서 사용되는 방법과 유사한 방법이다. 에이전트들이 각각의 가능한 상태와 그 상태 중 하나를 선택했을 때 처하게 되는 다음 상태를 동시에 평가하여 최선의 선택을 하는 것이다. 이 경우는 다음과 같이 설정한다.

$$\Delta\tau(r,s) = \gamma \cdot \text{Max}_{z \in J_k(s)} \tau(s,z) \quad (4)$$

위의 식 (4)에서 에이전트들이 현재 노드( $r$ )에 위치하면서 다음 선택하게 될 노드( $s$ )를 선택하는 경우지만, 임의의 노드( $s$ )를 선택한 후 노드( $s$ )에서 선택할 수 있는 임의의 노드( $z$ )를 위한 페로몬의 값까지 계산하여

다음의 상황도 고려하는 것을 알 수 있다. 이 전략을 Ant-Q라고 한다<sup>[5]</sup>.

두번째 전략은  $\Delta\tau(r,s)$ 의 값을 초기 페로몬 값으로 설정하는 것으로 단순 ACS(simply ACS)라고 한다<sup>[4]</sup>.

$$\Delta\tau(r,s) = \tau_0 \quad (5)$$

세번째 전략은  $\Delta\tau(r,s)$ 의 값을 0으로 설정하는 방법이다. 일반적으로 순회 외판원 문제에 적용된 초기 페로몬 결정은 단순 ACS를 사용한다. 이것은 실험에 의해 0으로 설정하는 것보다 우수한 성능을 가지고, Ant-Q와 단순 ACS는 비슷한 성능을 보이지만 Ant-Q가 계산량이 많아 시간이 많이 소요되므로 일반적으로 순회 외판원 문제에서는 단순 ACS를 적용한다<sup>[4]</sup>.

### 2.3 The Global Updating Rule

전역 갱신은 에이전트들이 모든 경로 사이클을 완성 후에 수행되는데, 가장 짧은 전체 경로를 완성한 에이전트의 경로에 대해 다음 전역 갱신 규칙을 적용한다.

$$\tau(r,s) \leftarrow (1 - \alpha) \cdot \tau(r,s) + \alpha \cdot \Delta\tau(r,s)$$

$$\text{여기서 } \Delta\tau(r,s) = \begin{cases} (L_{gb})^{-1} & \text{if } (r,s) \in \text{global\_best\_tour} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

$\alpha$  ( $0 < \alpha < 1$ )는 페로몬 자연 파라미터,  $L_{gb}$ 는 현재까지의 전역 최적 경로 길이이다.  $\tau(r,s)$ 는 노드( $r$ )과 노드( $s$ ) 사이의 페로몬 양으로, 전역 최적 경로에 속해 있으면  $(1/L_{gb})$ , 아니면 0으로 주어지며, 결국 자연 파라미터  $\alpha$ 에 의해 페로몬 양은 줄어들게 된다.

## III. 강화와 다양화의 조화

휴리스틱 알고리즘 연구에 있어서 중요한 분야 중 하나가 강화(Intensification)와 다양화(Diversification)의 균형을 맞추는 문제이다. 본 논문에서는 각 에이전트들이 상태전이를 통한 다양한 경로 탐색과, 모든 에이전트들이 경로 사이클을 구성 후 구성된 사이클에 대해 긍정적 강화를 받는 경로로 구분해 전역 강화하는 방법을 소개하고, 기존 방법과의 비교를 통해 그 효율성을 비교한다.

1. 강화(Intensification)와 다양화(Diversification)의 조화 강화(Intensification) 전략은 역사적으로 발견된 좋은

해의 장점을 이용하여 움직이게 하고, 다양화(Diversification)은 아직까지 탐색하지 않은 새로운 영역으로 다양하게 찾아가게 하는 것이다. 여기서 강화전략은 새로운 지역으로 탐색을 하게 하는 다양화 전략과는 근본적으로 반대되는 것 같다. 하지만 이 두 전략은 항상 상호 반대되는 것이 아니고 때때로 이 둘 사이의 지배권을 바꾸기 보다는 균형을 찾음으로써 더욱 효과적으로 다룰 수 있다.

사실, 강화 그 자체만으로 일반적인 최적화 문제에 가장 좋은 결과를 얻는 데는 불충분하다. 강화에만 의존하는 시스템은 최적해에 도달하는 올바른 움직임을 찾는데 실패한다. 즉 적당한 새로운 지역으로 탐색을 이끄는 충분한 다양화 도입에 실패한다. 현재 문제해결 방법으로 본능적인 대응은 임의적으로 움직여 운 좋게 효과적인 움직임을 찾는 것이다. 그러한 움직임도 가능성이 있다. 그러나 이것은 다양화를 얻기 위한 우연적인 방법이다. 이러한 임의적인 탐색방법은 좋은 움직임을 추구하는데 필요한 지속적인 균형을 잡는 효과를 잃는다. 임의적인 방법은 다양화를 추구하는 목적에 맞추지만, 눈먼 선택을 하는 것으로 나쁜 결과가 도출될 수 있다. 따라서, 다양화 전략은 강화와 상보적인 작용으로 가장 효과적인 탐색을 위하여 함께 사용되어야 한다. 다양화 성질로 해를 전략적으로 추구하는 것은 강화작용에서 약점을 보완하여 내부적으로 강화탐색단계와 반대영역을 가로질러 확장된 다양화 탐색단계 사이에 연속적으로 탐색방법을 추구해 본질적인 균형을 이뤄야 한다.

강화는 수렴 속도를 빠르게 하고 다양화는 좋은 해를 찾도록 하는데 필요한 요소이다. 하지만, 강화를 강조하면 에이전트들은 국부 최적해에 수렴하게 되고, 다양화를 강조하면 부적절한 상태(해의 발견)의 원인이 되기 때문에, 휴리스틱 탐색과정에 있어서, 특히 본 논문에서 적용되는 ACS 알고리즘의 성능을 향상시키는 중요한 요소 중의 하나로 강화(Intensification)와 다양화(Diversification)사이의 적절한 균형을 취하는 것이다.

ACS에서 상태전이 규칙을 적용하는 것은 탐색의 다양화 보장을 위한 방법이다. 이것은 먼저 발견된 좋은 해(또는 추정된 형태일 수 있는)를 포함한 지역에 집중적으로 탐색하는 대신에, 더 멀리 검토되지 않은 지역으로 진행을 유도한다. 이것은 새로운 시작점을 만들기 위하여 특수하게 설계된 개선기준을 휴리스틱 탐색과

정에 사용하는 것이다. 이로 인해 임의로 시작점을 찾는 대신 방법 대신에 목적을 갖고 그러한 점을 찾게 된다(확률배정방법에 의하여 제약된 무작위 방법을 사용하는 것).

따라서, ACS에서 순회 외관원 문제를 해결하기 위한 다양화 전략으로 에이전트들이 이웃 노드로 상태전이 과정시 확률배정방법을 적용할 때 해당 간선에 대해 다른 에이전트들의 방문횟수를 적용해 평가하는 것이다. 그 횟수에 기초하여 각 간선들에 평가값을 주는 것은 과거에 많이 사용된 간선들을 회피하려는 시도로, 이 방법은 상태전이 탐색 과정을 미탐사 지역으로 유도하기 위해 빈도수에 기초한 조건을 적용하여 사용하는 것이다.

## 2. 다양화를 통한 상태전이

ACS에서 노드( $r$ )에 있는 에이전트가 노드( $s$ )로 이동할 확률은 식 (1)의 상태전이 규칙에 의해 수행되는데 여기서 탐험(exploration)과 탐색(exploitation)의 균형은 매우 중요하다.

앞서 2절에서  $q$ 와  $q_0$ 에 대해 기술하였는데, 이것은 새로운 간선에 대한 다양한 탐험(다양화)과 축적된 정보에 대한 탐색(강화) 사이에서의 균형을 직접적으로 유지하기 위하여 이용된다.

에이전트들은 경로를 탐색동안 휴리스틱 정보와 페로몬 정보를 이용하는데, 만약  $q < q_0$ 인 경우 에이전트는 탐색 행동을 취하는데 그것은 오랜 기간동안 최선의 선택으로 페로몬 정보와 짧은 기간동안 경험적 지식으로서 거리와 관련된 휴리스틱 값을 사용한다.  $q < q_0$ 인 경우에는 에이전트는 다양한 탐험을 수행하기 위하여 확률 분포( $S$ )를 사용하여 길이가 더 짧고 많은 양의 페로몬을 가진 간선 선택을 선호하게 된다.

그러나, 식 (2)를 적용한 일반적인 ACS에서는 최적 경로에 속한 간선 선택 확률만 높게 만들기 때문에 에이전트들은 확률 분포( $S$ ) 적용의 기본 목적인 다양한 탐험 수행을 할 수 없게 된다.

따라서, 현재 최적 경로로 새로운 이웃 간선으로의 다양한 탐험의 역할을 고려해 각 간선에 에이전트들의 방문 횟수를 고려할 수 있는데, 이것은 탐험의 비율을 빠리하고 탐험의 정확성을 점차로 개선함으로써 강화 학습에 있어서 효과적일 수 있다.

다양한 탐험의 역할을 강화하기 위해서 확률분포로 방문 횟수를 적용할 수 있는데, 축적된 페로몬과 방문 횟수를 다양한 탐험에 동시에 적용하기 위해 식 (2)을

다음 식 (7)과 같이 수정할 수 있다. 여기서 파라미터  $\delta$ 는 현재 경로 사이클 동안 각 에이전트들이 간선  $E(r,s)$ 에 방문한 횟수를 확률적으로 적용한 것이다.

$$p_k(r,s) = \begin{cases} \frac{[\tau(r,s)]^\beta \cdot [\eta(r,s)]^\beta}{\sum_{u \in J_k(r)} [\tau(r,u)]^\beta \cdot [\eta(r,u)]^\beta} & \text{if } s \in J_k(r) \\ 0 & \text{otherwise} \end{cases}$$

여기서  $\text{if } f(r,s) \neq 0 \quad \delta = -1/f(r,s)$   
 $\text{else} \quad \delta = 1$  (7)

$f(r,s)$ 는 현재까지 에이전트들이 간선  $E(r,s)$ 를 방문한 빈도수로, 에이전트들이 해당 간선에 대해 한번이라도 방문했을 경우  $\delta$ 는 방문 빈도수에 대한 음의 역수(-1/방문 빈도수)를 취한다. 그러나, 현재 경로 사이클에서 해당 간선에 대해 한번도 방문하지 않은 경우  $\delta=1$ 로 초기화한다. 결국 식 (7)에 의해 에이전트들은 새로운 탐색공간으로의 다양한 탐색으로 인해 더욱 다양하게 새로운 간선을 선택할 수 있게 만든다. 즉, 학습 에이전트들은 더 많은 페로몬을 가진 상태 대신에 방문 횟수가 적은 상태를 선택할 수 있게 됨으로써, 임의의 적합한 정책에 대한 집중을 회피하고 최적해를 신속히 탐색하게 하는 장점이 있다.

3. 전역 갱신에서의 긍정적 강화

ACS알고리즘에서의 일반적인 전역갱신에 의한 강화 방법을 살펴보면, 각 에이전트들이 모든 경로 사이클을 완성 후 이루어지며, 전역 최적 경로에 속해 있는 간선만 갱신시키는 방법<sup>[6, 7, 9]</sup>과 현재 사이클의 최적 경로에 속해 있는 간선의 페로몬을 갱신시키는 방법<sup>[8]</sup>, 각 경로 사이클마다 우성과 열성 인자들로 구분하고 각각의 우성과 열성 인자들에 대해 전역 갱신 규칙에 기반해 생성된 모든 경로 사이클에 페로몬 값을 갱신시키는 방법<sup>[11]</sup>, 그리고 각 경로 사이클에서 방문한 각 간선들이 최적해를 이루는데 얼마나 적합한지를 구분하는 기준으로 현재 단계의 모든 경로 사이클을 완성 후 각 간선에 방문한 에이전트들의 방문 횟수를 강화값으로 표현하고, 그 강화값을 경로 사이클을 구성한 모든 간선의 페로몬 갱신에 추가하는 방법을 채택함으로써 방문 횟수가 많을수록 강화값을 높게 평가하는 방법이 있다<sup>[12]</sup>. <표 1>에 전역 갱신 방법이 잘 비교 설명되어 있다.

본 논문에서는 <표 1>에 나타나 있는 전역 강화 방

표 1. 전역 갱신 방법  
Table 1. Global Updating Method.

	기존 방법	제안 방법
전역 갱신 방법	현재 사이클 최적 경로 or 전역 최적 경로 or 우성열성 경로구분에 의한 갱신(50% 기준) or 모든 경로 갱신 강화값(방문횟수)	엘리트 경로 갱신 + 강화값(방문횟수)

법 대신에, 새로운 긍정적 강화를 받는 엘리트 경로 사이클 구분하는 방법과 전역 강화하는 방법을 제안한다.

엘리트 경로를 나누는 기준은 현재 경로 사이클에서 최적의 경로인 경우와 일반 경로 사이클에서 길이가 이전 사이클까지의 전역 최적 경로 길이보다 작거나 같은 경우 이 경로를 엘리트 경로로 간주한다.

강화방법은 에이전트들이 구성한 엘리트 경로 사이클의 모든 간선에 대해서 각 간선에 방문한 에이전트들의 방문 횟수를 확률적으로 표현하고 방문 횟수가 많을수록 강화를 높게 평가하는 방법이다. 갱신 방법은 다음의 식 (8)에 의해 갱신된다.

$$\tau(r,s) \leftarrow (1-\alpha) \cdot \tau(r,s) + \alpha \cdot (\Delta\tau(r,s) + \gamma \cdot \text{InfoT}(r,s))$$

여기서  $\Delta\tau(r,s) = \begin{cases} (L_k)^{-1} & \text{if } (r,s) \in \text{EliteTour} \\ 0 & \text{otherwise} \end{cases}$

$$\text{InfoT}(r,s) = \begin{cases} \frac{f(r,s)}{m} & \text{if } (r,s) \in \text{EliteTour} \\ 0 & \text{otherwise} \end{cases}$$

$$L_k = \frac{W}{E}, \quad W = \sum L_i \quad (8)$$

$L_k$ 는 에이전트들이 방문한 해당 간선을 포함하고 있는 엘리트 경로들의 평균 길이,  $W$ 는 전체 엘리트 경로들의 길이의 합,  $E$ 는 엘리트 경로 개수,  $L_i$ 는 각각의 엘리트 경로의 길이를 나타낸다.  $\text{InfoT}(r,s)$ 는 엘리트 경로에 대해서 전체 에이전트 수에 대한 간선별 방문 빈도수를 나타내며,  $f(r,s)$ 는 현재까지 에이전트들이 엘리트 경로의 간선  $E(r,s)$ 에 방문한 빈도수,  $m$ 은 에이전트 수를 나타낸다.  $\gamma$  ( $0 < \gamma < 1$ )는 할인계수(discount factor), 여기서 할인계수( $\gamma$ )를 사용하는 것은 지속적인 환경에서 보상을 무한대로 할 수 없기 때문에 일정

비율로 할인을 하는 것이다. 이 방법은 엘리트 경로에 대해 에이전트들이 방문한 각 간선에 방문한 횟수를 보상값으로 적용하여 페로몬을 갱신함으로써 에이전트들은 각 간선의 휴리스틱 정보를 이용해 탐색할 수 있게 된다.

IV. 성능측정 및 분석

본 논문에서 제안하고 있는 방법을 실험하기 위해서 도시들의 위치는 TSP 예제로 널리 알려진 TSPLIB<sup>[13]</sup>에서 추출하여 실험을 하였다. 실험 환경은 PentiumIV 1700MHz, 512MB RAM 환경에서 수행되었다. 사용된 파라미터들의 값은 실험에 의해 다음과 같이 결정되었으며, 실험에 의해 결정된 값은  $\beta=2$ ,  $\alpha=\rho=0.1$ ,  $q_0=0.9$ ,  $\gamma=0.1$ ,  $\tau_0=(n*L_m)^{-1}$  이다. 그리고 각각의 실험에 사용된 에이전트 수  $m=10$ 이고 초기 위치 결정은 에이전트들을 각 노드에 1개씩 무작위로 배정하였다. 종료 조건은 고정된 수행 횟수 또는 여러 실험에 의해 최적값으로 알려진 값을 찾았을 경우 종료한다.

전통적인 ACS방법과 본 논문에서 제안한 방법의 효율성 분석을 위해 최적 해에 얼마나 빨리 수렴하는가를 각각 비교 분석하였다.

<그림 1>은 KroA100.TSP를 이용한 수렴속도를 보여주는 것으로, 초기에는 전통적인 ACS방법이 빠른 경향이 있다. 그러나 시간이 지남에 따라 다양한 탐색을 거친 제안된 방법이 더 빠르게 수렴하는 것을 볼 수 있다. 이것은 상태전이 과정에서 각 에이전트들이 다양하게 다음 노드를 선택함으로써 초기 수렴속도는 느리지만 사이클 횟수가 많아 질수록 점점 빠르게 최적해에 수렴하게 된다.

<표 2> <표 3>은 문제 영역의 차이에 따른 성능 평

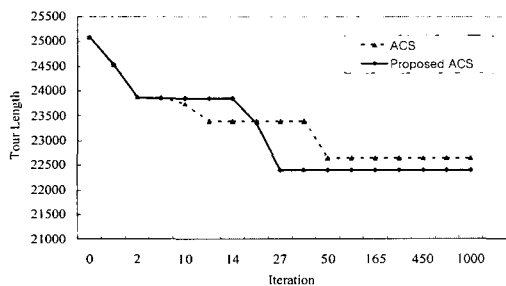


그림 1. KroA100.TSP를 이용한 수렴속도  
Fig. 1. Convergence Speed KroA100.TSP.

표 2. ACS 성능평가(Smaller Problem)  
Table 2. ACS Performance(Smaller Problem).

Node Set	ACS		Proposed ACS	
	Average Length	Best Length	Average Length	Best Length
City Set1	47.30	45.42	46.59	<b>44.12</b>
City Set2	51.88	48.98	50.43	<b>47.50</b>
City Set3	45.19	42.17	44.29	<b>42.31</b>
City Set4	49.54	48.46	46.17	<b>45.73</b>
City Set5	48.20	47.52	46.51	<b>43.75</b>
Eil51	434.82	428.31	428.65	<b>425.94</b>
St70	684.75	679.75	683.46	<b>675.51</b>

표 3. ACS 성능평가(Bigger Problem)  
Table 3. ACS Performance(Bigger Problem).

Node Set	ACS		Proposed ACS	
	Average Length	Best Length	Average Length	Best Length
KroA150	26949.25	26524	26844.32	<b>26524</b>
Rat195	2451.63	2341	2435.61	<b>2332</b>
Gil262	2486.75	2425	2426.45	<b>2402</b>
A280	2642.58	2629	2616.32	<b>2594</b>
Pr299	48497.78	48395	48401.39	<b>48292</b>
Lin318	43744.37	43417	43297.62	<b>42896</b>
Pr439	109441.02	108931	108927.47	<b>108443</b>

가를 보이고 있다. 10회 시행에 20000번, 100000번 사이클을 반복했을 경우의 결과로, 제안된 방법의 성능이 훨씬 우수하다는 것을 볼 수 있다. 결국, 강화와 다양화의 균형을 통한 탐색은 '좋은' 시작 경로들을 만들고, 미탐사 지역으로의 유도가 자유롭기 때문에 국부최적에서 쉽게 벗어날 수 있고, ACS 알고리즘 성능 개선에 효과적인 방법임을 보여준다.

V. 결론과 앞으로의 연구 방향

본 논문은 ACS 알고리즘에서 강화와 다양화의 조화를 통한 성능 향상 방법을 제안하고 있다.

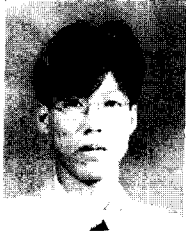
에이전트들이 구성한 경로 사이클에 대해 긍정적 강화를 받을 엘리트 경로를 구분하고, 구분된 엘리트 경로 사이클에 대해 방문 횟수를 추가 전역 강화하는 방법과 각각의 에이전트들이 경로 사이클을 이루는 동안 방문한 간선에 대한 방문 횟수를 상태전이 규칙에 적

용해 에이전트들이 탐색영역을 더욱 다양하게 검색하게 하는 제한된 방법이 ACS 알고리즘 성능 향상에 많은 도움이 되었다. 이것은 강화(Intensification) 전략을 통해 역사적으로 발견된 좋은 해의 장점을 이용하여 움직이게 하고, 다양화(Diversification) 전략을 통해 아직까지 탐색하지 않은 새로운 영역으로 다양하게 찾아가기 때문이다. 그러나 본 논문에서는 이러한 강화와 다양화를 적용하기 위해 각 간선에 방문한 빈도수를 기반으로 다양화 전략과 강화 전략을 채택했는데, 향후 연구과제는 병렬처리를 위한 다중 집단 에이전트에서 강화와 다양화의 조화를 이루기 위한 다른 요소에 대해 추가 연구가 필요하겠다.

### 참 고 문 헌

- [1] A. Colomi, M. Dorigo, and V. Maniezzo, "An investigation of some properties of an ant algorithm," Proceedings of the Parallel Problem Solving from Nature Conference(PPSN 92), 1992, R. Manner and B. Manderick (Eds.), Elsevier Publishing, pp. 509-520.
- [2] A.Colomi, M. Dorigo, and V. Maniezzo, "Distributed optimization by ant colonies," Proceedings of ECAL91 - European Conference of Artificial Life, Paris, France, 1991, F.Varela and P.Bourgine(Eds.), Elsevier Publishing, pp. 134-144.
- [3] B. Freisleben and P. Merz, "Genetic local search algorithm for solving symmetric and asymmetric traveling salesman problems," Proceedings of IEEE International Conference of Evolutionary Computation, IEEE-EC 96, 1996, IEEE Press, pp. 616-621
- [4] L.M. Gambardella and M. Dorigo, "Ant Colony System: A Cooperative Learning approach to the Traveling Salesman Problem" IEEE Transactions on Evolutionary Computation, vol. 1, No. 1, 1997.
- [5] L. M. Gambardella and M. Dorigo, Ant-Q: a reinforcement learning approach to the traveling salesman problem, Proceedings of ML-95, Twelfth International Conference on Machine Learning, A. Prieditis and S. Russell (Eds.), Morgan Kaufmann, 1995, pp. 252-260.
- [6] M. Drigo, V.Maniezzo, and A.Colomi, "The ant system: optimization by a colony of cooperation agents," IEEE Transactions of Systems, Man, and Cybernetics-Part B, vol. 26, No. 2, pp. 29-41, 1996.
- [7] M. Dorigo and G. D. Caro, "Ant Algorithms for Discrete Optimization" Artificial Life, vol5, no.3, pp. 137-172, 1999.
- [8] M. Dorigo & L.M. Gambardella (1997). "Ant Colonies for the Traveling Salesman Problem". BioSystems, 43:73-81.
- [9] P-C. Kanellakis and C.H. Papadimitriou, "Local search for the asymmetric traveling salesman problem," Operations Research, vol. 28, no. 5, pp. 1087-1099, 1980.
- [10] S. Lin and B.W. Kernighan, "An effective Heuristic algorithm for the traveling salesman problem," Operations Research, vol.21, pp. 498-516, 1973.
- [11] SeungGwan Lee, TaeUng Jung and TaeChoong Chung, "Improved Ant Agents System by the Dynamic Parameter Decision," Proceedings of IEEE International Conference on FUZZ-IEEE 2001, IEEE Press, pp. 666-669.
- [12] SeungGwan Lee and TaeChoong Chung, "A Study about Additional Reinforcement in Local Updating and Global Updating for Efficient Path Search in Ant Colony System," The KIPS Transactions : Part B, pp. 237-242, 2003. 06.
- [13] <http://www.iwr.uni-heidelberg.de/groups/comopt/software/TSPLIB95/>

## 저 자 소 개



李承寬(正會員)

1997년 : 경희대학교 전자계산공학과(공학사). 1999년 : 경희대학교 대학원 전자계산공학과(공학석사). 2001년~현재 : 경희대학교 대학원 전자계산공학과 박사과정 수료.  
 <주관심분야 : 인공지능, 지능에이

전트, 메타알고리즘, 스케줄링>



鄭泰忠(正會員)

1980년 : 서울대학교 전기공학과(공학사). 1982년 : 한국과학기술원 대학원 전자계산공학과(공학석사). 1987년 : 한국과학기술원 대학원 전자계산공학과(공학박사). 1987년~1988년 : KIST 시스템 공학센터

선임 연구원. 2001년 : 미국 Iowa대학 교환교수. 1988년~현재 : 경희대학교 컴퓨터공학과 정교수. <주관심분야 : 인공지능, 지능에이전트, 메타알고리즘>