https://doi.org/10.5392/JKCA.2020.20.12.278

이진 입자 군집 최적화를 이용한 반복 죄수 딜레마 게임 분석

Analysis on Iterated Prisoner's Dilemma Game using Binary Particle Swarm Optimization

이상욱
목원대학교 정보통신융합공학부

Sangwook Lee(slee@mokwon.ac.kr)

요약

죄수 딜레마 게임은 게임 이론의 대표적인 사례로 많은 경제학자, 사회과학자 및 컴퓨터 과학자가 관심을 가지고 연구하고 있다. 근래에는 죄수 딜레마 게임 분석을 위해 유전 알고리즘, 입자 군집 최적화 등의 진화 연산 기법을 적용한 계산적 접근에 대한 연구가 활발히 이루어져 왔다. 본 연구에서는 3가지의 서로 다른 이진 입자 군집 최적화 기법을 사용하여 2명 또는 그 이상의 플레이어가 참여하는 반복 죄수 딜레마 게임에 대한 전략을 진화시켜보고자 한다. 반복 죄수 딜레마 게임에 3가지 버전의 이진 입자 군집 최적화를 적용하여 실험한 결과 자신의 이득을 최대화하기 위한 이기적인 참가들 사이에서도 상호 협력 관계가 구축될 수 있음을 확인하였나 참여자가 많을수록 상호 협력 관계가 구축이 어려워 짐을 확인하였다.

■ 중심어 : | 반복 죄수 딜레마 게임 | 이진 입자 군집 최적화 | 진화 | 협동 | 배신 |

Abstract

The prisoner's dilemma game which is a representative example of game theory is being studied with interest by many economists, social scientists, and computer scientists. In recent years, many researches on computational approaches that apply evolutionary computation techniques such as genetic algorithms and particle swarm optimization have been actively conducted to analyze prisoner dilemma games. In this study, we intend to evolve a strategy for a iterated prisoner dilemma game participating two or more players using three different binary particle swarm optimization techniques. As a result of experimenting by applying three kinds of binary particle swarm optimization to the iterated prisoner's dilemma game, it was confirmed that mutual cooperation can be established even among selfish participants to maximize their own gains. However, it was also confirmed that the more participants, the more difficult to establish a mutual cooperation relationship.

■ keyword : | Iterated Prisoner's Dilemma Game | Binary Particle Swarm Optimization | Evolution | Cooperation | Defection |

I. 서 론

근래에 유전 알고리즘 등의 진화 프로그래밍을 사용하여 반복 죄수 딜레마 게임의 협동으로의 진화에 관한 연구가 많이 진행되어 왔다. 악셀로드(Axelrod)[1]는 해집단에서 각 개체는 자신을 제외한 다른 개체 모두와 2명 참가자의 반복 죄수 딜레마 게임을 진행하여 얻은 결과를 적합도로 계산하여 진화하는 유전 알고리즘을

접수일자 : 2020년 11월 10일
수정일자 : 2020년 11월 24일
심사완료일 : 2020년 11월 24일
교신저자 : 이상욱, e-mail : slee@mokwon.ac.kr

제안하였다. 해집단 내의 모든 전략들이 진화를 통해 동시에 변화하기 때문에 전략은 모든 세대에서 서로 다른 환경에 의해 평가되는 특징을 가지고 있다. 즉, 해집단 내의 모든 전략들은 그들의 동적인 환경에서 공진화하는 것이다. 악셀로드는 이러한 동적 환경에서도 전략들이 해집단 내의 다른 전략들을 상대로 우수한 성능을 내기 위해 적절히 진화한다는 사실을 발견하였다. 포겔(Fogel)[2]은 전략을 표현하기 위해 유한 상태 기계를 사용하였고 전략을 진화시키기 위해 진화 프로그래밍을 사용하였다. 유한 상태 기계 사용의 장점 중 하나는 해의 전략을 인코딩하기 위해 필요한 기억 용량에 대한 고민을 하지 않아도 된다는 점이었다.

본 연구에서는 유전 알고리즘과 유사한 해집단 기반 메타휴리스틱 기법인 이진 입자 군집 최적화 기법과 수정된 이진 입자 군집 최적화 기법을 사용하여 반복 죄수 딜레마 게임의 협동으로의 진화를 관찰하고자 한다.

II. 반복 죄수 딜레마 게임

1. 죄수 딜레마

형사 두 명이 공범으로 의심되는 용의자 두 명을 서로 다른 공간의 수사실에서 각각 범죄 사실에 대해 자백을 권유하고 있다. 만일 두 명 용의자의 자백 여부에 따른 형량은 다음과 같다고 하자.

가. 두 용의자 모두 범죄 사실 자백하지 않음
=> 둘 다 징역 1년

나. 두 용의자 모두 범죄 사실 자백
=> 둘 다 징역 3년

다. 용의자 1명 범죄 사실 자백, 1명 자백하지 않음
=> 자백한 용의자: 무혐의 석방
자백하지 않은 용의자: 징역 5년

이와 같은 상황에서 용의자는 범죄 사실을 자백하는 것(배신)과 자백하지 않는 것(협동) 중 어느 것이 더 최선인가. 이 문제에서 용의자 둘 전체의 이득을 생각한다면 둘 모두 자백하지 않는 것이 가장 최선이다. 왜냐하면 둘 모두 자백하지 않았을 때인 '가'의 형량 합은 2년으로 다른 두 경우의 형량 합인 '나' 6년, '다' 5년보다 적기 때문이다. 그러나 용의자 모두 상대의 협동 및 배신 여부를 알 수 없는 상황에서 본인의 협동 및 배신 여부를 결정해야 하므로 결국 둘 다 배신을 할 수 밖에 없다. 왜냐하면 상대의 협동 및 배신의 경우 수 모두에서 내가 배신하는 것이 더 형량을 적게 받기 때문이다. 이렇게 서로가 협동하면 최선이나 상대의 선택과 상관없는 절대 우위 전략이 배신이어서 결국 둘 모두 이 전략을 선택하는 것을 죄수 딜레마(Prisoner's Dilemma, PD)라고 한다.

2. 반복 죄수 딜레마 게임

2명이 참여하는 죄수 딜레마 게임(2-players PD, 2PD)은 서로 상대의 선택을 알 수 없는 상황에서 이루어지는 난-제로섬(non-zero sum) 게임이다. 게임을 통해 각 참여자가 얻는 이득은 다음과 같이 2×2 행렬로 표현할 수 있다.

표 1. 2PD 게임 참여자 이득 행렬

		참여자	
		협동	배반
참여자1	협동	(R, R)	(S, T)
	배반	(T, S)	(P, P)

단, $T > R > P > S$ 이고 $R > (T+P)/2$ 이다.

반복 죄수 딜레마 (Iterated Prisoner's Dilemma, IPD) 게임은 위와 같은 PD 게임 라운드를 여러 번 반복하는 게임을 말한다. 이때 참여자는 현재 라운드에서 상대방의 선택은 알 수 없으나 이전 라운드에서의 자신 및 상대방의 전략과 획득한 이득은 기억할 수 있다. 2명이 참여하는 반복 죄수 딜레마 게임(2-players IPD, 2IPD)은 많은 실세계의 사회, 경제 분야의 문제를 모델링 하는데 사용되어 왔다[3].

3. 여러 명이 참여하는 반복 죄수 딜레마 게임

2IPD가 사회, 경제 여러 분야의 문제를 모델링 할 수는 있으나 간단한 문제에만 가능하고 복잡한 문제에는 적용하기 어렵다는 단점이 있었다. 이러한 단점을 보완

하고 사회 및 경제 분야의 복잡한 문제를 모델링하기 위해 여러 명이 참여하는 반복 죄수 딜레마(n-players Iterated Prisoner's Dilemma, nIPD) 게임이 제안되었다. 데이비스(Davis)[4]는 nIPD는 에너지 절약, 생태학 및 인구 과잉 문제 등의 많은 실제 문제가 nIPD 패러다임으로 표현할 수 있다고 설명하였다. 그리고 콜먼(Colman)[5], 글랜스(Glance)[6] 및 허버만(Huberman)[7]은 nIPD는 2IPD와 질적으로 상이하며 2IPD에서 좋은 성능을 보여주는 전략이 여러 사람이 참여하는 IPD에서는 좋은 성능을 보여주지 못할 수도 있음을 보여주었다.

nIPD 게임은 다음과 같은 3가지 특징으로 정의할 수 있다.

가. 각 참여자는 협동(Cooperation, C)과 배신(Defection, D) 중 하나를 선택할 수 있다.

나. D 옵션은 각 참가자에 대해 우위 전략이다. 즉, 얼마나 많은 다른 참여자가 C를 선택하든 상관없이 각 참여자가 C보다 D를 선택하는 것이 유리하다.

다. 우위 전략인 D는 불완전 균형과 만난다. 모든 참여자가 C 전략을 선택하는 경우에는 이득 측면에서 D를 선택하는 것보다 C를 선택하는 것이 낫지만, D 전략에서 C 전략으로 움직이는 동기는 없다.

이러한 특징을 가지고 있는 nIPD 게임을 통해 참가자가 얻을 수 있는 이득은 다음 표와 같이 정의할 수 있다.

표 2. nIPD 게임 참여자 이득 행렬

		남은 n-1 참여자 중 협동하는 참여자 수				
		0	1	2	…	n-1
참여자1	C	C_0	C_1	C_2	…	C_{n-1}
	D	D_0	D_1	D_2	…	D_{n-1}

단, $0 < i < n$-1에 대하여 (1) $D_i > C_i$ (2) $D_{i+1} > D_i$, $C_{i+1} > C_i$ (3) $C_i > (D_i + C_{i-1})/2$ 이며 이득 행렬은 각 참여자에 대해 대칭구조를 갖고 있다.

[표 2]의 제약조건을 만족하는 행렬은 매우 많지만 본 연구에서는 [표 3]과 같이 [8]에서 사용한 행렬을 사용한다.

표 3. nIPD 게임 참여자 이득 행렬 (예시)

		남은 n-1 참여자 중 협동하는 참여자 수				
		0	1	2	…	n-1
참여자1	C	0	2	4	…	$2(n-1)$
	D	1	3	5	…	$2(n-1)+1$

n명이 참여하는 nIPD의 어느 게임 라운드에서 n_c명의 참여자가 협동하였다고 하면 참여자들의 평균 이득 g는 계차 수열 계산으로 다음과 같이 구할 수 있다.

$$g = 1 + \frac{n_c}{n}(2n-3) \tag{1}$$

이 결과로부터, 평균 이득 g를 구하기만 하면 협동하는 참여자의 비율 (n_c/n)을 다음과 같이 쉽게 구할 수 있다.

$$\frac{n_c}{n} = \frac{g-1}{(2n-3)} \tag{2}$$

III. 이진 입자 군집 최적화

입자 군집 최적화 (Particle Swarm Optimization, PSO)는 새 또는 벌 등의 곤충이 먹이를 구하기 위해 무리를 지어 이동하는 것에 영감을 얻어 개발된 해집단 기반 메타 휴리스틱 기법이다[9]. [그림 1]은 PSO의 개념도를 보여주고 있다. PSO는 초기화를 통해 각 해집단을 랜덤하게 생성하고 입자 속도 갱신 및 입자 위치 갱신 과정을 통해 해공간을 탐색한다. 평가를 통해 갱신한 해의 적합도를 구하고 지역 최적해 및 전역 최적해를 갱신하여 해를 개선해 나간다. 이러한 과정은 종료 조건을 만족할 때까지 수행된다.

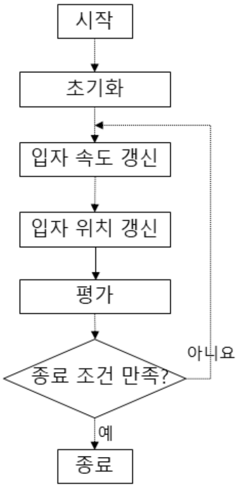

그림 1. 입자 군집 최적화 개념도

PSO의 입자 속도 갱신 함수는 식(1), 입자 위치 갱신 함수는 식(2)과 같다.

$$v_{i,j}(t+1) = wv_{i,j}(t) + c_1R_1(P_{best,i,j} - x_{i,j}(t)) + c_2R_2(G_{best,i,j} - x_{i,j}(t)) \quad (1)$$

$$x_{i,j}(t+1) = x_{i,j}(t) + v_{i,j}(t+1) \quad (2)$$

여기서 t는 세대(Generation) 수, i는 해집단 인덱스, j는 해차원 인덱스를 의미한다. 그리고 w는 관성 상수(Inertia Constant), c_1과 c_2는 가속 상수(Accelerator Constant), R_1과 R_2는 0과 1 사이의 난수를 의미한다. 또한 $v_{i,j}(t)$는 t 세대에서 i번째 해의 j차원 속도, $x_{i,j}(t)$는 t 세대에서 i번째 해의 j차원의 위치, $P_{best,i,j}$는 i번째 해의 j차원의 지역 최적해, $G_{best,i,j}$는 i번째 해의 j차원의 전역 최적해를 의미한다.

1. 이진 입자 군집 최적화

최초에 제안된 PSO는 연속 공간 문제 해결을 위해 제안된 기법으로 조합 최적화 문제 등의 불연속 공간에는 적용할 수 없다. 이러한 단점을 해결하고자 이진 입자 군집 최적화 (Binary Particle Swarm Optimization, BPSO)가 제안되었으며 PSO를 최초로 제안한 케네디(Kennedy)와 에버하트(Eberhart)가 1997년에 가장 먼저 BPSO를 제안하였다 [10]. 입자의 속도를 [0, 1]사이의 값을 반환하는 S자형 함수(Sigmoid function)에 대입하여 이진 값인 0 또는 1을 선택할 확률값으로 사용하는 방식으로 입자 속도 갱신 함수는 식(1)을 그대로 사용하나 입자 위치 갱신 함수는 식(3)과 같이 변경되었다.

$$x_{i,j}(t+1) = \begin{cases} 0 \text{ if } rand() \geq S(v_{i,j}(t+1)) \\ 1 \text{ if } rand() < S(v_{i,j}(t+1)) \end{cases}$$

(3)

여기서 $rand()$는 0과 1 사이의 난수를 의미하고 $S(\cdot)$는 Sigmoid function을 의미하며 식(4)와 같이 정의된다.

$$S(v_{i,j}(t+1)) = \frac{1}{1+e^{-v_{i,j}(t+1)}} \quad (4)$$

2. 수정된 이진 입자 군집 최적화 (버전 1)

BPSO는 PSO와는 달리 입자 위치 갱신에 이전 세대의 위치 정보를 활용하지 않고 속도 정보 사용한다. 이상욱[11]은 BPSO도 PSO와 같이 입자 위치 갱신에 이전 세대의 위치 정보를 사용하기 위하여 속도를 탐색 범위로 설정하고 속도와 위치를 유전자형-표현형 기법으로 해석한 수정된 이진 입자 군집 최적화 (Modified Binary Particle Swarm Optimization, MBPSO1)를 제안하였다. MBPSO1은 식(1), 식(2), 식(3)을 식(5), 식(6), 식(7)로 변경하여 구현한다.

$$v_{i,j}(t+1) = wv_{i,j}(t) + c_1R_1(P_{p,best,i,j} - x_{p,i,j}(t)) + c_2R_2(G_{p,best,i,j} - x_{p,i,j}(t)) \quad (5)$$

$$x_{g,i,j}(t+1) = x_{g,i,j}(t) + v_{i,j}(t+1) \quad (6)$$

$$x_{p,i,j}(t+1) = \begin{cases} 0 \text{ if } rand() \geq S(x_{p,i,j}(t+1)) \\ 1 \text{ if } rand() < S(x_{p,i,j}(t+1)) \end{cases} \quad (7)$$

3. 수정된 이진 입자 군집 최적화 (버전 2)

MBPSO는 입자 속도 갱신에 표현형을 사용하기 때문에 현재해와 최적해에 (0, 1, -1)의 3가지 경우의 값만 적용 가능하여 탐색 경로가 한정되어 있다는 문제가 있다. 임승균[12]은 이러한 단점을 보완하기 위하여 식(6)에 유전자형을 사용하는 MBPSO2를 제안하였다.

$$v_{i,j}(t+1)=wv_{i,j}(t)+c_1R_1(P_{g,best,i,j}-x_{g,i,j}(t)) \\ +c_2R_2(G_{g,best,i,j}-x_{g,i,j}(t)) \quad (8)$$

IV. 반복 죄수 딜레마 게임 관찰을 위한 이진 입자 군집 최적화 설계

1. 해 표현

유전 알고리즘, 입자 군집 최적화, 차분 진화 알고리즘 등 해집단 기반 진화 연산 기법을 해결하고자 하는 문제에 적용하기 위해서는 적절한 해 표현을 설계해야 한다. 단일 라운드 죄수 딜레마 게임에서 찾고자 하는 것은 이득을 극대화 할 수 있는 협동 또는 배신의 2가지 전략 중 하나의 선택이며 이는 0과 1의 이진수 한자리로 표현할 수 있다. 반복 죄수 딜레마 게임은 각 게임의 라운드마다 협동할 것인지 배신할 것인지를 선택하는 것이며 이전 라운드에서 진행했던 게임에 대한 정보인 나의 전략과 나를 제외한 상대들의 전략을 기억하고 있다가 현재 라운드의 전략을 결정하는데 사용할 수 있다.

n명 참여자가 진행한 반복 죄수 딜레마 게임에서 한 라운드의 전략을 기억하기 위해서는 n개의 이진 비트가 필요하며 이에 대한 경우의 수는 2^n이다. 그러나 '나를 제외한 참여자들이 어떤 전략을 사용하였는가'보다 '나를 제외한 참여자 중 몇 명이 협동하였는가'에 대한 정보만으로도 나의 전략을 결정할 수 있으므로 $\log_2 n$ 만큼의 이진 비트만으로 정보를 저장할 수 있다. 따라서 한 개 라운드의 정보를 저장하기 위해서는 (나의 전략) + (나를 제외한 참여자 중 협동자의 수) 만큼의 이진 비트가 필요하여 총 $\log_2 n+1$ 만큼의 이진 비트가 필요하다. 진행했던 이전 게임의 여러 라운드를 저장하는 경우 저장할 이전 라운드의 개수를 h라 하면 총 $h(\log_2 n+1)$ 만큼의 이진 비트가 필요하다.

저장한 이전 라운드 기억에 대한 각 경우에 대해 전략을 수립하기 위해서는 $2^{h(\log_2 n+1)}$만큼의 이진 비트가 필요하다. 따라서 h 만큼의 이전 라운드를 기억하는 nIPD게임을 분석하기 위한 해 표현은 (기억 저장에 필요한 이진 비트) + (기억에 대한 경우 수 만큼의 전략 수립에 필요한 이진 비트)으로 총 $2^{h(\log_2 n+1)}+h(\log_2 n+1)$ 만큼의 이진 비트를 사용한다.

2. 초기화

$2^{h(\log_2 n+1)}+h(\log_2 n+1)$ 길이 만큼의 이진 비트로 이루어진 해집단 해들의 비트 값을 0 또는 1로 랜덤하게 설정하여 생성한다.

3. 입자 속도 및 위치 갱신

BPSO, MBPSO1 및 MBPSO2의 입자 속도 및 위치 갱신 방법에 따라 입자 속도 및 위치 갱신한다.

4. 평가

반복 죄수 딜레마 게임에 한 번 참여하였을 때 라운드별 게임에서 획득한 이득의 총합을 총 라운드 수로 나눈 값을 '한 게임에 대한 개체의 적합도'로 정의한다. 만약 한 세대에서 여러 번 게임에 참여한 경우는 게임 당 획득한 적합도의 총합을 참여한 게임 수로 나눈 값을 '한 세대에 대한 개체의 적합도'로 정의한다. 한 세대에 대한 개체의 적합도가 이전 세대까지 발견한 지역 최적해 및 전역 최적해와 대비하여 개선되었을 경우 지역 최적해 및 전역 최적해를 갱신한다.

본 연구의 nIPD 게임은 해집단에서 n 명의 개체(참여자)를 랜덤하게 선택하여 100번의 라운드 반복하는 것을 한게임으로 하고 이러한 게임을 미리 정한 게임 수 만큼 한 세대에서 진행하도록 설계하였다.

5. 종료 조건

식(1)에서 협동하는 참여자의 비율(n_c/n)이 95%가 넘는 평균 적합도 값이 5세대 이상 나타날 경우는 협동

으로 수렴하였다고 판단하고 종료한다. 또한 설정한 최대 세대 수까지 수렴에 도달하지 못할 경우도 종료한다. 여기서 평균 적합도는 해당 세대 모든 개체의 '한 세대에 대한 개체의 적합도'를 평균하여 계산한다.

V. 실험 결과

1. 실험 환경

실험환경은 Intel(R) Core(TM) i7-9700 CPU 3.00GHz 성능의 옥타코어 CPU와 8GB메모리의 컴퓨팅 환경에서 Visual studio 2019 소프트웨어로 코딩하여 구현하였다.

BPSO 구현에 필요한 변수인 $c_1 = c_2 = 0.69$, $w = 1.43$으로 설정하였으며 입자 속도의 발산을 방지하기 위해 입자 속도 갱신 후 속도의 값을 제한하는 속도의 최대값은 $V_{\max} = 6$으로 설정하였다.

본 연구에서는 BPSO 기반 nIPD 게임에서 어떤 요인들이 참여자들을 협동으로 수렴하게 하는지를 확인하기 위하여 아래와 같은 실험을 진행하였다. 각 실험을 총 20번을 반복하여 협동으로 수렴한 횟수를 작성하였으며 평균 수렴 세대 수는 협동으로 수렴한 경우의 세대 수를 평균하여 작성하였다.

- nIPD 게임 참여자 수 변화에 따른 수렴 특성
- 기억 라운드 수 변화에 따른 수렴 특성
- 해집단 수 변화에 따른 수렴 특성
- 세대 당 게임 수 변화에 따른 수렴 특성

마지막으로는 2IPD 실험에서 상위 10개의 우위 전략 특성을 살펴보았다.

입자 군집 최적화 기법은 관성 상수 값에 따라 속도가 발산할 위험이 있기 때문에 최대 속도를 설정하여 속도 갱신이 설정된 최대 속도 ($V_{\max}$) 넘지 못하도록 제한한다. 이에 따라 속도는 [$-V_{\max}$, $V_{\max}$] 사이의 값을 가지며, 식(1)로 계산된 속도의 값이 이 범위를 벗어난 경우 $-V_{\max}$ 또는 $V_{\max}$ 값으로 제한된다.

2. 기억 라운드 수 변화에 따른 수렴 특성

[표4]는 기억 라운드 수 2, 해집단 100, 게임 수 1,000일 때 게임 참여자 수에 따른 수렴 특성을 보여주고 있다. 참여자의 수가 증가하면 협동으로 수렴하는 횟수가 감소함을 확인할 수 있으며 특히 참여자가 32명인 경우에는 3개의 BPSO 모두 협동으로 수렴하지 못하였다. 큰 차이는 없으나 MBPSO1, MBPSO2가 BPSO에 비해 좋은 수렴 특성을 보임을 확인하였다.

표 4. 기억 라운드 수: 2, 해집단: 100, 게임 수: 1,000

nIPD 크기	협동 수렴 횟수			평균 수렴 세대 수		
	BPSO	MBPSO 1	MBPSO 2	BPSO	MBPSO 1	MBPSO 2
2IPD	16	17	18	120.5	118.6	130.5
3IPD	12	16	15	130.7	125.9	128.3
4IPD	13	14	15	125.3	140.5	130.9
5IPD	9	12	13	152.3	168.2	159.4
6IPD	10	11	11	210.5	220.8	232.4
8IPD	8	10	9	207.8	219.4	240.5
16IPD	0	5	6	-	251.7	249.9
32IPD	0	0	0	-	-	-

[표 5]는 [표 4] 대비 기억 라운드 수를 증가시켜 수렴 특성을 확인한 결과이며 기억 라운드 수 변경은 수렴 특성에 별다른 영향을 주지 않음을 확인하였다.

표 5. 기억 라운드 수: 3, 해집단: 100, 게임 수: 1,000

nIPD 크기	협동 수렴 횟수			평균 수렴 세대 수		
	BPSO	MBPSO 1	MBPSO 2	BPSO	MBPSO 1	MBPSO 2
2IPD	15	16	18	111.6	108.5	123.1
3IPD	13	17	17	87.8	114.5	101.4
4IPD	9	10	12	168.1	177.2	152.5
5IPD	11	13	15	220.4	190.3	174.6
6IPD	7	12	13	192.3	211.6	252.2
8IPD	8	9	10	321.2	294.5	302.4
16IPD	0	7	7	-	281.6	299.2
32IPD	0	0	0	-	-	-

3. 해집단 수 변화에 따른 수렴 특성

[표 6]는 [표 4] 대비 해집단 수를 증가시켜 수렴 특성을 확인한 결과이다. 5IPD, 6IPD에서 협동 수렴 횟수가 다소 늘어났음을 확인할 수 있으나 다른 결과는

큰 차이가 없어 해집단 수 변화는 수렴 특성에 큰 영향을 주지 않음을 확인하였다.

표 6. 기억 라운드 수: 2, 해집단: 200, 게임 수: 1,000

nIPD 크기	협동 수렴 횟수			평균 수렴 세대 수		
	BPSO	MBPSO 1	MBPSO 2	BPSO	MBPSO 1	MBPSO 2
2IPD	16	17	17	89.1	97.2	96.7
3IPD	17	17	16	110.2	108.4	112.7
4IPD	15	15	16	135.2	143.6	133.8
5IPD	13	14	14	210.6	182.5	190.2
6IPD	14	14	15	176.8	201.4	213.7
8IPD	10	9	11	238.6	261.8	249.8
16IPD	0	6	7	-	182.6	190.3
32IPD	0	0	0	-	-	-

4. 게임 수 변화에 따른 수렴 특성

[표 7]은 [표 4] 대비 세대 당 게임 수를 증가시켜 수렴 특성을 확인한 결과이다. 해집단 수 변화와 마찬가지로 5IPD, 6IPD에서 협동 수렴 횟수가 다소 늘어났음을 확인할 수 있으나 다른 결과는 큰 차이가 없어 게임 수 변화는 수렴 특성에 큰 영향을 주지 않음을 확인하였다.

표 7. 기억 라운드 수: 2, 해집단: 100, 게임 수: 10,000

nIPD 크기	협동 수렴 횟수			평균 수렴 세대 수		
	BPSO	MBPSO 1	MBPSO 2	BPSO	MBPSO 1	MBPSO 2
2IPD	17	18	18	66.7	69.2	79.4
3IPD	16	17	18	126.3	135.6	141.2
4IPD	15	16	16	212.4	178.4	193.5
5IPD	12	13	14	230.4	199.7	201.2
6IPD	11	12	13	210.4	203.8	217.3
8IPD	10	10	10	267.3	283.8	251.7
16IPD	0	5	5	-	186.4	192.5
32IPD	0	0	0	-	-	-

5. 우위 전략 특성 분석

기억 라운드 수 2, 해집단 수 100, 게임 수 10,000의 2IPD 게임에서 해표현은 [그림 2]와 같다. 전략 비트 및 기억 비트의 인덱스 0, 1번 비트에서 0은 협동, 1은 배신을 의미한다.

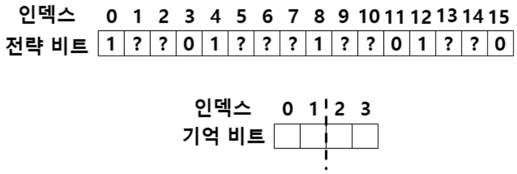

그림 2. 기억 라운드 수 2, 해집단 수 100, 게임 수 10,000의 2IPD 게임의 해표현

[그림 2]에서 기억 비트 인덱스가 의미하는 바는 다음과 같다.

- 0 : 나의 2라운드 이전 전략
- 1 : 나의 1라운드 이전 전략
- 2 : 2라운드 이전 협동 참여자 수
- 3 : 1라운드 이전 협동 참여자 수

다음은 협동으로 수렴했을 때 상위 10개의 전략을 표시한 것이다.

1 1 0 0 1 0 1 1 1 0 1 0 1 0 1 0
1 0 1 0 1 1 1 0 1 1 0 0 1 1 0 0
1 1 0 0 1 1 0 0 1 0 1 0 1 0 1 0
1 1 1 0 1 1 1 0 1 1 0 0 1 1 1 0
1 0 1 0 1 0 1 1 1 1 0 0 1 0 0 0
1 1 0 0 1 1 0 1 1 0 1 0 1 1 1 0
1 0 1 0 1 1 1 0 1 1 0 0 1 1 0 0
1 0 1 0 1 0 1 0 1 0 1 0 1 1 1 0
1 1 0 0 1 1 0 1 1 1 0 0 1 0 1 0
1 0 0 0 1 0 1 0 1 1 1 0 1 1 0 0

상위 10 전략에서 예외 없이 나타나는 비트들은 다음과 같다.

1 ? ? 0 1 ? ? ? 1 ? ? 0 1 ? ? 0

이것을 분석하면 다음과 같다.

- 0번, 4번, 8번, 12번 전략 비트의 값이 1이 의미하는 바는 이전 두 라운드에서 나를 제외한 참여자 중 협동자가 없었던 경우 나의 이전 두 라운드 전략과 상관

없이 배신한다는 뜻임

- 3번, 11번, 15번 전략 비트의 값이 0이 의미하는 바는 이전 두 라운드에서 나를 제외한 참여자 모두가 협동한 경우 내가 직전 라운드에서 배신하지 않았다면 협동한다는 뜻임
- 위의 경우를 제외한 나머지 경우는 협동할 수도 배신할 수도 있다는 뜻임

VI. 결 론

본 연구에서는 이진 입자 군집 최적화를 활용하여 반복 죄수 딜레마 게임을 분석해 보았다. 분석 결과 반복 죄수 딜레마 게임의 참여자 수가 증가할수록 협동으로 수렴하기가 어려움을 확인하였으며 그 외에 입자 군집 최적화 및 죄수 딜레마 게임의 변수 들은 수렴 특성에 별다른 영향을 주지 않음을 확인하였다. 또한 반복 죄수 딜레마 게임에서 이전 라운드에서 협동자가 없었을 때는 배신하고 모두 협동하였을 때는 협동하는 전략들이 우위를 보임을 확인하였다.

게임 참여자 수가 증가할수록 협동으로 수렴하기가 어려운 이유는 탐색 공간의 증가하여 문제의 복잡도가 증가하였기 때문으로 추정된다. 이러한 결과를 실세계 환경에 비추어 보면 참여자 수가 많을수록 대다수 참여자가 협동하더라도 소수가 배반할 경우 배반하는 개체의 적합도가 높은데 이를 협동으로 이끌어 가기 위해서는 다수의 협동자가 계속해서 협동으로 피해를 보아야 하므로 협동하는 개체는 손해를 견디지 못해 결국 배반을 선택하게 되어 배반으로 수렴하는 결과를 초래하게 되는 것과 유사하다. 본 연구를 통해 이진 입자 군집 최적화를 통한 반복 죄수 딜레마 게임 분석이 실세계 모습과 유사한 결과를 보여줌을 확인하였다. 추후 연구로 참여자 수가 증가하더라도 협동으로 진화할 수 있는 기법에 대해 연구할 필요가 있다.

참 고 문 헌

[1] R. Axelrod, “The evolution of strategies in the iterated prisoner's dilemma,” in Genetic Algorithms and Simulated Annealing (L. Davis, ed.), ch.3, pp.32-41, San Mateo, CA: Morgan Kaufmann, 1987.

[2] D. B. Fogel, “Evolving behaviors in the iterated prisoner's dilemma,“ Evolutionary Computation, Vol.1, No.1, pp.77-97, 1993.

[3] A. Rapoport, “Optimal policies for the prisoner's dilemma,“ Tech. Rep. 50, The Psycho metric Lab., Univ. of North Carolina, Chapel Hill, NC, USA, July 1966.

[4] J. H. Davis, P. R. Laughlin, and S. S. Komorita, “The social psychology of small groups,” Annual Review of Psychology, Vol.27, pp.501-542, 1976.

[5] A. M. Colman, *Game Theory and Experimental Games*, Oxford, England: Pergamon Press, 1982.

[6] N. S. Glance and B. A. Huberman, “The outbreak of cooperation,“ Journal of Mathemat ical Sociology, Vol.17, No.4, pp.281-302, 1993.

[7] N. S. Glance and B. A. Huberman, “The dynamics of social dilemmas,” Scientific American, pp.58-63, March 1994.

[8] X. Yao and P. Darwen, “Genetic algorithms and evolutionary games,” Commerce, Complexity and Evolution, W. Barnett, C. Chiarella, S. Keen, R. Marks, and H. Schnabl (eds.), Chapter 16, pp.313-333, Cambridge University Press, 2000.

[9] J. Kennedy and R. Eberhart, “Particle Swarm Optimization,” Proceedings of The IEEE International Conference of Neural Networks, IV (pp.1942-1948), Piscataway:IEEE, 1995.

[10] R. C. Eberhart, “A discrete binary version of the particle swarm algorithm,” Proceedings of 1997 conference systems man cybernetics, pp.4104-4108, 1997.

[11] S. W. Lee, S. M. Soak, S. H. Oh, Witold Pedrycz, and M. G. Jeon, “Modified binary particle swarm optimization,” Progress in Natural Science, Vol.18, No.9, pp.1161-1166, 2008.

[12] 임승균, 이상욱, "유전자형-표현형 개념을 적용한 수정된 이진 입자군집최적화 (버전 2)," 한국콘텐츠학회논문지, Vol.14, No.11, pp.541-548, 2014.

저 자 소 개

이 상 욱(Sangwok Lee) **종신회원**

- 2000년 2월 : 한국과학기술원 기계공학과(공학사)
- 2002년 2월 : 광주과학기술원 기전공학과(공학석사)
- 2007년 8월 : 광주과학기술원 정보기전공학부(공학박사)
- 2007년 8월 ~ 2007년 9월 : 조지아공대 전산학과 박사후연구원
- 2008년 11월 ~ 2009년 2월 : 삼성전자 통신연구소 책임연구원
- 2009년 3월 ~ 현재 : 목원대학교 정보통신융합공학부 교수

〈관심분야〉 : 휴리스틱 알고리즘, 인공지능, 최적화