

기생체 숙주 이론 기반의 경쟁 공진화 신경망

박정은⁰ 박민재 오경환

서강대학교 컴퓨터학과

{fayemint@ailab⁰, pmj219@ailab, kwoh@ccs}.sogang.ac.kr

Competitive Co-Evolving Neural Network : Host and Parasites

Jung-Eun Park⁰ Minjae Park Kyung-Whan Oh

Dept. of Computer Science, Sogang University

요 약

유전자 알고리즘을 사용하여 신경망의 가중치를 학습하는 방법은 역전파 알고리즘이 가지는 여러 가지 문제점을 해결하기 위해 제안되었으나, 유전자 알고리즘 역시 전역 탐색이 아니기 때문에 실세계의 데이터에 적용하기 어려운 가장 큰 장애 요소인 지역 최소점 문제를 완벽하게 해결할 수는 없다. 이러한 지역 최소점 문제를 완화하기 위해 본 논문에서는 기생체-숙주 공진화 현상을 기반으로 한 유전자 알고리즘을 사용한 경쟁 공진화 신경망 학습 방법을 제시하고 있다. 경쟁 공진화는 서로 다른 개체간의 경쟁적인 진화를 통해 궁극적으로 보다 적합도가 높은 개체가 생성되는 이론을 기반으로 하고 있다. 이러한 경쟁 공진화를 통한 신경망 가중치의 학습이 일반적인 유전자 알고리즘을 사용하여 신경망을 학습시키는 것보다 더욱 우수한 가중치 집단을 탐색할 수 있음을 두 종류의 기계 학습 데이터를 통해 입증하였다.

1. 서 론

진화 인공 신경망은 역전파 알고리즘의 몇 가지 문제점들 때문에 제안되었다. 진화 인공 신경망에 대한 기존의 연구 결과는 많은 경우의 데이터에 대해 최적화 신경망 구조 탐색이 가능함을 보이고 있으나, 항상 좋은 성능을 보이는 것은 아니다.[2][3][4]. 이는 유전자 알고리즘 자체도 전역 탐색을 실시하지 못하기 때문에 종종 지역 최소점에 빠지는 경우가 발생하기 때문이다. 숙주-기생체의 공진화 현상은 이러한 문제를 해결하는 또 다른 접근 방법으로 제안될 수 있다. 숙주-기생체 공진화 현상은 자연계에서 자신을 공격하는 기생체를 방어하도록 진화한 유기체를 모방한 것으로, 이러한 공진화 현상에 의해 진화하는 개체들에게 더욱 어려운 도전을 제시하여 진화 연산의 지역 최소점 문제를 어느 정도 해결할 수 있게 되었다. 본 논문의 연구 목적은 숙주-기생체 공진화 현상의 적용으로 유전자 알고리즘을 사용한 신경망 학습의 지역 최소점 문제를 개선하여 보다 적합도가 높은 신경망 구조의 탐색에 있다.

2. 유전자 알고리즘을 사용한 신경망 가중치 학습

전통적인 신경망의 가중치 학습은 기울기 강하법이 가지는 문제점 때문에 종종 지역 최소점에 수렴하는 단점이 있다. 이러한 단점을 극복하기 위해 제안된 것이 유전자 알고리즘을 사용하여 신경망의 가중치를 진화하는 것이다. 신경망 학습을 위한 기존의 유전자 알고리즘은 다음의 요소들로 구성되어 있다.

- (1) **개체 표현** : 유전자 하나의 실제 값은 하나의 가중치 값의 집합에 일치한다. 각 유전인자는 하나의 신경망의 가중치(weight) 또는 바이어스(bias)값에 대응되며, 이진 표현이나 실수 표현으로 유전자 개체를 표현할 수 있다.
- (2) **적합도 함수** : 유전자로 표현된 신경망 가중치 집합의 적합도 함수는 해당 신경망에 의한 출력의 오차를 사용한다
- (3) **진화 연산자** : D. Montana의 신경망 가중치 진화를 위한 진화 연산은 두 개의 부모로부터 하나의 자식만을 생성한다.[2] 생성된 자식은 돌연변이를 거쳐 새로운 개체로 변화된 기존의 개체들과 비교되어 더 낮은 적합도를 가지는 개체와 교체된다.

3. 기생체-숙주 이론 기반의 경쟁 공진화 신경망

3.1 경쟁 공진화 모델

본 논문에서 제안하고 있는 기생체-숙주 이론 기반의 경쟁 공진화 신경망에서 신경망 자체는 숙주로, 망을 학습하기 위한 학습 데이터 집단은 기생체로 모델링 할 수 있다. 각 집단은 서로 다른 진화 연산을 통해 진화하며 각 집단에서 가장 적합도가 우수한 개체가 다른 집단의 진화를 위해 사용된다.

3.1.1 숙주 진화 : 신경망

신경망의 가중치 진화를 위한 유전자 알고리즘의 메커니즘은 Davis, J. Montana가 그의 연구에서 사용한 유전자 알고리즘과 유사한 형태를 사용하였으며, 추가적인 진화 연산자 몇 개를 사용하였고, 숙주 진화를 위한 구체적인 사항은 다음과 같다.

- **유전형 표현** : 각 신경망은 각 뉴런을 연결하는 가중치와 뉴런에 연결된 바이어스 값을 가지게 된다. 가중치는 -1.0에서 1.0사이의 값 중 임의로 배정하였고, 바이어스는 0에서 1.0사이의 값에서 임의의 값이 선택되었다. 각 가중치 값과 바이어스 값은 다음과 같은 규칙을 기반으로 유전형으로 표현되었다.

- 출력 계층에서 입력 계층으로 (위에서 아래로)
- 좌측 뉴런에서 우측 뉴런으로
- 각 계층에서 가중치 먼저 표현 후 바이어스 표현

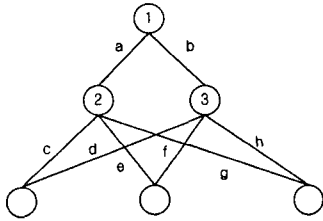
그림 1은 유전형 표현의 예를 보이고 있다. 위의 3가지 규칙에 의해 작성된 유전형이 그림 1의 (b)에 표현되어있다.

- **적합도 함수** : 유전형의 적합도 함수는 유전형에 의해 구성되는 신경망의 오차 제곱을 누적한 값의 역수에 10의 값을 곱하여 크기조정(scaling)을 하였다. 우수한 신경망에 대한 적합도는 전체 오차가 작아지므로 큰 값을 가지게 된다.
- **진화 연산자** : 유전형의 진화를 위한 진화 연산자는 D. Montana가 사용한 진화 연산자와 다른 연구에서 사용된 진화 연산자가 복합적으로 사용되었다. 유전자 알고리즘에 사용된 진화 연산자는 다음과 같다.

- (1) **선택 연산자** : 유전형이 가진 적합도에 비례하여 선택될 수 있도록 두 가지 종류의 선택 연산자-룰렛 휠 선택 연산자와 순위 기반 선택 연산자-가 사용되었다.
- (2) **교배 연산자** : 유전형 교배를 위해 교배 연산자와 선형

(liner) 교배 연산자가 사용되었다.[5]

- (3) 돌연변이 연산자 : n개의 뉴런이 돌연변이를 위해 선택되며, 선택된 뉴런의 입력으로 연결된 모든 가중치와 해당 뉴런의 바이어스값이 임의의 값으로 교체되는데, 초기화 돌연변이, 잠형성 돌연변이 중 하나를 선택하여 사용하였다.
- (4) 교체 연산자 : 진화 연산자에 의해 생성된 자식 유전형을 전체 유전형 집합에 포함시키고 기존의 유전형 중 하나를 집합에서 제거하는 교체 연산자는 비교 또는 무작위 교체 연산자가 사용되었다.



(a) 신경망에서의 가중치와 바이어스



(b) 가중치와 바이어스의 유전형

그림 1. 유전형 표현의 예

3.1.2 기생체 진화 : 학습 데이터 집단

학습 데이터 집단의 진화는 더욱 오차가 큰 학습 데이터를 보유하는 방향으로 진행된다. 이것은 망 집단이 진화 할수록 더욱 오차를 줄이기 힘든 데이터를 적용함으로써 망 집단의 오차 감소 능력을 강화시킨다.

- **유전형 표현** : 학습 데이터 집단(기생체)은 원래의 학습 데이터가 가지고 있는 패턴수와 동일한 패턴수를 가지게 된다. 각 집단을 표현하는 유전형은 학습 패턴의 인덱스를 포함하고 있다. 그림 2는 학습 데이터 집단의 유전형 표현의 예를 보이고 있다. 전체 m개의 기생체는 서로 다른 학습 패턴들의 집단으로 구성된다.

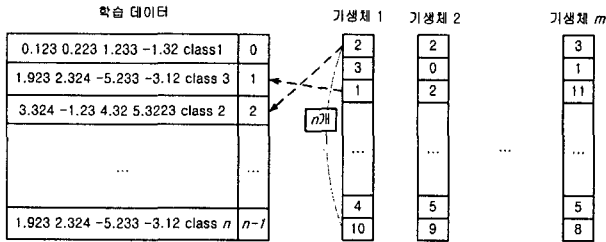


그림 2. 학습 데이터 집단의 유전형 표현의 예

- **적합도 함수** : 기생체의 적합도는 주어진 신경망으로부터 구해지는 모든 오차의 누적으로 결정된다.
- **진화 연산자** : 기생체의 진화를 위해 사용된 진화 연산자는 다음과 같은 표준 유전자 알고리즘 연산자를 사용하였다.
 - (1) 선택 연산자 : 선택 확률 0.5의 순위 기반 선택 연산으로 2개의 부모 유전형을 선택하였다.
 - (2) 교배 연산자 : 일점 교배 연산자를 통해 기생체의 유전형을 교배 하였다. 교배점은 임의로 선택된다.

- (3) 돌연변이 연산자 : 돌연변이 연산은 임의의 n개의 점을 선택하여 임의의 학습 패턴으로 교체한다.
- (4) 교체 연산자 : 선택, 교배, 돌연변이에 의해 생성된 두 자식 유전형은 유전형 집단내에서 제일 작은 적합도를 가지는 개체들과 교체된다. 만일 새로 생성된 유전형이 기존의 개체들보다 적합도가 낮으면 그 자식 유전형은 폐기된다.

3.1.3 공진화 시스템

기생체와 숙주는 동일한 세대동안 서로 다른 진화 연산자를 사용하여 진화된다. 각 진화의 결과에 의해 생성되는 최적합 개체는 서로 다른 종(種)의 적합도 계산을 위해 전달된다. 기생체-숙주 공진화 시스템은 다음과 같은 흐름으로 진화를 진행한다.

- 1) N개의 숙주(신경망)집단 임의 발생.
- 2) 학습 데이터로 각 숙주의 적합도를 계산.
- 3) M개의 기생체(학습 집단) 집단 임의 발생.
- 4) 세대 수 만큼 다음의 진화를 반복
 - a) 첫 번째 세대가 아니라면 단계 4-d)를 통해 얻어진 기생체 중 가장 적합도가 높은 기생체를 사용하여, 각 숙주의 적합도 계산.
 - b) 다음의 숙주 진화를 N번 반복.
 - ① 선택된 두 개의 부모에 돌연 변이를 수행
 - ② 두 개의 부모를 교배, 하나의 자식을 생산.
 - ③ 생성된 자식의 적합도를 숙주 집단내의 최저 적합도와 비교하여 그 값이 크면 해당 유전형과 자식 유전형을 교체, 작으면 자식 유전형을 폐기.
 - c) 단계 4-a)를 통해 얻어진 최고 적합도를 가진 숙주를 사용하여 기생체들의 적합도를 계산.
 - d) 다음의 기생체 진화를 M/2번 반복.
 - ① 두 개의 부모를 교배, 두 개의 자식을 생산.
 - ② 생성된 자식 유전형에 돌연변이를 수행
 - ③ 생성된 자식의 적합도를 숙주 집단내의 최저 적합도와 비교하여 그 값이 크면 해당 유전형과 자식 유전형을 교체, 작으면 자식 유전형을 폐기.

4. 실험 및 결과 분석

제안된 경쟁 공진화 신경망 학습 모델의 적합성을 검증하기 위해 두 개의 UCI Machine Repository 기계학습 데이터를 사용하였다.[6] 신경망 학습이 용이한 속성의 수가 적고 각 변수가 가지는 값의 편차가 작으며, 목적 속성(class)의 종류가 적은 데이터로 Iris를 선택하였고, 신경망 학습이 어려운 데이터로 Glass를 선택했다.

실험에서 사용되는 평가 기준은 반복수의 증가에 따른 최적 적합도 값의 변화를 사용한다. 일반적인 유전자 알고리즘을 사용한 신경망 학습의 경우에는 두 번의 망평가가 한번의 역전파 반복과 동일하다고 볼 수 있다.[2] 이는 주어진 학습 데이터에 대한 역전파 학습은 활성화의 전방향 전파(출력 뉴런에서의 에러 계산)와 역방향 에러 전파(가중치의 조정)인 두 가지 부분으로 이루어졌기 때문이다. 유전자 알고리즘은 단지 첫 번째 부분만을 수행한다. 두 번째 부분이 더 많은 연산을 요구하기 때문에, 두 번의 유전자 알고리즘 평가에 걸리는 시간이 한 번의 역전파 알고리즘 반복시 걸리는 연산시간의 절반보다 적게 걸린다. 기생체-숙주 경쟁 공진화 신경망 학습은 일반적인 유전자 알고리즘 학습 방법보다는 많은 계산을 요구한다. 만일 기생체와 숙주의 집단수가 동일하다면 일반적인 유전자 알고리즘 학습의 경우보다 두 배의 계산량이 요구된다. 즉, 이러한 경우 경쟁 공진화 신경망에 의한 한 번의 망 평가는 한 번의 역전파

알고리즘 수행과 동일하게 된다. 그러나 기생체의 집단수가 숙주의 집단수보다 적으면 한번의 공진화 신경망 학습은 역전과 알고리즘의 한 학습기(epoch)보다는 적은 시간이 소요된다.

4.1 IRIS 데이터를 통한 실험

학습되기 위한 신경망은 두 개의 은닉층을 가지고 있으며 첫 번째 은닉층은 3개의 은닉 뉴런을, 두 번째 은닉층은 2개의 은닉 뉴런을 가지고 있다. 각 진화 연산을 사용한 실험을 통해 숙주 진화를 위해 표준 교배와 잠행성 돌연변이 연산이 사용되었다. 기생체 진화 연산자는 0.6의 선택 확률을 가지는 순위 기반 선택과 비교 교체가 적용되었다. 유전자 알고리즘을 위한 매개 변수 설정은 표 1과 같다.

표 1. IRIS 데이터의 매개 변수 설정

매개변수	설정
교배율	0.8
돌연변이율	0.2
숙주 집단 수	50
기생체 집단 수	10
세대 수	200
엘리트 기법	2% 사용

IRIS 데이터에 대한 세 가지 학습 기법-역전과 알고리즘(BP), 유전자 알고리즘 학습(GA), 경쟁 공진화 학습(Co-evolution)-의 적합도 탐색 결과는 그림 3과 같다. 역전과 알고리즘은 학습률 0.6으로 5000 학습기(epoch)동안 수행되었다. 유전자 알고리즘은 전체 50개의 망을 200세대 동안 평가하므로 총 10,000개의 반복 평가가 이루어진다. 역전과 알고리즘 1회 수행은 유전자 알고리즘은 2회 수행과 같으므로 두 모델 모두 같은 횟수만큼 망을 평가하게 된다.

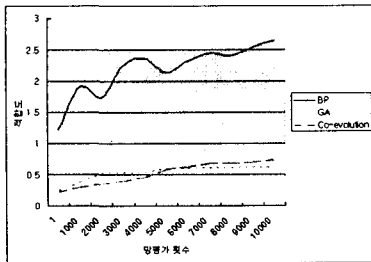


그림 3. IRIS 데이터에 대한 세 가지 학습 기법 결과

오차가 작은 학습 데이터의 경우에 역전과 알고리즘의 성능이 월등함을 알 수 있다. 일반적인 유전자 알고리즘을 사용하는 경우와 경쟁 공진화의 경우를 비교하면 약 5,500회 정도의 망 신경망 학습까지는 유전자 알고리즘의 성능이 우수하지만 그 이후로는 경쟁 공진화 알고리즘의 성능이 우수함을 알 수 있다. 이것은 경쟁 공진화 알고리즘이 더 많은 신경망 평가를 요구하기 때문이다. 같은 횟수의 망 평가를 수행할 경우에 경쟁 공진화 알고리즘은 대략 유전자 알고리즘의 절반 정도의 세대밖에 수행하지 못한다.

4.2 Glass 데이터를 통한 실험

신경망 구조는 Iris의 경우와 동일한 구조가 사용되었다. 숙주 진화를 위해 룰렛 휠 선택, 잠행성 돌연변이와 비교 교체가 사용되었다. 그림 4는 Glass 데이터에 대한 세가지 학습 모델의 적합도 탐색을 보여주고 있다.

오차가 비교적 큰 데이터에 대해서 역전과 알고리즘에 비해 유전자 알고리즘을 사용하여 학습한 모델의 성능이 우수함을 알 수 있다. 또한 경쟁 공진화를 적용한 학습 모델이 일반적인

유전자 알고리즘을 사용한 학습 모델에 비해서 더 나은 성능을 보임을 알 수 있다. 앞선 실험과 비교하여 오차가 큰 데이터에 대해서 경쟁 공진화 학습 모델의 성능과 유전자 알고리즘 학습 모델의 성능간의 격차가 더욱 큼을 관찰 할 수 있다. 이는 오차가 큰 데이터에 대해 유전자 알고리즘이 지역 최소점에 머무를 확률이 커지기 때문이다.

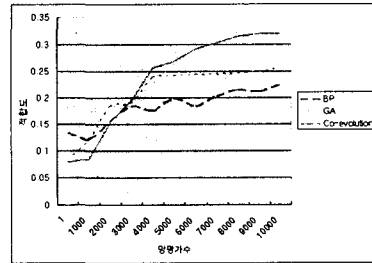


그림 4. GLASS 데이터에 대한 세 가지 학습 모델의 적합도 탐색

5. 결론 및 향후 연구 과제

진화 인공 신경망은 역전과 알고리즘과 같은 전통적인 학습 기술 대신에 신경망의 진화를 통해 최적화된 신경망 구조를 탐색하기 위해 연구되어 왔다. 경쟁 공진화는 자연계의 진화 메커니즘에서 서로 다른 종(種)간의 상호 영향을 기반으로 이루어지는 진화로써 실세계의 자연에서 쉽게 관찰되는 현상이다. 본 논문에서 제안하고 있는 경쟁 공진화 신경망 알고리즘은 이러한 경쟁 공진화를 기반으로 삼고 있으며, 실제 데이터를 통한 실험에 의해 일반 유전자 알고리즘에 비해 우수한 적합도를 가지는 개체를 탐색할 수 있음을 알 수 있다. 경쟁 공진화 이론은 비단 이러한 신경망 학습에 제한되지 않고 진화 연산을 수행하는 대부분의 경우에 적용할 수 있을 뿐 아니라 자연계의 진화 메커니즘의 보다 사실적인 모방을 위한 모델로 사용될 수 있을 것이다.

감사의 글

본 연구는 과학 기술부 주관 뇌신경 정보학 사업에 의해 지원되었음.

참고 문헌

[1] Xin Yao, "Evolving Artificial Neural Networks", proceeding of the IEEE, Vol. 87, No. 9, pp.1423-1439, September 1999
 [2] David. J. Montana, Lawrence Davis, "Training Feedforward Neural Network Using Genetic Algorithm", Proceedings of the International Conference on Artificial Intelligence, pp. 762-767, 1989
 [3] Van Rooij, A.J.F., Jain, L.C. and Johnson, R.P., "Neural Network Training Using Genetic Algorithms", Guidance, Control and Fuzing Technology International Meeting, April 1996.
 [4] Whitley, D., Starkweather, T. and Bogart, C., "Genetic Algorithms and Neural Networks: Optimizing Connection and Connectivity", Parallel Computing, Vol. 14, pp. 347-361, 1990
 [5] Wright, A. H., "Genetic Algorithms for Real Parameter Optimization", Foundation of Genetic Algorithms, Morgan Kaufmann Publishers, 205-218, 1991
 [6] <http://www.ics.uci.edu/~mllearn/MLRepository.html>