

개선된 코호넨 신경회로망 학습 알고리즘을 이용한 송전선 고장 종류 분류

조흥식 민상원 장용진 박종근
서울대학교 전기공학부

김광원
울산대학교 전기전자정보시스템공학부

Fault Type Classification using Improved KNN Learning Algorithm

Hong-Shik Cho Sang-Won Min Yong-Jin Jang Jong-Keun Park Gwang-Won Kim
Seoul National University University of Ulsan

Abstract - 송전선에 사고가 발생하면 계전기가 동작하고 여러 가지 후속 조치가 적절히 수행되기 위해서는 빠른 시간 내에 고장의 검출, 고장 종류의 분류, 고장 위치의 판정 등이 요구된다. 신경회로망을 이용한 고장 해석은 빠르고 정확한 해석이 가능하기 때문에 최근에 각광을 받고 있다. 본 논문에서는 기존의 코호넨 신경회로망의 지도 경쟁 학습 알고리즘을 개선한 새로운 학습 알고리즘이 제시된다. 제안된 학습 알고리즘은 전력계통의 송전선 고장 종류 분류에 적용된다.

개선된 알고리즘이 본 논문에서 제시된다. 제시된 알고리즘의 효율성을 검증하기 위하여 사례 연구로 154 [kV]급 송전계통에서의 고장 종류 분류가 수행된다.

1. 서 론

송전선에 고장이 발생하면 전압과 전류 신호에 과도 직류 성분과 고조파 성분이 기본주파수 (60Hz) 성분은 추가되어 정확하고 신속한 사고의 검출, 고장 종류의 분류, 고장 위치의 판정 등을 어렵게 한다. 최근에 고장 진단 분야에 인공 신경회로망 (Artificial Neural Network: ANN)을 적용하는 연구가 많이 진행되어 긍정적인 결과를 얻어내었으며 기존의 방법들에 비하여 처리 속도에 있어서 우월함을 보였다. 신경회로망은 학습 능력에 의한 임의의 비선형 함수의 근사화와 병렬 처리에 의한 계산 속도에 있어서 뛰어난 성능을 갖고 있다. 또한 노이즈로 인하여 변조된 입력에 대해서조차 출력을 정확하게 이끌어낼 수 있는 고유의 능력을 갖고 있기 때문에 측정 노이즈가 심각한 문제가 될 수 있는 상황에서 이상적으로 사용될 수 있는 방법이다. 이러한 장점으로 신경회로망을 바탕으로 한 패턴 인식 방법은 온라인 전력계통 고장진단에 있어서 가장 전망있는 대체 수단으로 생각할 수 있으며 많은 연구가 진행되고 있다. Dalstein 등은 역전파 알고리즘 (Back-Propagation algorithm)을 이용하여 학습한 다층퍼셉트론 (Multi-layer perceptron)을 고속 보호 계전을 위한 고장 종류 분류와 고장 지역 추정에 적용하였다 [1,2]. Narendra 등은 HVDC 계통에서의 고장 진단에 Radial Basis Function (RBF) 신경회로망을 적용하였다 [3]. Aggarwal 등은 결합 지도/비지도 신경회로망 (Combined Supervised/Unsupervised Neural Network) 을 이용하여 송전선의 고장 종류를 분류하였다 [4].

본 논문에서는 코호넨 신경회로망 (KNN: Kohonen Neural Network)을 이용한 송전선의 고장 종류 분류 방법이 제시된다. 그 학습 알고리즘인 지도경쟁학습법은 다층퍼셉트론의 역전파 알고리즘에 비해 이해하기 쉽고 학습시간이 짧으며 그 성능을 조절하기 쉬운 장점을 갖고 있다. 기존의 지도경쟁학습법인 학습 양자화 (LVQ) 알고리즘은 기존의 분류기들에 비해 좋은 성능을 보이지만 [5,6], 고장 종류 분류와 같이 클래스와 클래스의 구분이 판단경계면으로 분명히 나뉘어지는 문제에 있어서는 그 참조벡터의 효과적인 이용 측면에 있어서 결점을 보인다. 그러한 문제점을 해결하기 위하여 본 논문의 저자들에 의해 경계면탐색법이 개발되었으며 [7], 그러한 경계면탐색법의 분류 성능을 더욱 향상시킬 수 있는

2. 개선된 코호넨 신경회로망 알고리즘

2.1 학습벡터양자화 알고리즘 (LVQ)

기존의 대표적인 코호넨 신경회로망의 지도학습 방법으로 학습 벡터 양자화 (LVQ: Learning Vector Quantization) 알고리즘이 있다 [5,6]. LVQ 알고리즘은 지도 학습에 근거한 패턴인식기의 일종으로, 학습과 인식의 두 부분으로 이루어져 있다. 학습부에서는 소속 클래스를 알고 있는 학습 패턴 벡터를 양자화하여 클래스별로 표준패턴(이를 참조벡터 (reference vector)라고 함)을 구성하고, 인식부에서는 최소거리 (nearest neighbor) 규칙으로 입력된 미지패턴에 가장 가까운 참조벡터를 탐색하여 참조벡터의 클래스를 미지 패턴에 할당한다.

LVQ 알고리즘은 학습 벡터들의 확률적 분포를 묘사하여 참조벡터들을 입력 패턴들의 공간에 분포시키는 방법으로서, parametric Bayes classifier나 k-nearest-neighbor classifier와 같은 기존의 분류기에 비하여 더 좋은 성능을 보인다. 그렇지만, 클래스와 클래스의 구분은 판단 경계면 (Decision Boundary)으로 나뉘어지는 분류 문제에 있어서는 문제점을 갖고 있다. LVQ 알고리즘의 학습 과정에서는 주로 판단 경계면 근방의 참조벡터들이 선정되어 학습 받기 때문에 판단 경계면에서 먼 참조벡터들은 분류에 이용되지 못하고 불필요하게 되는 것이다. 이는 같은 규모의 코호넨 신경회로망의 성능을 저하시키는 큰 요소가 된다.

그림 1 (b)의 예에서 보여지는 것처럼 LVQ 알고리즘에 의한 참조벡터는 대체적으로 입력 패턴 공간에 고르게 분포되게 된다. 코호넨 신경회로망은 유클리디언 거리 (Euclidean distance)을 기초로 하여 코호넨층의 노드들간의 경쟁의 승리자 (winner)를 결정하기 때문에 판단 경계면에서 상대적으로 먼 참조벡터들을 포함하지 않더라도 같은 성능을 보일 것이다. 이러한 모든 참조벡터를 분류에 제대로 활용하지 못한다는 점이 판단 경계면에 의하여 분명하게 나뉘어 지는 분류 문제에 적용될 때 LVQ 알고리즘의 단점이 되는 것이다.

2.2 경계면탐색법 (BSA)

LVQ 알고리즘의 그러한 문제점을 해결하기 위하여 클래스 간의 분류에 있어서 중요한 판단 경계면 주위에서의 분류에 더 효율적으로 대처하고 코호넨 신경회로망의 참조벡터로 판단 경계면에 가까운 학습 데이터를 효과적으로 이용하는 학습 알고리즘인 경계면탐색법 (Boundary Search Algorithm : BSA)을 제안하였다 [7]. LVQ 알고리즘에서 나타나는 불필요한 참조벡터들을 제거하고, 모든 참조벡터들을 판단 경계면 근방에 분포시켜 분류에 활용시킨다면 같은 수의 참조벡터를 가지

고 더 우수한 성능의 코호넨 신경회로망 분류기를 구현할 수 있을 것이다.

경계면탐색법은 기존의 LVQ 알고리즘의 불필요한 참조벡터를 남기는 문제점을 개선하고 모든 참조벡터를 클래스간의 분류에 중요한 지점인 경계면 근방에 배치함으로써 분류기로서의 성능을 더욱 향상시킬 수 있었다.

그러나 그 알고리즘은 주어진 학습데이터 중에서 적당한 참조벡터를 선택하는 것이기 때문에 분류에 있어서 더 필요한 지점이 있다 하더라도 그 지점이 학습 데이터 집합에 포함되어 있지 않으면 주어진 학습 데이터로 얻어지는 참조벡터만으로는 최적의 분류기를 구현하는 것은 불가능하다. 학습 데이터 집합이 충분히 많이 그리고 고밀도 데이터 공간에서 골고루 채득된다면 그러한 문제점은 그리 크게 부각되진 않겠지만, 그것을 만족할만큼의 많은 학습 데이터를 얻기에는 다소 어려움이 따른다.

2.3 개선된 경계면탐색법

본 논문에서 제시하는 경계면탐색법 기반 학습 벡터 양자화 알고리즘 (BSA Based Learning Vector Quantization: BLVQ)은 경계면탐색법에 LVQ 알고리즘의 미세 조정 (Fine-Tuning) 과정을 결합함으로써 경계면탐색법에서 나타나는 문제점을 해결할 수 있다. 경계면탐색법의 경계면을 중심으로 하는 참조벡터의 획득과 LVQ 알고리즘의 최적의 지점을 찾아 참조벡터를 미세 조정할 수 있는 과정을 결합함으로써 두 알고리즘의 장점을 모두 가지고 있다고 할 수 있다.

BLVQ는 두 단계 학습과정으로 구성되며 각 학습과정이 순차적으로 수행된다. 단계 1에서는 주어진 학습 데이터 중에서 판단 경계면에 근접해 있다고 추정되는 N개의 데이터를 코호넨 신경회로망의 참조벡터로 취하는 과정이다. 단계 2에서는 분류 성능을 향상시키기 위한 참조벡터의 미세 조정이 이루어져 단계1의 결과를 개선한다.

2.3.1 단계 1: 경계면 탐색 단계

단계 1의 목적은 학습 데이터 중에서, 미리 결정된 참조벡터 수만큼의 경계면 근방에 있는 데이터를 찾아 신경회로망의 참조벡터로 지정하는 것이다. 주어진 학습 데이터 중에서 경계면에 위치한 데이터는 다른 클래스의 데이터와 근접해 있다는 사실로부터 어떠한 데이터가 경계면에 근접해 있다는 판정 기준을 다른 클래스에 속한 데이터와의 최단 거리로 정하였다.

경계면에 위치한 N개의 데이터를 구하는 과정은 다음과 같다.

- 절차1-1: 학습데이터들을 학습데이터 집합에 옮겨놓고 임의적으로 정렬한다
- 절차1-2: 처음 N개의 학습데이터를 코호넨 신경회로망의 참조벡터의 초기값으로 지정한다. 여기서 N은 미리 정해진 참조벡터의 수이다.
- 절차1-3: 식 (1)을 이용하여 각 참조벡터에 대해서 DO_i 를 계산한다.

$$DO_i = \min_j (\| R_i - R_j \|), \quad 1 \leq i \leq N \quad (1)$$

여기서 DO_i 는 i번째 참조벡터 R_i 와 다른 클래스에 속하는 다른 참조벡터들과의 거리 중 최소치이며, j는 i번째 참조벡터와 다른 클래스에 속하는 참조벡터의 인덱스이다.

- 절차1-4: 학습데이터 집합에서 다음의 학습데이터를 받아들인다.
- 절차1-5: 입력된 학습데이터를 N+1번째 참조벡터로 지정한다.
- 절차1-6: DO_i 를 계산하고, 현재의 학습 데이터와 다른 클래스에 속하는 참조벡터들에 대해서 그 DO_j 를 식 (2)를 이용하여 바꾼다.

$$DO_i = \min (DO_i, \| R_i - R_{N+1} \|) \quad (2)$$

- 절차1-7: N+1개의 참조벡터 중 가장 큰 DO_i 를 가지고 있는 참조벡터를 찾는다.
- 절차1-8: 그 참조벡터가 N+1번째 참조벡터이면 절차 1-9로 진행하고, 그렇지 않으면 그 참조벡터와 그에 해당하는 DO_i 를 N+1번째 참조벡터와 교체한다.
- 절차1-9: 절차 1-4로 진행하며, 모든 학습 데이터에 대하여 계속 수행한다.

단계 1에서는 작은 DO_i 를 가지고 있는 학습데이터를 판단 경계면 근방에 있다고 추정하고, 학습 데이터 집합에서 미리 지정된 수만큼의 작은 DO_i 를 가진 데이터를 찾아 참조벡터로 지정한다.

2.3.2 단계 2: 미세 조정 단계

단계 2에서는 단계 1의 결과로 얻어진 참조벡터 집합을 대상으로 학습벡터양자화 2 알고리즘 (LVQ2)을 이용하여 참조벡터를 다른 입력 데이터의 분류에 더 좋은 위치로 미세 조정 (Fine-Tuning) 시킨다.

주어지는 학습 데이터의 분류가 기존의 참조벡터 집합으로 오분류가 될 경우 다음과 같은 과정으로 참조벡터들을 개선한다.

- 절차2-1: 학습 패턴 집합의 각 패턴 x_j 에 대하여 다음을 수행한다:
- 절차2-2: 패턴 x_j 에 가장 근접한 코드북 벡터 R_i 를 선정하여 $R_c = R_i$ 로 한다;
- 절차2-3: 만약 x_j 의 클래스와 R_c 의 클래스가 다르면 (즉, 기존의 코드북 벡터 집합으로 주어진 학습 데이터의 분류가 잘못 되면), x_j 와 두번째로 근접한 코드북 벡터 R_i 를 선정하여 $R_c = R_i$ 로 한다;
- 절차2-4: x_j 의 클래스와 R_c 의 클래스가 같으면, 다음과 같은 식으로 R_c 와 R_c^* 를 갱신한다;

$$R_{c^*}(k+1) = R_{c^*}(k) + a(k)[x_j - R_{c^*}(k)] \quad (3)$$

$$R_c(k+1) = R_c(k) + a(k)[x_j - R_c(k)] \quad (4)$$

- 절차2-5: 그 외의 코드북벡터는 갱신하지 않는다.
- $$R_i(k+1) = R_i(k) \text{ for } i \neq c, c^* \quad (5)$$

2.3 제안된 알고리즘의 비교

앞에서 제시한 세가지 알고리즘을 비교 분석하기 위하여 그림 1(a)와 같은 다소 복잡한 2차원 입력 모델의 클래스 분류에 기존의 학습 벡터 양자화 알고리즘과 제안된 알고리즘을 적용해보았다. 학습 벡터 양자화 알고리즘 (LVQ1, LVQ2), 경계면 탐색법 (BSA), 경계면 탐색법 기반 학습 벡터 양자화 알고리즘 (BLVQ)에 대하여 코호넨 신경회로망의 참조벡터의 수를 200, 300, 400개로 바꿔 가며 학습데이터 5000개에 대하여 학습하였고, 학습에 사용되지 않은 새로운 5000개의 데이터로 그 성능을 평가하였다. LVQ 알고리즘은 참조벡터의 초기값의 설정에 따라 그 성능의 차이를 보이기 때문에 각각 10번의 실험의 평균값으로 하였다.

표 1은 각 알고리즘의 오분류율 (Misclassification Rate)이며 그림 1은 입력 공간에서의 참조벡터의 분포를 나타낸다.

표 1. 2차 입력 모델에 대한 오분류율 (%)

참조벡터 수	LVQ1	LVQ2	BSA	BLVQ
200	3.52	3.60	2.74	1.70
300	3.55	3.48	1.86	1.44
400	3.45	3.35	1.86	1.68

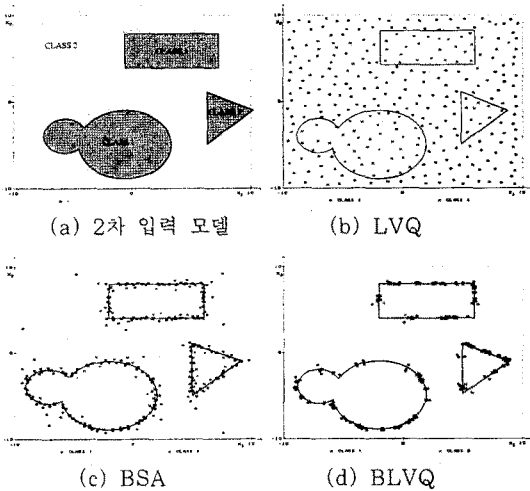


그림 1. 2차 입력 모델에서의 참조벡터의 분포

3. 송전선 고장 종류 분류

본 논문에서는 Electromagnetic Transients Program (EMTP)를 이용하여 송전선에서의 사고를 모의하였다. 그림 2와 같은 154[kV]급 송전 계통에 길이가 80[km]인 송전선에서 다음과 같은 세개의 매개변수를 변화시켜 가며 10가지 유형 (a-g, b-g, c-g, a-b, b-c, c-a, a-b-g, b-c-g, c-a-g, a-b-c)의 사고가 모의되었다.

- (a) 고장시 전압 위상각 (0, 45, 90°)
- (b) 사고 지점 (10, 30, 50, 79, 90%)
- (c) 고장 저항 (0, 5, 10, 20 Ω)

신경회로망의 학습과 성능평가에 사용될 고장 패턴을 구하기 위하여 다양한 EMTP 모의가 이루어져 각각 다른 상황에 대하여 학습데이터로 500개의 사고 패턴과 검증데이터로 100개의 사고패턴을 얻었다.

신경회로망의 입력 패턴으로 사용될 특징변수로는 다음 두가지 방법을 이용하였다.

- (case 1) 3상 전압의 기본파 성분, 3상 전류의 기본파 성분, 중성전류의 기본파성분 (입력차수 : 7)
- (case 2) 3상 전압, 전류, 중성전류의 6 샘플링 시간 동안의 연속값 (입력차수: 42)

두가지 경우의 입력 패턴에 대하여 BSA와 BLVQ로 코호넨 신경회로망을 학습하여 고장 종류를 분류하였으며, 그에 따른 오분류율(Misclassification Rate)은 표 2와 같다. 표에서 보는 바와 같이 BLVQ는 기존에 제시한 BSA 보다 우월한 성능을 보인다. 그 차이는 참조벡터의 수가 크지 않을 때 더 크게 나타난다. case 1과 같은 입력의 특징 변수는 작은 수의 참조 벡터로 유형의 분류가 가능하지만, 전압 전류의 기본파 성분을 구하기 위해서는 DFT를 비롯한 전처리 과정 (pre-processing)이 필요하다. case 2와 같은 입력의 특징 변수는 정확한 분류를 위해서는 많은 수의 참조 벡터를

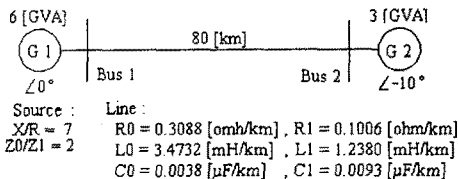


그림 2. 154kV 송전 계통

필요로 하지만, 전압, 전류의 측정값만으로 분류가 가능하기 때문에 실시간 적용이 가능하다.

표 2. 고장 종류 분류에 대한 오분류율 (%)

	참조벡터 수	BSA	BLVQ
case 1	51	6.3	1.0
	77	2.0	0.7
case 2	1300	3.06	0.98
	1650	2.12	0.88

4. 결 론

본 논문은 코호넨 신경회로망을 이용하여 전력계통 송전선에서의 고장 분류를 효과적으로 수행하였다. 코호넨 신경회로망의 효과적인 학습을 위하여 경계면탐색법 기반 학습벡터양자화알고리즘 (BLVQ)이 개발되었으며, 이는 기존의 학습 알고리즘인 학습벡터양자화알고리즘 (LVQ)과 경계면탐색법 (BSA)의 단점을 해결하여 분류기로서의 성능을 향상시킨 것이다. 제안된 알고리즘의 타당성은 154kV급 송전 계통의 고장 종류 분류에서 입증되었다. 고장 분류에 있어서 신경회로망의 빠른 계산 능력에 정확도를 향상시킴으로써 고속 보호계전 시스템이나 변전소 모니터링 시스템에서의 자동사고분석에 적용할 수 있는 가능성을 높였다.

(참 고 문 헌)

- [1] Thomas Dalstein and Bernd Kulicke, "Neural Network Approach to Fault Classification for High Speed Protective Relaying", IEEE Trans. on Power Delivery, Vol.10, No.2, pp.1002-1011, 1995
- [2] Thomas Dalstein, Thomas Friedrich, Bernd Kulicke, and Dejan Sobajik, "Multi Neural Network Based Fault Area Estimation for High Speed Protective Relaying", IEEE Trans. on Power Delivery, Vol.11, No.2, pp.740-747, 1996
- [3] K. G. Narendra, V. K. Sood, K. Khorasani, and R. Patel, "Application of a Radial Basis Function (RBF) Neural Network for Fault Diagnosis in a HVDC System", IEEE Trans. on Power Systems, Vol.13, No.1, pp. 177-183, 1998
- [4] R. K. Aggarwal, Q. Y. Xuan, R. W. Dunn, A. T. Johns, and A. Bennett, "A Novel Fault Classification Technique for Double-circuit Lines Based on a Combined Unsupervised/Supervised Neural Network", IEEE Trans. on Power Delivery, Vol.14, No.4, pp. 1250-1256, 1999
- [5] T. Kohonen, Self-Organization and Associative Memory, 2nd ed., Berlin : Springer-Verlag, 1988.
- [6] J. A. Kangas, T. Kohonen, and J. T. Laaksonen, "Variants of Self-Organizing Maps," IEEE Trans. on Neural Networks, Vol. 1, No. 1, pp. 93-99, 1990.
- [7] Y. M. Park, G. W. Kim, H. S. Cho, and K. Y. Lee, "A New Algorithm for Kohonen Layer Learning with Application to Power System Stability Analysis", IEEE Trans. on System, Man, and Cybernetics - Part B, Vol. 27, No. 6, pp. 1030-1034, 1997