# 이진신경회로망에서 MSP Term Grouping 알고리즘의 Time Complexity 분석

# Time Complexity Analysis of MSP Term Grouping Algorithm for Binary Neural Networks

박 병 준, 이 정 훈
Computational Vision and Fuzzy Systems Laboratory
한양대학교 전자컴퓨터공학부
425-791 경기도 안산시 사1동 1271번지
{bpark, frhee}@fuzzy.hanyang.ac.kr

**Abstract**

본 논문은 Threshold Logic Unit (TLU)를 기본 뉴런으로 하여 최소화된 이진신경회로망을 합성하는 방법인 MSP Term Grouping(MTG) 알고리즘의 time complexity를 분석하고자 한다. 이를 전체 패턴 탐색을 통한 이진신경회로망 합성의 경우와 비교하여 MTG 알고리즘의 효용성을 보여준다.

## I. 서론

초창기 신경회로망의 단위 뉴런이 되었던 Threshold Logic Unit은 Sigmoid와 같은 미분 가능한 함수를 활성화함수(activation function)로 사용하지 않고 hard limit 함수를 사용한다[1]. 따라서 Threshold Logic Unit을 단위 뉴런으로 하는 신경회로망은 역전파 알고리즘(Error Back propagation Algorithm)처럼 일반화된 학습 알고리즘이 불가능하고 몇몇 합성 알고리즘이 존재한다[2][3][4]. 그러나 VLSI 상에서의 하드웨어 구현을 고려한다면 sigmoid 함수보다는 hard limit 함수를 활성화 함수로 하였을 때 보다 실제에 가까운 뉴런 모델이 될 수 있다[5]. TLU를 사용한 신경회로망을 합성하는 알고리즘으로는 BLTA[2]와 ETL[3] 방법이 나와 있으며, 이보다 향상된 방법으로 MSP Term Grouping Algorithm[4] 이 고안되었다.

MSP Term Grouping 알고리즘은 학습 패턴을 이진 함수로 취급하여, 이 함수를 디지털 논리설계 기법으로 사용되는 Minimal Sum of Products (MSP)형태로 변환하였을 때 만들어진 항을 그룹화하는 과정을 통해 신경회로망을 합성하는 방법이다. 이는 다른 기하학적인 접근방법과는 다르게 전체 이진함수를 최적화된 조건에서 TLU로 만들 수 있는지 검사할 수 있는 최소화된 이진 신경회로망 합성 알고리즘이다[4][6].

본 논문에서는 이런 MSP Term Group 알고리즘과 전체 탐색을 이용한 방법의 time complexity 분석을 통해, MSP Term Group 알고리즘의 효용성을 보이고자 한다.

## II. 본론

### A. Binary Neural Network using TLUs

Threshold Logic Unit 한 개는 $N$차원 2진 hypercube를 두 개의 영역으로 분할하는 hyperplane을 형성한다. 이 TLU를 그림으로 표현하면 그림1과 같게 되고, 이 TLU의 입출력 관계식은 식(1)이 된다.

$$\begin{aligned} y=1 &\quad \text{if and only if} \quad \textstyle\sum_{i=1}^{N} w_i x_i - T \geq 0 \\ y=0 &\quad \text{if and only if} \quad \textstyle\sum_{i=1}^{N} w_i x_i - T < 0 \end{aligned} \qquad (1)$$

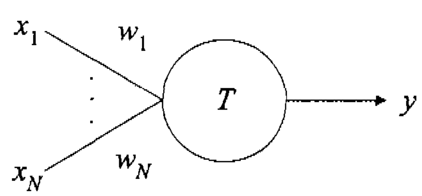


그림1. TLU의 기본구조

그러므로 두 영역을 분할하는 hyperplane의 식은 다음과 같으며,

$$\sum_{i=1}^{N} w_i x_i - T = 0 \qquad (2)$$

이 hyperplane을 만족하는 가중치(weight) $w_i$와 문턱값(threshold) $T$를 찾는 것이 TLU를 학습시키는 과정이다.

그러나 일반적인 패턴 분류 문제에서 여러 개의 Threshold Logic Unit을 사용하여야 하며, TLU의 적절한 개수와 인수를 결정하는 것은 매우 어렵다.

주어진 2진 함수로부터 한 개의 TLU의 가중치와 문턱값을 구해내는 방법에는 몇 가지가 있다[6]. MSP Term Grouping알고리즘에서는 이 중에서 RHS알고리즘을 사용하여, 한 개의 TLU로 표현이 가능한지 여부를 확인하고 해당 TLU의 가중치 및 문턱값을 구한다[7]. TLU의 자세한 특성과 RHS의 과정을 살펴보려면, [6]을 참조하면 된다.

## B. MSP Term Grouping Algorithm

MSP Term Grouping 알고리즘은 2진 패턴 함수를 먼저 MSP형태로 바꾼 다음 TLU의 특성을 이용하여 MSP항을 그룹화 함으로써 최소화된 2진 신경회로망을 얻어낸다[4][6].

2진 패턴 분류문제를 MSP형태로 표현할 때 얻어지는 장점으로는 패턴의 그룹화 시도를 위한 초기 그룹형성을 해주는 것과 동시에 unate 특성을 확인할 수 있는 방법을 제공해 준다는 것이다. Unate특성은 2진 함수를 MSP형태로 표현하였을 때, 하나의 변수라도 complement와 uncomplement 형태가 동시에 나온다면 TLU하나로 표현할 수 없는 특성이다[1][6]. 그러므로 이 unate특성을 만족하지 못하는 항들은 그룹화를 시도할 필요가 없다[4][6].

MSP Term Grouping 알고리즘에서는 가장 먼저 전체 MSP항들을 각각 한 개의 TLU를 이용하여 표현하도록 한다. 그리고 나서 이 TLU들의 집합에서 TLU의 쌍을 그룹화하여 하나의 TLU로 묶는 과정을 반복한다. 이 때 그룹화의 우선순위를 정하여 주는 $d$ 함수를 고안하였으며, 그룹화 가능여부를 확인하고, 새로운 TLU를 생성하기 위해 RHS 알고리즘을 이용하였다.

$N$차원의 이진 함수에서 $A$와 $B$의 두개의 MSP Term 그룹이 있을 때, 그 그룹의 변수들의 집합을 각각 $A'$과 $B'$이라고 하면, $d$함수는 다음과 같이 정의 된다.

$$d(A,B) = N \cdot \min\{|A'-B'|, |B'-A'|\} + \max\{|A'-B'|, |B'-A'|\} \qquad (3)$$

이 $d$함수의 값이 작으면 하나의 TLU로 묶을 수 있는 충분조건에 해당된다. $d$함수에 대한 자세한 내용과 알고리즘의 세부적인 사항은 [4]나 [6]을 참조하면 된다.

일단 한 번 그룹화에 실패한 TLU쌍은 다시 그룹화 시도를 반복하지 않도록 표시하여 둔다. 더 이상 그룹화 할 수 있는 TLU의 쌍이 없을 때까지 수행하면 bottom-up 방식의 최소화된 2진 신경회로망을 얻을 수 있다. 다음은 전체 알고리즘을 나타낸 것이다.

```
MSP term Grouping Algorithm
Obtain the initial LTU set where its elements are
terms of the MSP function;
REPEAT
   FOR all existing set element pairs that are unate
   and available for grouping DO
       Obtain pair that has the smallest d value;
   END FOR
   Apply the RHS method to the selected pair;
   (i.e., to check if pair is linear separable)
   IF the pair is linearly separable THEN
       Partition the LTU set which include the pair;
       (i.e., the partition becomes an element of the
       LTU set)
   ELSE
       Set pair is not available for grouping;
ENDIF
UNTIL(there exists no available pairs for
grouping)
FOR all LTU set elements DO
   Obtain a linear threshold unit using the RHS
   method;
END FOR
```

## C. Time Complexity Analysis

MSP Term Grouping 알고리즘에서는 먼저 주어진 이진 패턴을 MSP 형태로 만들어 놓고, 각 MSP 형태의 항을 TLU 하나로 표시한다. 이 TLU들을 서로 합쳐가는 과정이 MSP Term Grouping 알고리즘이다. 즉 Bottom-Up 방향의 접근 방식이다.

먼저 $f = t_1 + \cdots + t_m$ 과 같이 $m$ 개의 항을 갖는 MSP형태의 함수가 있다고 가정해 보자.

$n$차원의 이진 문제의 경우에 $m$의 최대값은 $2^{n-1}$이 된다. 이는 그림2와 같은 예제를 통해 이해할 수 있는데, 2진 함수의 MSP항이 가장 많은 경우는 패리티를 구하는 경우이며, 이 때는 2진 함수의 최소항(minterm)이 모두 MSP의 항이 된다. 여기서는 이 사항을 염두에 두고 $m$을 사용하여 식을 전개한 후에 값을 대입하도록 한다.

M개의 MSP 항에 대해서 서로 묶음을 시도해 보는 모든 항의 쌍은 $\binom{m}{2}$ 이다. 즉, $m$개의 항에 대해 그룹화를 시도하는 회수를 $T_m$ 이라

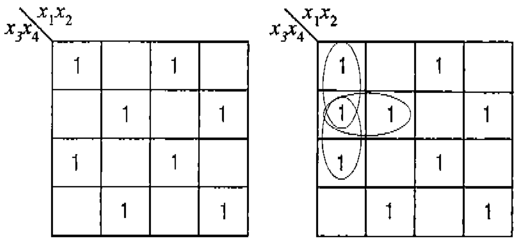


그림 2. K-map으로 표현한 4-dim의 경우에 대한 최대 MSP항. 최소항의 수를 증가시켜도 MSP항의 수는 증가하지 않는다

고 하면

$$T_m \le \binom{m}{2} \qquad (4)$$

가 되고, 일단 그룹화에 실패한 항의 쌍은 다시 시도할 필요가 없다.

이 중에 한 번의 시도가 성공하게 되면 이제 항의 수는 두 개가 합쳐져서 하나로 되었으므로 $m$-1로 바뀐다. $m$-1개의 항 중에서 바로 앞 단계에서 합쳐진 1개의 항을 제외한 $m$-2 개의 항에 대해서는 이미 $m$개의 항이 있던 단계에서 그룹화 시도를 하였다가 실패한 것들이므로 이 단계에서 또 다시 그룹화를 시도할 필요가 없다. 결국 $T_{m-1}$ 은

$$T_{m-1} \le \binom{m-1}{2} - \binom{m-2}{2} \qquad (5)$$

가 된다. 같은 원리로

$$T_{m-2} \le \binom{m-2}{2} - \binom{m-3}{2} \qquad (6)$$

가 되고, 이를 일반화하여 쓰면

$$T_{m-i} \le \binom{m-i}{2} - \binom{m-i-1}{2} \; (0 < i < m-1 \text{인 정수}) \qquad (7)$$

와 같이 표현할 수 있다.

이 때, 항의 수가 1개 일 때는 그룹화를 시도할 필요가 없으며, 2개 일 때는 1회만 그룹화를 시도하면 된다는 것을 직관적으로 알 수 있다. 그러므로 $T_1$ 은 0 이고, $T_2$는 1 임을 알 수 있다.

결국 $m$ 개의 MSP항을 갖는 이진 함수의 그룹화 시도 회수는

$$\begin{aligned} T_{total} = \sum_{i=1}^{m} T_i &\le \binom{m}{2}_m C_2 + \sum_{i=3}^{m-1} \left\{ \binom{i}{2} - \binom{i-1}{2} \right\} + 1 \\ &= \binom{m}{2} + \left\{ \binom{m-1}{2} - \binom{m-2}{2} \right\} + \left\{ \binom{m-2}{2} - \binom{m-3}{2} \right\} \\ &\quad + \Lambda + \left\{ \binom{3}{2} - \binom{2}{2} \right\} + 1 \\ &= \binom{m}{2} + \binom{m-1}{2} \end{aligned} \qquad (8)$$

$$\therefore T_{total} \le \binom{m}{2} + \binom{m-1}{2}$$

가 된다.

이 식을 전개하면

$$\begin{aligned} T_{total} &\le \frac{m(m-1)}{2} + \frac{(m-1)(m-2)}{2} \\ &= \frac{m^2-m}{2} + \frac{m^2-3m+2}{2} \\ &= m^2 - 2m + 1 \\ &= (m-1)^2 \end{aligned} \qquad (9)$$

와 같아서 MSP항의 개수 $m$에 대해 $O(m^2)$이 된다.

앞에서 기술한 바와 같이 $n$차원 이진공간을 표시하는 데 필요한 MSP항의 수의 최대값은 $2^{n-1}$ 이 되고, 결국 $n$차원 이진공간을 분할하는 신경회로망을 합성하기 위해 필요한 time complexity의 최대값은 $O(4^{n-1})$의 한계를 가진다.

그러나 패리티 문제를 제외한 대부분의 경우에 MSP항의 개수는 $2^{n-1}$개보다 상당히 적으므로 실질적인 time complexity는 $O(4^{n-1})$ 보다 낮다.

## III. 결론

본 논문에서는 TLU를 이용한 2진 신경회로망을 합성하는 알고리즘 중에서 최소화된 신경회로망을 합성해주는 MSP Term Grouping 알고리즘의 합성과정에 대한 time complexity 분석을 수행하였다. 그 결과로 MSP Term Grouping 알고리즘은 패턴 전체 탐색을 통한 방법과 같은 수준으로 최소화된 2진 신경회로망을 합성하여 주면서 그 탐색 시간을 $O(m)=m^2$으로 줄였음을 확인하였다.

## IV. 부록

Top-down방식 전체탐색 time-complexity

Top-down방식으로 그룹화를 시도할 경우에는 처음 모든 MSP항을 한 개의 그룹으로 시도를 한 후, 실패할 때마다 그룹을 한 개씩 늘려 나간다.

한 그룹을 시도할 때의 시도 회수는

$$A_1 = \binom{m}{0} \qquad (A1)$$

이고, 두 개의 그룹이 되도록 시도할 경우에는 1개와 나머지, 두개와 나머지, 의 순서로 시도해 보면

$$A_2 \le \binom{m}{1} + \binom{m}{2} + \Lambda + \binom{m}{m/2} \qquad (A2)$$

가 된다. 같은 방법으로 세 개의 그룹이 되도록 시도할 경우에는 1개를 제외한 나머지로

두 그룹을 시도해 보고, 두개를 제외한 나머지로 두 그룹을 시도해 보는 식으로 진행이 되어 재귀적으로 시도 회수가 증가하게 된다.

$$A_3 \le \left[\binom{m}{1} + \left\{\binom{m-1}{1} + \binom{m-1}{2} + \Lambda + \binom{m-1}{\lfloor \frac{m-1}{2} \rfloor}\right\}\right]$$
$$+ \left[\binom{m}{2} + \left\{\binom{m-2}{1} + \binom{m-2}{2} + \Lambda + \binom{m-2}{\lfloor \frac{m-2}{2} \rfloor}\right\}\right] \quad \text{(A3)}$$
$$+ \Lambda$$

이 중에서 만약 두 개의 그룹으로 시도하는 경우인 $A_2$만을 고려해 보더라도

$$\begin{aligned} A_2 &\le \sum_{i=1}^{\lfloor m/2 \rfloor} \binom{m}{i} \\ &= \frac{m}{1!} + \frac{m(m-1)}{2!} + \Lambda \qquad \text{(A4)} \\ &\quad + \frac{m(m-1)\Lambda\,(m - \lfloor m/2 - 1 \rfloor)}{\lfloor m/2 \rfloor} \\ &\approx \sum_{i=1}^{\lfloor m/2 \rfloor} \frac{m^i}{i!} \qquad (\text{for large } m) \end{aligned}$$

가 되는데 이것을 테일러급수를 이용하여 풀면

$$e^{x+m} \approx \sum_{i=0}^{\lfloor m/2 \rfloor} \frac{m^i}{i!} e^x \qquad \text{(A5)}$$
$$\therefore e^m \approx 1 + \sum_{i=1}^{\lfloor m/2 \rfloor} \frac{m^i}{i!}$$

와 같게 되며 결과적으로

$$A_2 \le e^m - 1 \qquad \text{(A6)}$$

이 되어 $O(e^m)$의 time-complexity를 가지게 된다. 이것은 단지 두 그룹을 시도하였을 때만을 표시한 것이며, top-down방식의 그룹화 접근은 실제로 더욱 큰 time complexity를 가지는 NP complete문제임을 알 수 있다.

## V. 참고문헌


[1] Z. Kohavi, *Switching and Finite Automata Theory*: 2nd ed., McGraw-HILL, 1970.

[2] D. Gray and A. Michel, "A training algorithm for binary feedforward neural networks," *IEEE Trans. Neural Networks,* vol. 3, no. 2, pp. 176-194, Mar. 1992.

[3] J. H. Kim and S. K. Park, "The geometrical learning of binary neural networks," *IEEE Trans. Neural Networks*, vol. 6, no. 1, pp.237-247, Jan. 1995.

[4] F. C. Rhee and B. Park, "An optimal method for linear threshold network synthesis," *Proc. of IJCNN'98*, Anchorage, Alaska, pp. 1905-1909, May. 4-9, 1998. as a part of IEEE WCCI'98

[5] J. H. Kim, S. K. Park, Y. Han, H. Oh, and M. S. Han, "Efficient VLSI implementation of a 3-layer threshold network," *Proc. ICNN'97*, Houston, TX, vol. 2, pp. 888-893, Jun. 1997.

[6] 박병준, "선형 신경회로망의 최적 합성 방법," 석사학위논문 한양대학교 대학원 1998.

[7] S. K. Park and J. H. Kim, "Geometrical learning algorithm for multilayer neural networks in a binary field," *IEEE Trans. Computers*, vol. 42, no. 8, pp. 988-992, Aug. 1993.