

신경망 최소-합 복호 알고리즘의 최적화에 대해

윤홍서*, 이채영*, 곽희열†, 김상호**†

On Optimization of Neural Min Sum Decoding Algorithm

Hong-Seo Yun*, Chae-Young Yi*, Hee-Yeoul Kwak†, Sang-Hyo Kim**†

*Department of Electronics and Electrical Engineering

**Department of Electrical and Computer Engineering

***Sungkyunkwan University, 16419 Suwon, Republic of Korea

†University of Ulsan, 44610 Ulsan, Republic of Korea

‡ Email : iamshkim@skku.edu

요 약

본 연구에서는 BCH (63, 36) 부호의 Neural MS 디코딩을 대상으로 Straight-Through Estimator(STE)와 Log-Likelihood Ratio(LLR) Scaling 기법을 적용하고, 각 기법이 신경망 디코딩의 수렴 효율성과 복호 성능에 미치는 영향을 정량적으로 비교하였다. 실험 결과, LLR scaling 기법은 신경망 기반 복호기의 수렴 속도를 개선하고 성능을 향상시켜, 보다 효율적인 neural decoding 설계 가능성을 보여주었다.

1. 서론

신경망 기반 반복 복호기 구조는 선형 블록 부호의 복호 성능을 향상시키기 위한 새로운 패러다임으로 자리잡고 있다. 특히 Belief Propagation(BP) 및 Min-Sum(MS)과 같은 기존 알고리즘에 학습 가능한 가중치를 도입한 Neural BP, Neural MS 구조는 계산량과 복호 성능 사이의 균형을 통해 다양한 응용 가능성을 확보하고 있다 [1,2].

그러나 NMS 구조는 반복 횟수가 적거나 네트워크 깊이가 얇은 상황에서 학습 안정성과 수렴 속도 측면에서 성능 저하를 겪는 경우가 있다. 본 연구에서는 이를 개선하고자, Neural MS에 STE와 LLR Scaling을 적용하고 그 성능을 정량적으로 비교하였다. 목적은 Neural MS의 계산 효율성은 유지하며 학습 효율과 복호 성능을 동시에 향상시키는 신경망 최적화 전략을 도출하는 데 있다.

2. 본론

2.1 BCH 부호

BCH 부호는 대수적 구조를 바탕으로 한 선형 블록 부호로, 고정된 부호 길이와 명확히 정의된 오류 정정 능력을 바탕으로 예측 가능한 성능을 제공하여 제어 채널이나 저장 장치 등에서 널리 사용된다. 일반적으로 이진 갈루아 필드 $GF(2^m)$ 위에서 정의되며, $n = 2^m - 1$ 형태의 코드 길이를 갖는다.

2.2 반복 기반 디코딩 알고리즘

BCH 코드에 적용될 수 있는 반복 기반 디코딩 알고리즘으로는 BP, MS, 그리고 해당 코드에 신경망 기반 학습이 적용된 Neural BP, Neural MS 알고리즘이 있다. 이들은 Tanner 그래프 기반 메시지 전달 구조를 공유하며, 변수 노드와 검사 노드 간 메시지를 반복적으로 갱신하여 오류를 정정한다.

BP 알고리즘은 베이지 추론 기반의 확률적 디코딩 방

식으로 식은 다음과 같다.

$$m_c = 2 \tanh^{-1} \left(\prod_{i=1}^n \tanh \left(\frac{m_{v,i}}{2} \right) \right) \quad (1)$$

이때 m_c 는 검사노드에서 변수 노드로 전달되는 메시지, m_v 는 변수노드에서 검사 노드로 전달되는 메시지, $m_{v,i}$ 는 해당 검사노드로 한 개 edge에 해당하는 메시지이다.

BP의 복잡한 계산식을 단순화한 MS 알고리즘은 최소 절댓값 기반 근사화를 이용하며 식은 다음과 같다.

$$m_c = \prod_{i=1}^n \min |m_{v,i}| \operatorname{sgn}(m_{v,i}) \quad (2)$$

Neural BP와 Neural MS 알고리즘은 Trellis Diagram 기반으로 변수 노드 연산을 홀수 항, 검사 노드 연산을 짝수 항으로 간주하여 신경망 형태로 확장한 뒤 각 메시지 전달 단계에 학습 가능한 가중치를 도입한 디코딩 성능 개선안이다[2]. Neural BP의 식은 다음과 같다.

이때 식(3)은 홀수 항, 식(4)는 짝수 항으로 l_v 는 로그우도비(LLR), e 는 edge, v 와 c 는 각각 변수노드, 검사노드를 뜻하며 '기호는 해당 항목을 제외한 나머지 항목, 그리고 w 가 가중치로 곱해지는 형식이다.

$$x_{i,e=(v,c)} = \tanh \left(\frac{1}{2} \left(w_{i,v} l_v + \sum_{e'=(v,c'), c \neq c'} w_{i,e,e'} x_{2L,e'} \right) \right) \quad (3)$$

$$x_{i,e=(v,c)} = 2 \tanh^{-1} \left(\prod_{(e'=(v,c'), v \neq v')} x_{i-1,e'} \right) \quad (4)$$

Neural MS의 경우 (3)을 (5)로 대체하여 계산한다.

$$x_{i,e=(v,c)} = \min_{(e'=(v,c'), v \neq v')} |x_{i-1,e'}| \prod_{(e'=(v,c'), v \neq v')} \operatorname{sgn}(x_{i-1,e'}) \quad (5)$$

식 (6)은 i 를 $2L+1$ (L 은 코드길이)까지 진행했을 때의 최종 결과값(o_v)이며, 이때 σ 는 sigmoid 함수이다.

$$o_v = \sigma(w_{2L+1,v} l_v + \sum_{e'=(v,c')} w_{2L+1,v,e'} x_{2L,e'}) \quad (6)$$

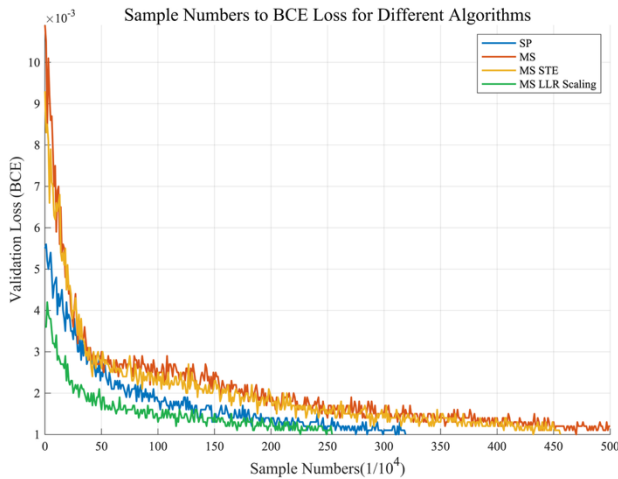


그림 1. Sample 수에 따른 BCH (63, 36)의 BCE loss 비교

2.3 안정화 알고리즘

Neural MS 구조의 학습 과정에서 비선형 함수의 gradient 단절 문제를 해결하기 위해 Straight-Through Estimator(STE) 기법이 활용될 수 있다. STE는 forward pass에서는 기존 Min-Sum의 hard decision을 유지하고, backward pass에서는 연속적인 soft 함수로 근사하여 gradient를 전달한다. 본 연구에서는 아래와 같은 수식을 사용하였다[3].

$$STE(x) = x_{hard, detach} - (x_{soft} - x_{soft, detach}) \quad (7)$$

또한, 검사 노드에서 변수 노드로 전달되는 메시지를 soft-min 함수로 근사하여 미분 가능하도록 구성하였다. 이때 soft-min은 Log-Sum-Exp (LSE) 형태로 정의되며 파라미터 β 가 무한대로 수렴할수록 true minimum 함수에 가까워진다[4].

$$\min_{soft}(x_1, x_2, \dots, x_n) = -\frac{1}{\beta} \log \sum_i \exp(-\beta x_i) \quad (8)$$

LLR Scaling은 반복 횟수가 적은 초기 단계에서 메시지 값의 반영 비율을 제한하여, 불안정한 추론 신호로 인한 성능 저하를 방지한다. 특히, MS 구조에서 발생할 수 있는 메시지 과대 반응을 완화하기 위한 scaling factor k 는 iteration에 따라 점진적으로 증가하는 하이퍼볼릭 탄젠트 함수로 정의되는데, 이 값은 이후 식(9)와 같이 m_v 에 곱해져 값을 조절한다.

$$k = \tanh(0.1 * \min(10, iteration)) \quad (9)$$

3. 신경망 메시지 전달 복호 성능 분석

본 실험에서는 PCM(Parity Check Matrix)로 BCH (63,36), max Iteration 10, training num 10000, epoch 500(총 $5 * 10^6$ sample 수), target uncorrected error 200로 설정하였다. 그림 1은 iteration이 10일 때 BCE loss 0.001에 도달하기까지의 sample 수 대비 loss 변화를 나타낸다. STE 적용 시 전반적으로 수렴 속도가 향상되었으며, threshold loss 달성에 필요한 sample 수 역시 $4.5 * 10^6$ 수준으로 감소하였다. LLR scaling 적용 모델은 모든 epoch 구간에서 loss 감소 폭이 가장 컸고 Neural MS 대비 절반의 sample 수로 threshold loss에 도달했다.

그림 2는 동일한 실험 조건에서 $E_b/N_0=6$ dB 일 때, 학습 샘플

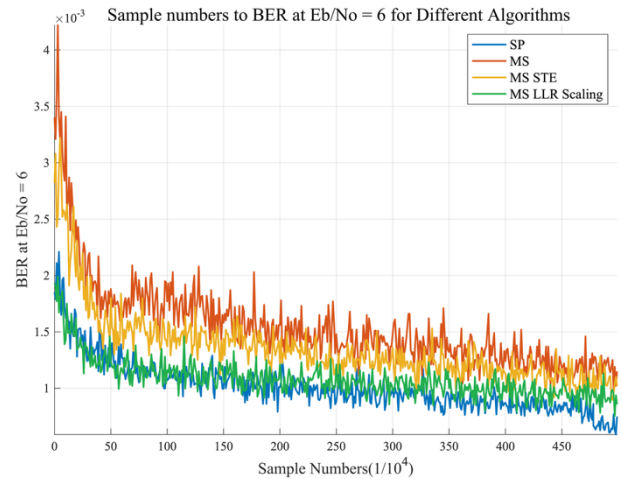


그림 2. $E_b/N_0=6$, 샘플수에 따른 BCH (63, 36)의 BER 비교

수에 따른 BER(Bit Error Rate) 변화를 비교한 결과를 나타낸다. 샘플 수가 $2.5 * 10^6$ 을 초과하면, 모든 모델에서 BER 성능이 수렴하는 양상을 보이며, 이후 성능 차이는 통계적으로 미미한 수준으로 축소된다. LLR Scaling을 적용한 Neural MS 모델은 초기 학습 단계에서 BCE loss가 더 낮게 형성되었음에도 불구하고, BER 기준으로는 Neural SP와 유사한 성능 수준에 도달하였다. 이는 LLR Scaling 기법이 학습 수렴 속도에는 기여하나, 최종 복호 성능 향상에는 제한적인 영향을 미칠 수 있음을 시사한다.

4. 결론

본 연구는 Neural MS Decoding의 학습 최적화를 위한 기법으로 STE와 LLR scaling을 적용하고, 그 효과를 sample 수 별 BCE loss, BER 분석을 통해 정량적으로 비교하였다. 실험 결과, LLR scaling 기법이 기존 Neural MS 구조의 성능을 향상시키는 것을 확인하였다.

향후에는 다양한 부호율 및 코드 길이에 대한 일반화 가능성 검토, LLR scaling factor의 정규화 혹은 다중 계층 적용 가능성, 하드웨어 효율성 평가 등을 통해 해당 기법의 실용성을 보다 넓은 스펙트럼에서 검증할 필요가 있다.

ACKNOWLEDGEMENT

이 논문은 2025년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원(RS-2024-00398449, RS-2024-00397216)을 받아 수행된 연구임.

5. 참고 문헌

- [1] E. Nachmani *et al.*: "Deep Learning Methods for Improved Decoding of Linear Codes," *IEEE J. Sel. Top. Signal Process.*, vol. 12, no. 1, pp. 119–131, Feb. 2018.
- [2] R. C. Bose and D. K. Ray-Chaudhuri, "On a class of error correcting binary group codes", *Information and Control*, vol. 3, no. 1, pp. 68–79, March 1960.
- [3] G. C. Calafiore *et al.*: "A universal approximation result for difference of log-sum-exp neural networks,"
- [4] Y. Bengio *et al.*: "Estimating or propagating gradients through stochastic neurons for conditional computation," Dept. of Computer Science and Operations Research, Université de Montréal, Montréal, QC, Canada