

# 해밍 코드 디코딩을 위한 이진 신경망 기반 디코더

이시윤, 권나현, 김정현

세종대학교

23012130@sju.ac.kr, nhkwon9803@sju.ac.kr, j.kim@sejong.ac.kr

## A Binary Neural Network Decoder for Hamming Codes

Siyoon Lee, Nahyeon Kwon, Junghyun Kim

Sejong Univ.

### 요약

본 논문에서는 저전력 통신 환경을 위한 이진 신경망(binary neural network, BNN) 기반의 BNN 디코더(BNN decoder)를 제안한다. 제안하는 디코더에는 이진화로 인한 성능 저하 문제를 해결하기 위해 이진 포컬 손실(binary focal loss)과 패리티 일관성 손실(parity consistency loss)을 적용하고, 이중 소프트 임베더(double soft embedder) 구조를 도입하였다. 실험 결과 Hamming(15,11) 코드 기준으로 Classical hard-decision(HD) Decoder 대비 블록 오류율(block error rate, BLER)이 최대 71% 감소함을 확인하였다.

## I. 서론

차세대 통신 시스템은 초저지연과 저전력을 핵심적으로 요구한다. 이에 따라 기존 알고리즘의 병렬 처리 이점을 유지하면서도 디코딩 성능을 유의미하게 개선할 수 있는 딥러닝 기반 오류 정정 부호 디코딩 기술의 중요성이 커지고 있다 [1]. 하지만 딥러닝은 대부분 실수형 연산에 기반하기 때문에 전력 소모와 메모리 사용량이 크다는 한계가 있다 [2]. 이에 대한 대안으로 가중치와 연산을 1비트로 제한하여 메모리와 연산 복잡도를 줄이는 이진 신경망(binary neural network, BNN)이 제안되고 있다 [2, 3]. 하지만 이진 신경망은 내부 가중치와 연산의 완전한 이진화로 인해 모델의 표현력이 제한되기 때문에, 코드의 대수적 구조에 기반한 Classical hard-decision(HD) Decoder의 성능 수준에 도달하는 데 한계가 있다.

본 논문에서는 이러한 문제를 해결하기 위해 다음 두 가지를 제안한다. 첫째, 제한된 표현력 내에서도 모델이 코드의 대수적 성질을 학습할 수 있도록 이진 포컬 손실(binary focal loss) [4]과 패리티 일관성 손실(parity consistency loss)을 결합한 BNN 디코더(BNN decoder)를 제안한다. 둘째, 실수형 채널 출력값을 사용하는 연속 입력 환경에서 급격한 이진화로 인한 정보 손실을 방지하기 위해 입력 처리부인 임베더(embedder)를 이중화한 이중 소프트 임베더(double soft embedder) 구조를 제안한다.

실험 결과, 제안하는 BNN Decoder는 이진 입력 조건에서 Classical HD Decoder와 동등한 수준의 블록 오류율(block error rate, BLER) 성능을

달성하였다. 나아가 연속 입력 환경에서는 이중 소프트 임베더 구조를 통해 Hamming(15, 11) 코드 기준 BLER를 약 71% 감소시킴으로써, 이진화된 신경망 구조가 연속 정보를 효과적으로 활용하여 디코딩 성능을 향상시킬 수 있음을 확인하였다.

## II. 본론

송신 메시지 벡터를  $m$ , 해밍 인코딩 함수를  $g(\cdot)$ , 이진 위상 편이 변조(binary phase shift keying, BPSK) 함수를  $f(\cdot)$ 라 정의하면, 송신된 코드워드  $c$ 는  $c = g(m)$ 이며, 변조된 신호  $x$ 는  $x = f(c)$ 가 된다. 채널을 통과하며 가산 백색 가우시안 잡음(additive white Gaussian noise, AWGN)  $n$ 이 더해진 최종 수신 신호  $y$ 는 식 (1)과 같이 정의된다.

$$y = f(g(m)) + n \quad (1)$$

본 논문에서는 이러한 통신 환경에서 이진 입력만을 사용하는 Classical HD Decoder와 동등한 성능을 달성하는 BNN Decoder를 제안한다. 제안하는 디코더는 이진화된 가중치와 활성화 함수로 인한 학습 어려움을 완화하기 위해 직통 추정기(straight-through estimator, STE) [3], XNOR 스케일링(XNOR scaling) [2], 확률적 이진화(stochastic binarization) [3], 드롭커넥트(DropConnect) [6]를 적용하였다. 또한, 높은 신호 대 잡음비(signal-to-noise ratio, SNR) 환경에서는 오류 발생 확률이 낮아 희소한 오류 패턴이 충분히 학습되지 않는 문제가 있다. 이를 해결하기 위해 이진

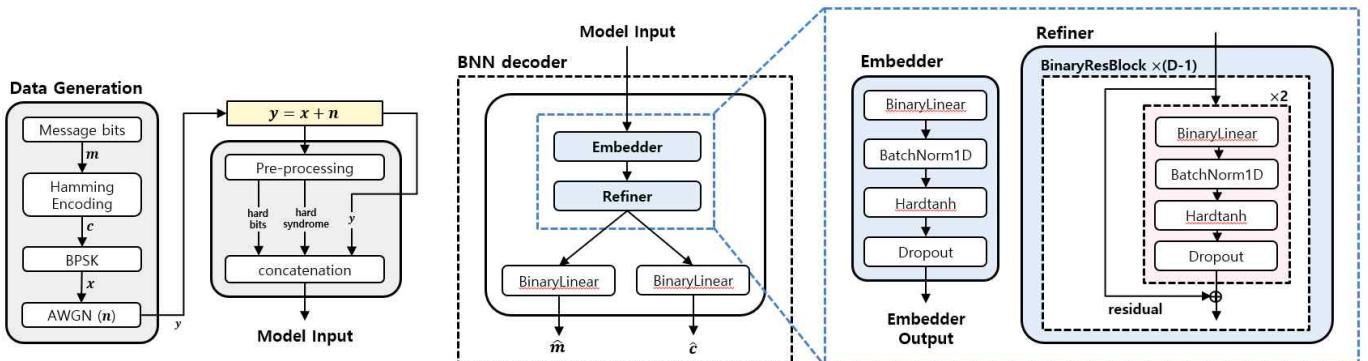


그림 1. 하드 임베더 및 소프트 임베더 기반 디코더에 공통으로 적용되는 BNN Decoder의 전체 구조

포컬 손실 [4]을 도입하여 오류에 대한 학습 비중을 강화하였다. 또한, 디코더가 해밍 코드의 구조를 학습할 수 있도록 패리티 일관성 손실을 식(2)와 같이 정의하여 전체 손실 함수에 반영하였다.

$$L_{\text{parity}} = \frac{1}{M} \sum_{j=1}^M \left( 1 - \prod_{k \in I_j} \tanh\left(\frac{z_k}{2}\right) \right) \quad (2)$$

여기서  $z_k$ 는  $k$ 번째 비트의 로짓(logit),  $I_j$ 는  $j$ 번째 패리티 검사에 포함된 비트의 인덱스 집합,  $M$ 은 패리티 검사식의 총 개수를 의미한다. 이는 기준의 신드롬 기반 디코딩에서 사용된 패리티 제약 조건을, 신경망 학습이 가능하도록 미분 가능한 손실 함수 형태로 재구성한 것이다.

그림 1은 하드 임베더(hard embedder)와 소프트 임베더(soft embedder) 기반 디코더에 공통으로 적용되는 BNN Decoder의 아키텍처를 나타낸다. 제안하는 모든 BNN 기반 디코더는 입력으로 실수형 수신 신호  $y$ 를 사용하며, 임베더 블록에서의 이진화에 방식에 따라 하드 임베더와 소프트 임베더 구조로 구분된다. 기본 BNN Decoder는 임베더의 첫 번째 계층에서 입력을 즉시 이진화하는 하드 임베더 구조를 사용한다. 반면, 소프트 임베더는 연속 표현을 유지한 상태에서 특징을 추출한 후 이진화를 수행하며, 소프트 임베더 블록의 개수에 따라 단일 소프트 임베더(single soft embedder)와 이중 소프트 임베더(double soft embedder)로 구분한다.

디코더의 입력은 신드롬 기반 접근법(syndrome-based approach) [5]에 따라, 수신 신호  $y$ 로부터 얻은 이진 비트( $\pm 1$ )와 패리티 검사 행렬과의 연산을 통해 얻은 이진 신드롬( $\pm 1$ )을 결합하여 구성된다. 디코더는 입력 신호를 고차원 이진 특징으로 변환하는 임베더, 오류 패턴을 보정하는 리파이너(refiner), 그리고 최종 메시지와 코드워드를 각각 예측하는 예측단으로 구성된다. 이때, 네트워크 전반의 활성화 함수로는 출력값을  $[-1, 1]$  범위로 제한하는 Hardtanh 활성화 함수를 사용하여 부호 함수(sign function)와의 동작 일관성을 확보하였다. 예측단은 두 개의 분기 구조로 이루어져 각각 메시지  $\hat{m}$ 과 코드워드  $\hat{c}$ 를 예측한다.

평가 단계에서는 예측 로짓에 시그모이드(sigmoid) 함수를 적용한 값을 신뢰도로 정의하고, 두 출력 중 신뢰도가 더 높은 결과를 선택한다. 예측된 코드워드가 패리티 검사를 만족하지 못하는 경우에는 신드롬 기반 단일 비트 오류 정정을 수행하여 디코딩 결과가 유효한 코드 공간을 만족하도록 보정한다. 또한 추론 과정의 불확실성을 완화하기 위해 몬테카를로 드롭아웃(MC-Dropout) [7] 기반 양상을 추론을 적용하였다.

실험 결과, 제안하는 BNN Decoder는 이진 입력만을 사용하는 조건에서 모든 신호 대 잡음비에 대해 Classical HD Decoder와 동등한 수준의 BLER 성능을 달성하였으며, 이는 그림 2(a)에 제시하였다.

Classical HD Decoder의 성능 한계를 극복하기 위한 연속 입력 기반 구조 분석은 그림 2(b)에 제시되어 있으며, 단일 및 이중 소프트 임베더 구조의 성능 비교를 통해 이를 확인하였다. 단일 소프트 임베더 구조는 연속 입력을 처리하는 소프트 임베더 계층이 하나로 구성되어 있어 유의미한 특징 추출에 제한적이다. 이를 보완하기 위해 소프트 임베더를 두 개의 계층으로 확장한 이중 소프트 임베더 구조를 도입하였다. 이중 소프트 임베더에서는 두 단계의 특징 추출을 통해 급격한 이진화로 인한 정보 손실을 완화한다. 또한, 입력 단계에서의 정보 보존을 위해 소프트 임베더 계층에서는 드롭아웃을 제거하고, 과적합을 방지하기 위해서 리파이너 계층에서는 드롭아웃을 적용하는 방법을 사용하였다.

실험 결과, Hamming(15, 11) 코드에서 SNR 7.0 dB 기준으로 연속 입력을 단일 소프트 임베더 구조로 처리했을 때의 BLER은  $4.40 \times 10^{-5}$ 로, Classical HD Decoder의 BLER인  $6.61 \times 10^{-5}$  대비 약 33% 감소하였다. 나아가 이중 소프트 임베더 구조를 적용한 경우에는 동일 조건에서의

BLER이  $1.93 \times 10^{-5}$ 로, Classical HD Decoder 대비 약 71%의 상대적 성능 개선을 달성하였다.

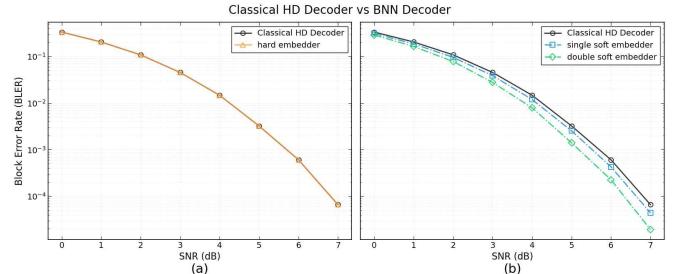


그림 2. 기존 HD 디코더와 제안한 BNN 기반 디코더의 BLER 성능 비교:  
(a) 하드 임베더, (b) 소프트 임베더.

### III. 결 론

본 논문에서는 이진 신경망 기반 해밍 코드 디코더의 성능 저하를 극복하기 위한 학습 전략과 디코더 구조를 제안하였다. Classical HD Decoder와 비교하였을 때, 제안하는 BNN Decoder는 이진 입력 조건에서 기준 디코더와 동등한 BLER 성능 수준을 달성하였으며, 이중 소프트 임베더 구조를 적용한 경우에는 BLER를 약 71% 감소시켰다. 이러한 결과는 추가적인 복잡한 디코딩 알고리즘의 변경 없이도 임베더의 구조적 심층화와 초기 연속 입력 정보 보존만을 통해 이진 신경망이 연속 정보를 효과적으로 활용할 수 있음을 시사한다. 나아가 본 연구는 차세대 통신 시스템에서 딥러닝 기반 채널 코딩에 이진 신경망을 적용하기 위한 기초 연구로서 의의를 가진다.

### ACKNOWLEDGMENT

이 논문은 2025년도 교육부 및 서울특별시의 재원으로 서울RISE센터의 지원을 받아 수행된 지역혁신중심 대학지원체계(RISE)의 결과입니다.  
(2025-RISE-01-019-04)

### 참 고 문 헌

- [1] E. Nachmani, Y. Be'ery, and D. Burshtein, "Learning to decode linear codes using deep learning," in *Proc. Allerton Conf. Commun., Control, and Computing*, pp. 341 - 346, Sep. 27 - 30, 2016.
- [2] M. Rastegari, V. Ordonez, J. Redmon, and A. Farhadi, "XNOR-Net: ImageNet classification using binary convolutional neural networks," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, pp. 525 - 542, Oct. 11 - 14, 2016.
- [3] I. Hubara, M. Courbariaux, D. Soudry, R. El-Yaniv, and Y. Bengio, "Binarized neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, pp. 4107 - 4115, Dec. 5 - 10, 2016.
- [4] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, pp. 2980 - 2988, Oct. 22 - 29, 2017.
- [5] A. Bennatan, Y. Choukroun, and P. Kisilev, "Deep learning for decoding of linear codes—A syndrome-based approach," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 1595 - 1599, Jun. 17 - 22, 2018.
- [6] L. Wan, M. Zeiler, S. Zhang, Y. LeCun, and R. Fergus, "Regularization of neural networks using DropConnect," in *Proc. Int. Conf. Mach. Learn. (ICML)*, pp. 1058 - 1066, Jun. 16 - 21, 2013.
- [7] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian approximation: Representing model uncertainty in deep learning," in *Proc. Int. Conf. Mach. Learn. (ICML)*, pp. 1050 - 1059, Jun. 19 - 24, 2016.