

전이 학습을 이용한 Error Correction Code Transformer 성능 개선

심윤범, 곽희열
울산대학교

syb0915@naver.com, ghy1228@gmail.com

Performance Improvement of the Error Correction Code Transformer by Transfer Learning

Sim Yun Beom, Kwak Hee Youl
University of Ulsan

요약

본 논문은 Error Correction Code Transformer(ECCT)의 성능 개선을 위해 다른 부호에서 이미 학습된 임베딩 가중치를 전이시키는 방법을 제시한다. 이 방법은 임베딩 가중치는 각 부호 bit의 차원을 확장해주는 역할을 하며 최적화된 임베딩 가중치가 부호에 따라 크게 다르지 않기에 전이 학습에 적합하다. 모의 실험을 통해 BCH(63,36) 코드에서 학습된 가중치를 BCH(31,16) 코드에 적용함으로써 학습 속도와 복호 성능을 향상시킬 수 있음을 입증했다.

I. 서론

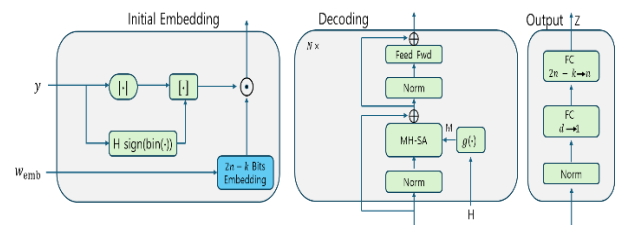
AI 분야의 핵심 기술로 떠오른 Transformer 기술의 발전은 많은 분야에서 혁신을 가져왔으며, 특히 오류 정정 코드(Error Correction Code, ECC) 분야에서의 적용은 Error correction code transformer(ECCT)라는 새로운 종류의 복호기를 만들어냈다. 본 논문에서는 ECCT의 성능을 더 향상시키기 위해 임베딩 가중치를 Transfer learning 하는 방법을 제안한다. 임베딩 가중치는 각 bit를 더 높은 차원으로 확장시켜 주는 가중치이다. 즉, 서로 다른 bit들을 Transformer가 구분할 수 있게 해주는 역할을 하며 부호의 종류와 무관할 것으로 기대된다.

따라서 다른 부호에서 학습된 임베딩 가중치를 초기 임베딩 가중치로 사용하는 Transfer learning을 제안한다. 이로써 더 빠른 학습 속도와 더 우수한 복호 성능을 보일 수 있다.

II. ECCT의 구성

ECCT의 구성은 그림 1과 같다. 채널로부터 수신 받은 길이 n 인 y 수신 벡터로부터 preprocessing 과정을 거쳐 길이 $2n-k$ 인 \hat{y} 를 생성한다. 이후 여기에 차원이 $(2n-k) \times d$ 인 임베딩 가중치를 곱하여 Decoding layer의 입력으로 넣어준다. 이 과정은 Transformer 모델이 데이터를 효과적으로 처리할 수 있도록 준비하는 핵심 단계이다. 디코딩 모듈은 여러 개의 마스크된 자기 주의(Masked Self-Attention) 및 피드 포워드(Feed Forward) 레이어를 포함하고 있으며, 각 레이어 사이의 정규화 레

이어는 학습 안정성과 성능을 개선한다. 최종 출력 모듈은 임베딩된 데이터를 먼저 $2n-k$ 차원의 벡터로 축소한다. 다음, 이를 n 차원 벡터로 변환하여 소프트 디코딩된 결과를 출력한다.



<그림 1> ECCT 구성

III. 임베딩 가중치의 중요성

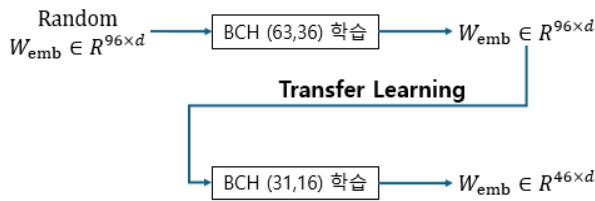
임베딩이란, 데이터의 특성을 더 잘 표현하기 위해, 원래의 데이터를 더 높은 차원의 공간으로 매핑하는 과정을 말한다. 임베딩 가중치는 그림 2와 같이 $2n-k \times d$ 차원의 행렬로 표현할 수 있다. 그림 1의 초기 임베딩 단계에서 임베딩 가중치인 W_{emb} 를 곱해 $2n-k$ 비트 임베딩을 시킨다. 즉, 임베딩 가중치는 각 데이터 비트를 독립적으로 더 높은 차원의 벡터로 매핑시켜 주며, 따라서 부호의 종류나 길이에 따라 크게 달라지지 않을 것으로 보인다. 따라서, 다양한 부호들 간에 임베딩 가중치를 공유할 수 있을 것으로 생각되어 Transfer learning을 적용해보았다.

$$W_{emb} = \left(\begin{array}{c} \dots \\ \vdots \\ \dots \end{array} \right) \left\} \begin{array}{c} d \text{ model} \\ 2n - k \end{array} \right.$$

<그림 2> 임베딩 가중치 구조

IV. 전이 학습의 구현

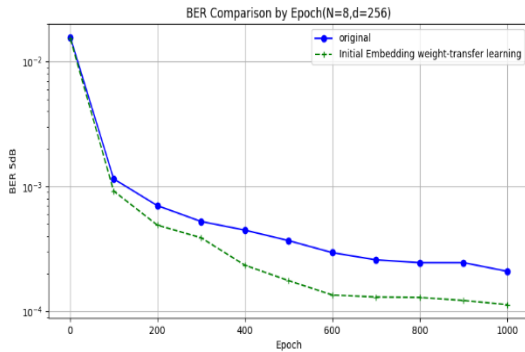
전이 학습이란, 한 분야에서 학습된 지식을 다른 분야에 적용하는 기법이다. 그림 3은 BCH(63,36) 코드에서 학습된 임베딩 가중치를 BCH(31,16) 코드에 적용하는 전이 학습 과정을 설명한다. 구체적으로 BCH(63,36) 코드의 96개 행 중에서 46개 행을 선택하여 추출할 것이다. 첫 번째 행부터 순차적으로 필요한 수만큼의 행을 잘라내는 방법을 사용하여 추출한 후 BCH(31,16) 코드의 임베딩 가중치의 초기값으로 활용한다.



<그림 3> Transfer Learning 과정

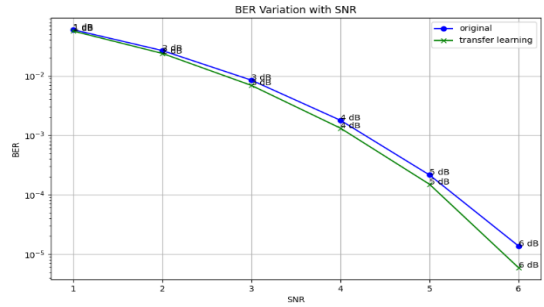
V. 모의 실험 결과

BCH(63,36) 코드로 학습된 임베딩 가중치를 BCH(31,16) 코드에 적용하는 전이 학습의 효과를 모의 실험 결과로 검증했다. 본 실험은 N=8, d=256의 매개변수를 가진 모델이다. <그림 3>은 Epoch에 따른 BER(Bit Error Rate) 5dB의 변화를 보여준다. 결과는 그림 4와 같다. Epoch 100 이상부터는 확실한 성능 차이를 확인할 수 있다.



<그림 3> 전이 학습을 진행한 ECCT와 일반 ECCT의 Epoch에 따른 BER 비교

본 실험에서 사용한 임베딩 가중치는 기존에 BCH(63,16) 코드에서 학습을 마치고, 최적화를 거쳤으며, 이는 새로운 BCH(31,16) 코드의 학습 과정에서의 그림 5와 같이 더 높은 복호 성능을 보여준다. 이를 통해 기존 ECCT보다 성능 또한 우수함을 알 수 있다.



<그림 5> SNR에 따른 BER 변화

VI. 결론

본 논문에서는 ECCT의 성능 개선을 위해 초기 임베딩 단계에서 임베딩 가중치의 전이 학습을 활용하는 방법을 제안하였다. 제안한 방법은 더 낮은 BER로 수렴하는 점이 전이된 임베딩 가중치가 이미 일부 유용한 패턴을 학습했기 때문으로 해석된다. 이는 새로운 코드에 대한 학습을 가속화하며, 특히 초기 학습 단계에서의 성능 향상이 두드러진다. 결론적으로, 학습된 가중치 전이 학습을 통해 기존 ECCT보다 학습 시간을 단축시키고, 더 효율적인 학습 과정임을 실험을 통하여 검증하였다.

ACKNOWLEDGMENT

이 논문은 2024년 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No. RS-2023-00247197, LDPC 부호의 한계 극복으로 응용 분야 확대)

참고 문헌

- [1] Choukroun, Yoni, and Lior Wolf. "Error correction code transformer." *Advances in Neural Information Processing Systems* 35 (2022): 38695–38705.
- [2] Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems* 30 (2017).
- [3] Torrey, Lisa, and Jude Shavlik. "Transfer learning." *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*. IGI global, 2010. 242–264.