다중사용자 상향링크 다중안테나 시스템을 위한 업무 지향적 통신 기술

이도윤°, 김승욱*, 이훈°
°울산과학기술원, *부경대학교

ehdbs8964@unist.ac.kr, swkim@pknu.ac.kr, hoonlee@unist.ac.kr

Task-Oriented Communication Systems for Multi-User Uplink MIMO Network

Doyun Lee°, Seung-Wook Kim*, and Hoon Lee°
°Ulsan National Institute of Science and Technology, *Pukyong National University

요 약

최근 수신단의 업무 수행 관점에서 통신 모듈을 통합적으로 설계하는 업무 지향적 통신 기술이 각광받고 있다. 본 논문에서는 인공지능 기술을 활용하여 다중사용자 상향링크 환경에서 사용자의 수에 대한 확장성을 보장하는 다중안테나 업무 지향적 통신 기법을 제안한다. 모의실험을 통해 영상 분류를 수행하는 환경에서 제안하는 기법의 효과를 입증한다.

Ⅰ. 서 론

전통적인 통신 기술은 수신단의 최종 목표 업무와 관계없이 데이터의 단순 복원 성능에 초점을 두어 설계되어 많은 비효율이 발생한다. 최근 인공지능 기술을 사용하여 수신단의 업무 수행 성능을 직접적으로 극대화할 수 있는 충분통계량 (sufficient statistic)을 추출하여 전송하는 업무 지향적 (task-oriented) 시스템이 주목받고 있다[1-4]. 성공적인 업무 수행에 필요한 필수 정보만을 전송하므로 높은 통신 효율을 달성할 수 있다. 그러나, 인공지능 구조의 특성상 확장성이 매우 낮아 사용자 수가 달라지면 새로운 모델 설계와 훈련 작업이 필수적이다.

본 논문에서는 임의의 사용자 수에 적용 가능한 다중사용자 다중안테나 상향링크 업무 지향적 통신 시스템을 설계한다. 모든 사용자에게 동일한 인공지능 모델을 적용하는 파라미터 공유 전략을 활용한다. 이를 통해 테스트 환경에서 사용자 수가 변경되어도 추가적인 미세조정 단계 없이 단일 인공지능 모델을 바로 사용하는 것이 가능하다. 모의실험 결과를 통해 제안하는 기법의 효율성을 입증한다.

Ⅱ. 시스템 모델

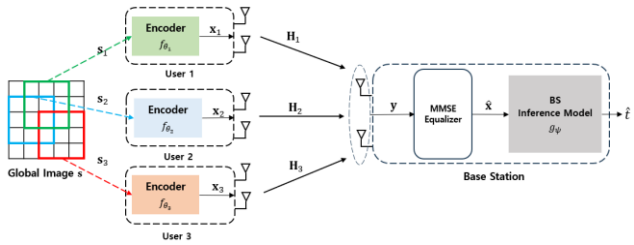

그림 1. 다중사용자 상향링크 업무 지향적 통신 시스템

그림 1 은 제안하는 다중사용자 상향링크 다중안테나 업무 지향적 통신 시스템을 나타낸다. M_T 개의 송신 안테나를 갖는 N 개의 사용자가 개별적으로 취득한 관찰값 정보 (예: 영상, 센싱 데이터 등)을 상향링크 채널을 통해 M_R 개의 수신안테나를 갖는 기지국에게 전송한다. 기지국은 수신 신호를 기반으로 관찰 정보의 레이블 예측을 수행한다.

전역 관찰 정보 $\mathbf{s}$ 중 사용자 i 는 그 일부분인 지역 관찰 정보 $\mathbf{s}_i \in \mathbb{R}^S$ 를 획득한다. 각 사용자는 인코더 신경망 $f_{\theta_i}: \mathbb{R}^S \to \mathbb{C}^{M_T}$ 을 통해 복소 채널 심볼 $\mathbf{x}_i = f_{\theta_i}(\mathbf{s}_i) \in \mathbb{C}^X$ 을 생성한다. 이때, 다음의 전력 제한 조건을 만족하도록 인코더 신경망을 설계한다.

$$\|\mathbf{x}_i\|^2 \le P \tag{1}$$

사용자별 상향링크 채널 $\mathbf{H}_i \in \mathbb{C}^{M_R \times M_T}$ 을 통해 기지국에서 수신한 신호 벡터 $\mathbf{y} \in \mathbb{C}^{M_R}$ 는 다음과 같다.

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{2}$$

이때 $\mathbf{x} = [\mathbf{x}_1^T, \mathbf{x}_2^T, \cdots \mathbf{x}_N^T]^T \in \mathbb{C}^{NM_T}$ 는 다중사용자 상향링크 송신 신호 벡터, $\mathbf{H} = [\mathbf{H}_1, \mathbf{H}_2, \cdots \mathbf{H}_N] \in \mathbb{C}^{M_R \times NM_T}$ 는 상향링크 레일리 페이딩 채널 행렬, 그리고 $\mathbf{n} \in \mathbb{C}^{M_R}$ 은 가우시안 잡음 벡터를 의미한다.

기지국에서는 minimum mean-squared-error (MMSE) 수신기를 통해 아래와 같이 송신 신호를 추정한다.

$$\hat{\mathbf{x}} = \mathbf{H}(\mathbf{H}\mathbf{H}^H + \sigma^2\mathbf{I})^{-1}\mathbf{y} \tag{3}$$

이때 $\hat{\mathbf{x}} = [\hat{\mathbf{x}}_1^T, \hat{\mathbf{x}}_2^T, \cdots \hat{\mathbf{x}}_N^T]^T \in \mathbb{C}^{NM_T}$ 는 추정된 신호 벡터를 나타낸다.

최종 레이블을 예측하기 위해 기지국은 추론 신경망 $g_\psi: \mathbb{C}^{NM_T} \to \mathbb{R}$ 을 사용하여 예측을 수행한다. 최종 예측 레이블은 $\hat{t} = g_\psi(\hat{\mathbf{X}}) \in \mathbb{R}$ 로 계산된다.

Ⅲ. 제안하는 업무 지향적 통신 기술

상기 신경망 모델은 사용자 수 N 에 비례하여 인코더 신경망 개수와 기지국 추론 신경망의 입력 차원이 증가하여 그 확장성이 낮다. 이를 해결하기 위해 파라미터 공유 전략을 사용하여 확장 가능한 신경망 구조를 개발한다. 인코더 신경망의 개수를 사용자 수와 무관하도록 설계하기 위해 모든 사용자가 동일한 인코더 모델을 사용하도록 다음과 같이 제한한다.

$$f_\theta = f_{\theta_i}, \forall i \in \mathcal{N} \tag{4}$$

사용자 인덱스 집합을 $\mathcal{N} \triangleq [1, N]$ 으로 정의한다. 상기 파

라미터 공유를 통해 훈련 과정에서 개별 사용자들에서만 관찰할 수 있었던 지역 관찰 정보를 모든 사용자가 공유할 수 있어 일반화 능력 향상에 크게 기여할 수 있다.

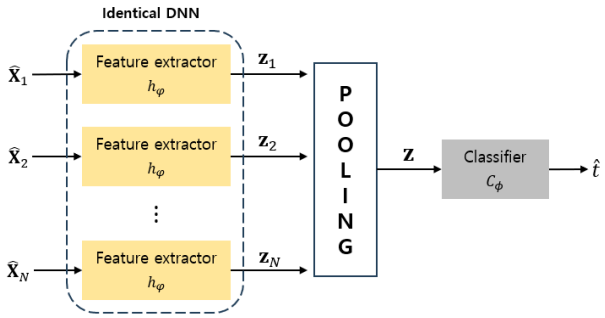

그림 2. 제안하는 기지국 추론 신경망 구조

기지국 추론 신경망의 확장성을 증대하기 위해 그림2와 같은 모듈화된 구조를 제안한다. 제안하는 기지국 추론 신경망은 특징 추출기, pooling, 그리고 분류기로 구성된다. 개별 사용자에 대한 복원 신호 $\hat{\mathbf{x}}_i$ 를 동일한 특징 추출기에 입력하여 특징 벡터 $\mathbf{z}_i$ 를 $\mathbf{z}_i = g_\varphi(\hat{\mathbf{X}}_i) \in \mathbb{R}^Z$ 와 같이 추출한다. 이 결과를sum pooling 연산을 통해 취합된 특징 벡터 $\mathbf{z}$ 를 생성한다. 이를 분류기의 입력으로 사용하여 예측 레이블 $\hat{t}$ 을 추론한다. 따라서 전체 클라우드 모델의 신경망 파라미터 집합은 $\psi \triangleq \varphi \cup \phi$ 이다.

IV. 모의실험 결과

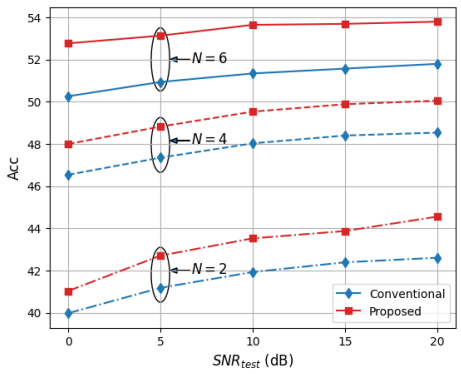

그림 3. SNR 대비 영상 분류 정확도 성능

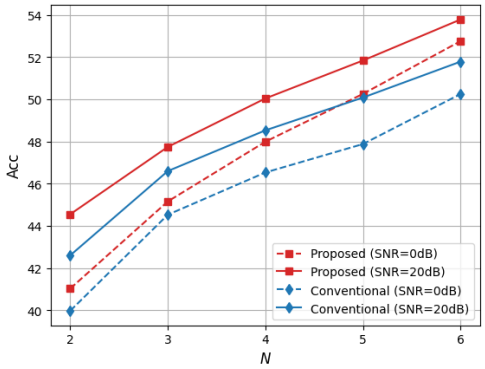

그림 4 사용자 수 대비 영상 분류 정확도 성능

본 절에서는 CIFAR10 데이터집합을 활용하여 제안하는 기법을 영상 분류 상황에 적용하여 그 효과를 입증한다. Cross-entropy 손실함수를 기반으로 인코더 신경망 및 기지국 추론 신경망의 통합 훈련을 실시한다. 배치 크기는 512, 학습률 8e-5 의 Adam 알고리즘을 사용한다. 인코더 신경망은 3 계층 합성곱 신경망으로 구성되어 있으며 $M_T = 2$ 의 수신 안테나를 장치하고 있다. 기지국은 $M_R = NM_T$ 개의 수신안테나를 사용한다. 훈련 시 SNR 을 $[0dB, 20dB]$ 사이에서 무작위로 발생시켜 가우시안 잡음의 분산에 대한 강건성을 확보한다. Conventional 기법은 확장성이 없는 신경망 구조를 사용하며, 제안하는 기법과 동일한 계층 수와 파라미터 수를 갖도록 설정한다.

그림 3 은 영상 분류 정확도를 SNR 에 대해 도시하였다. SNR 이 증가함에 따라 모든 기법의 성능이 향상되며, 이러한 현상은 사용자 수 N 에 무관하게 관찰된다. 사용자 수가 증가하면 예측에 활용하는 정보량이 많아지므로 정확도가 증가한다. 모든 실험 환경에서 제안하는 확장 가능한 신경망 구조가 기존 기법 대비 높은 정확도를 달성한다.

그림 4 는 사용자 수에 따른 분류 정확도를 나타낸다. 예상했던 대로, 사용자 수가 증가함에 따라 정확도가 상승함을 확인할 수 있다. 모든 사용자 수에 대해 제안하는 기법이 기존 기법보다 우월한 성능을 보인다.

Ⅲ. 결론

본 논문에서는 임의의 사용자 수에 유동적으로 적용 가능한 상향링크 다중안테나 업무 지향적 통신 기술을 개발하였다. 적절한 파라미터 공유 전략을 통해 인코더 신경망 및 기지국 추론 신경망의 비효율성을 제거하여 높은 추론 성능을 달성할 수 있다. 모의실험 결과 제안하는 클라우드 추론 구조가 기존 기법 대비 높은 실효성과 확장성을 입증하였다.

ACKNOWLEDGMENT

이 논문은 2022 년도 정부(방위사업청)의 재원으로 국방기술진흥연구소 (21-106-A00-007, 우주계층 지능통신망 특화연구실) 및 2023 년도 정부(과학기술정보통신부) (NRF-2022R1F1A1074009)의 지원을 받아 수행된 연구임.

참 고 문 헌

[1] J. Shao, Y. Mao, and J. Zhang, "Learning Task-Oriented Communication for Edge Inference: An Information Bottleneck Approach," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 1, pp. 197-211, Jan 2022

[2] H. Xie, Z. Qin, X. Tao, K. and B. Letaief, "Task-Oriented Multi-User Semantic Communications," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 9, pp. 2584-2597, Sep 2022

[3] H. Lee, and S. W. Kim, "Task-Oriented Edge Networks: Decentralized Learning Over Wireless Fronthaul," *IEEE Internet Things J.*, vol. 11, no. 9, pp. 15540-15556, May 2024

[4] J. Shao, Y. Mao, and J. Zhang, "Task-Oriented Communication for Multidevice Cooperative Edge Inference," *IEEE Trans. Wireless Commun.*, vol. 22, no. 1, pp. 73-87, Jan 2023