

시맨틱 통신 연구 동향: 기회 및 해결해야 할 문제들

박소정, 주호성, 장종규, 양현중*
포항공과대학교

sojeong@postech.ac.kr, zxcqa123@postech.ac.kr, igjang@postech.ac.kr,
hyunyang@postech.ac.kr

A Survey on Semantic Communications: Opportunities and Challenges

Sojeong Park, Hosung Joo, Jonggyu Jang, Hyun Jong Yang*
POSTECH

요약

본 논문에서는 최신 통신 네트워크 기술 중 하나인 시맨틱 통신의 최신 연구 동향에 대해 통찰하고자 한다. 먼저, 기존 통신 기술의 문제점에 대해 서술하고, 시맨틱 통신 기술로 해당 문제를 해결하는 방법에 대해 소개한다. 이를 위해 최신 시맨틱 통신 연구들의 접근법을 소개하며, 해당 기술들의 한계점과 향후 연구 방향에 대해 서술하고자 한다. 기존 통신 기술과 달리 시맨틱 통신에서는 bit 기반으로 정보를 전달하는 것이 아니므로 통신 성능을 측정해 줄 새로운 metric이 필요하다. 자원할당 성능 측정을 위해 semantic transmission rate, semantic spectral efficiency와 같은 새로운 정의를 제안한 연구들이 있지만, 다양한 용도에 대한 통합 시맨틱 통신 기술 벤치마크 metric의 개발의 필요성은 여전히 남아있다. 또한 텍스트에 국한되지 않는 multi-modal 시맨틱 통신에 대한 자원할당 기술 개발 역시 필요한 실정이다.

I. 서론

VR/AR 등 최신 서비스들의 등장으로 모바일/IoT 디바이스를 중심으로 하여 데이터 트래픽이 폭발적으로 증가할 것으로 예상된다. 그러나 사용 가능한 주파수 자원은 한정되어 있으므로 주파수 사용 효율을 높여 많은 데이터를 전송하는 것이 필요하다. 하지만, 기존 bit 기반 주파수 효율 증대 기술들은 이미 수렴단계에 이르렀으며, 현 시점에서 통신 효율을 한 단계 더 향상시키는 방법은 주로 기지국과 안테나 수를 늘리는 것인데, 이는 설비 시간과 비용적으로 비효율적일 수 밖에 없다. 따라서, 주파수 자원을 더욱 효율적으로 사용할 수 있는 새로운 기술이 필요한 시점이다.

통신의 주요한 목적은 특정 업무 수행을 위한 정보를 주고 받는 것이다. 예를 들어 대부분의 텍스트 전송의 경우 정확한 표현 전달 보다는 의미 전달이 목적이다. 하지만 기존 통신은 데이터의 의미 및 업무와 무관하게 bit로 변환된 데이터를 정확하고 효율적으로 전송하는 것에 초점을 둔다. 이에 반해 시맨틱 통신의 목표는 업무 수행에 필요한 정보를 추출하고 불필요한 정보를 걸러냄으로써 중요 정보의 효율적 전송을 목표로 한다.

시맨틱 통신의 가장 두드러지는 장점 두 가지는 핵심 정보만을 전송하여 주파수 자원을 효율적으로 사용할 수 있다는 것과 정확한 bit를 복원하는 기존 통신 대비 낮은 신호 대 잡음 비 (signal-to-noise ratio, SNR)를 가지는 환경에서 보다 정확한 의미전달이 가능하다는 점이다 [1]. 이를 위해 시맨틱 통신에서 가장 중요한 것은 적절한 시맨틱 정보를 원본 데이터로부터 추출하는 것이다. 최근 딥 러닝 기술의 발전으로 딥 러닝을 이용한 시맨틱 정보 추출이 가능하게 되었고 이를 이용한 시맨틱 연구가 활발하게

이루어지고 있다. 본 논문에서는 시맨틱 통신의 연구 동향을 분석하고 해결해야 할 문제들을 제안한다.

II. 본론

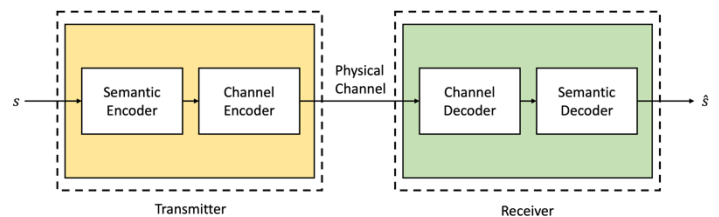


그림 1. 시맨틱 통신의 구조

본 단락에서는 먼저 딥러닝 시맨틱 통신에 관한 연구들을 소개하고 무선 자원의 효율적 사용에 관한 문제를 다룬 연구들을 통찰한다. 그 후, 앞으로 해결해야 할 문제들을 제안한다.

A. 딥 러닝 기반 시맨틱 통신

그림 1은 시맨틱 통신의 전반적인 구조를 나타낸다. 그림 1에서 s 는 송신해야 하는 원본 데이터, \hat{s} 는 수신한 복원 데이터를 의미한다. 시맨틱 인코더에서는 원본 데이터로부터 업무 수행에 필요한 정보를 추출한다. 추출된 정보는 채널 인코더, 채널, 채널 디코더를 순차적으로 거친 후 시맨틱 디코더에서 복원된다. 시맨틱 인코더/디코더는 딥러닝을 이용하여 설계하며 텍스트, 이미지, 스피치 각각의 데이터 또는 여러 종류의 데이터에 대한 시맨틱 정보를 추출하는 multi-modal 시맨틱 인코더/디코더가 개발되었다[2]. 논문 [1]에서는 시맨틱 오류를 줄이고

시스템 용량을 최대화하는 동시에 시맨틱 레벨에서의 복원 성능을 높이기 위하여 시맨틱 코딩과 채널 코딩을 동시에 설계하였다.

기존 Autoencoder 기반 무선 통신 기술과 두드러지는 차이점은 입력 bits 와 출력 bits 간 에러를 최소화 하는 Autoencoder 기술과 달리 시맨틱 인코더/디코더는 원본 데이터와 수신된 데이터 사이의 의미의 유사도를 최소화 한다는 점이다. 따라서 설계한 시맨틱 인코더/디코더의 성능을 측정하기 위해서는 원본 데이터와 수신된 데이터 사이의 시맨틱 유사도를 측정해주는 metric 이 필요하다. 논문 [1]에서는 두 문장 사이의 시맨틱 유사도를 비교해주는 새로운 metric 으로 sentence similarity 를 제안하였고 수학적 식은 아래와 같다.

$$\text{match}(s, \hat{s}) = \frac{\mathbf{B}_{\Phi}(s) \cdot \mathbf{B}_{\Phi}(\hat{s})^T}{\|\mathbf{B}_{\Phi}(s)\| \|\mathbf{B}_{\Phi}(\hat{s})\|}$$

여기서 \mathbf{B}_{Φ} 는 사전 학습된 BERT 모델을 의미한다[3].

이외에도 이미지간 또는 스피치간 시맨틱 유사도를 비교해주는 metric 들이 존재한다[4]. 이를 이용하면 시맨틱 유사도를 수치적으로 표현이 가능하게 되어 시맨틱 인코더 및 디코더 학습에 효과적으로 사용할 수 있다.

B. 시맨틱 통신에서의 자원 할당 문제

시맨틱 통신에서는 bit 로 된 데이터를 전송하는 것이 아니라 실수 값으로 이루어진 feature 를 전송하기 때문에 기존 통신의 자원할당 방법을 적용할 수 없어 새로운 정의를 도입해 문제를 해결할 필요가 있다. 논문 [5]에서는 텍스트 시맨틱 통신에서의 semantic transmission rate 과 semantic spectral efficiency 를 정의하고 이를 이용해 시맨틱 통신에서의 자원할당 문제를 해결한다. 또한 bit-based spectral efficiency 를 semantic spectral efficiency 로 변환하는 방안을 고안하여 기존 통신과 시맨틱 통신의 성능을 비교하였다. 그 결과 채널 수, transmit power 에 관계 없이 시맨틱 통신이 기존 통신보다 더 높은 semantic spectral efficiency 를 보여주었다.

C. 해결해야 할 문제들

Ref	Text data	Image data	Resource Allocation
[1],[6]	O		
[7]		O	
[2]	O	O	
[5]	O		O

표 1. 시맨틱 통신 연구 현황

표 1 은 최근 시맨틱 통신의 연구 동향에 대해 나타낸다. 텍스트와 이미지 데이터 각각에 대한 여러 연구들이 제안되어왔지만, 각 기술마다 사용한 유사도 metric 이 통일되어있지 않아 통합된 성능 벤치마크가 불가능하다. 따라서 여러 기술에 포괄적으로 사용될 수 있는 일반적인 유사도 metric 고안이 필요하며 metric 의 신뢰성을 위해 작업 수행 등을 통한 검증이 수반되어야 할 것으로 보인다.

무선 자원 할당 기술은 논문 [5]에서 제안하였는데, 해당 논문에서는 모든 채널의 주파수 대역폭이 동일하고 하나의 채널은 하나의 사용자에게만 할당된다는 가정하에 자원할당 문제를 해결하였다. 그러나 각 사용자마다 전달하고자 하는 데이터의 크기가 다르기 때문에 하나의 사용자가 요구에 따라 여러 채널을 할당 받을 수 있는 문제들에 대한 해결책이 필요하다.

또한, 표 1 에 기재된 바와 같이 이미지와 텍스트 데이터 모두를 고려한 multi-modal 시맨틱 통신에 대한 자원할당

문제가 다뤄지지 않았음을 확인할 수 있다. 따라서 여러가지 데이터를 한번에 다루는 multimodal 시맨틱 통신에 대한 연구가 필요할 것으로 보이며, 다양한 유형의 데이터에 대한 자원할당 문제 역시 필요할 것으로 예상된다.

III. 결론

본 논문에서는 시맨틱 통신의 연구동향을 소개하고 해결해야 할 문제점 및 연구 방향을 제시하였다. 시맨틱 통신은 원본 데이터를 그대로 전송하지 않고 작업 수행에 필요한 데이터만 전송한다는 점에서 기존 통신과 차이가 있다. 이러한 특성 때문에 시맨틱 유사도를 측정하는 metric 이나 semantic spectral efficiency 와 같은 새로운 정의가 필요하다. 이를 정형화하고 시맨틱 통신을 발전시킨다면 미래에 새로운 통신 기술로써 유용하게 사용될 수 있을 것으로 기대된다.

ACKNOWLEDGMENT

본 논문은 과학기술정보통신부 및 정보통신기획평가원의 대학 ICT 연구센터지원사업의 연구결과와(IITP-2021-0-02048) 2022 년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2018-0-00958, 타겟 감지 및 정밀 자세 제어를 이용한 전기적/기계적 드론 범포밍 기술 개발).

참 고 문 헌

- [1] H. Xie, Z. Qin, G. Y. Li and B. -H. Juang, "Deep Learning Enabled Semantic Communication Systems," in *IEEE Transactions on Signal Processing*, vol. 69, pp. 2663-2675, 2021
- [2] H. Xie, Z. Qin, X. Tao and K. B. Letaief, "Task-Oriented Multi-User Semantic Communications," in *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 9, pp. 2584-2597, Sept. 2022
- [3] Devlin, Jacob, et al. "BERT: Pre-training of deep bidirectional transformers for language understanding." *arXiv preprint arXiv:1810.04805*, 2018
- [4] Qin, Zhijin, et al. "Semantic communications: Principles and challenges." *arXiv preprint arXiv:2201.01389*, 2021
- [5] L. Yan, Z. Qin, R. Zhang, Y. Li and G. Y. Li, "Resource Allocation for Text Semantic Communications," in *IEEE Wireless Communications Letters*, vol. 11, no. 7, pp. 1394-1398, July 2022
- [6] Q. Zhou, R. Li, Z. Zhao, C. Peng and H. Zhang, "Semantic Communication With Adaptive Universal Transformer," in *IEEE Wireless Communications Letters*, vol. 11, no. 3, pp. 453-457, March 2022
- [7] D. Huang, X. Tao, F. Gao and J. Lu, "Deep Learning-Based Image Semantic Coding for Semantic Communications," *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 1-6