

차세대 통신을 위한 생성 모델 기반 시멘틱 통신 연구 동향

임채문, 김중헌

고려대학교

anscodla0314@korea.ac.kr, joongheon@korea.ac.kr

A Trends on Generative-AI based Semantic Communication For Next-generation Communication systems

Chaemoon Im, Joongheon Kim

Korea Univ.

요약

고전 통신 이론은 송신 정보량이 증가함에 따라 대역폭 증가 및 비트와 정보 간 결합으로 인한 의미의 왜곡 등의 문제를 갖고 있다. 이에, 해당 문제를 해결하기 위해 비트가 아닌 신호의 의미 전달을 중심으로 하는 시멘틱 통신이 제시되었다. 그러나, 초기 단계인 시멘틱 통신은 정의 및 평가 지표 부재에 따른 어려움을 겪고 있다. 본고는 이러한 시멘틱 통신의 특징 및 연구 동향을 탐색하고, 시멘틱 통신의 향후 연구 방향에 대해 제시한다.

I. 서론

고전적 통신 이론은 비트의 왜곡 없는 정확한 정보 전달을 목적으로 지속적으로 발전해 왔다. Shannon은 채널의 상태가 정해졌을 때 최대로 전달할 수 있는 정보량에 대한 이론적 한계를 제시하였고, 기존 통신 이론은 해당 이론적 한계에 근접할 수 있는 방법을 제시하며 발전하였다 [1]. 그러나 고전적 통신은 보내야 하는 데이터의 양이 증가함에 따라 추가적인 대역폭을 필요로 한다. 이러한 전송량과 대역폭의 상관관계는 Shannon이 제시한 정보 이론의 틀을 기반으로 한 고전적 통신 이론에서는 자명한 사실로 여겨져 왔다. 또한, 고전 통신의 경우 비트의 일부만이 손실됨에 따라 전체 정보가 손실되는 등 비트와 정보 간에 강한 결합성이 존재하는 등의 문제가 있다. 이는 고전 통신이 비트를 최대한 손실 없이 전송하는 것에 주안점을 두고 있기 때문이다.

이에, 해당 문제를 해결하기 위해서 시멘틱 통신이 제기되었다. 시멘틱 통신은 비트열을 송수신하는 것이 아닌, 신호에 담겨 있는 의미를 전송함으로써 해당 문제를 해결한다. 이러한 시멘틱 통신은 주어진 데이터로부터 의미 정보를 효과적으로 추출하는 머신 러닝 기법과 결합하여, 기존의 고전 통신 이상의 송신 속도를 달성하기 위해 발전하고 있다. 본고는 이러한 시멘틱 통신의 특징 및 연구 동향에 대해 논하고, 마지막으로 현재 제시되는 시멘틱 통신의 한계 및 연구 방향에 대해 논한다.

II. 의미(Semantic)의 정의

시멘틱 통신은 고전 통신 이론과 다르게 비트를 보내는 것보다, 해당 비트를 통해 어떤 정보를 전달하려 했는지를 보다 중점적으로 본다. 또한 시멘틱 통신은 정보의 전송 단위를 비트가 아닌 의미로 보고, 결론적으로 통신을 통해 손실 없이 전달되어야 하는 것은 비트 그 자체가 아니라 의미라는 점에 집중한다.

신호의 의미가 무엇인지, 통신 과정에서 어떻게 의미가 왜곡되는지에 대해서는 현재까지 정확하게 합의된 정의가 존재하지 않으며, 신호의

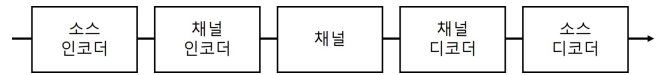


그림 1. 기존의 통신 방식 개념도.



그림 2. 시멘틱 통신 개념도.

의미를 어떻게 정의할지에 대한 다양한 논의가 진행되고 있다. .

Y. Shao는 문자열에서 어떻게 언어 및 의미망이 형성되는지에 대한 과정을 수학적으로 모델링하였고, 해당 과정을 통해 시멘틱 채널, 시멘틱 정보의 왜곡 등을 정의하였다 [2]. 그에 따르면, 시멘틱 정보가 형성되기 위해서는 문자열의 조합에 사전 지식이 결합되어야 하고, 따라서 그의 모델링에서는 사전 지식이 인코딩과 디코딩의 과정에 필수적으로 포함된다.

E. Beck은 Y. Shao의 논의를 발전시켜, 어떻게 시멘틱 정보가 수학적으로 정의될 수 있을지 탐구하였다 [3]. 그러나 그는 Y. Shao와 다르게 시멘틱 정보를 전달하기 위해 문자열이 구성되는 과정을 시멘틱 채널이라고 정의하였고, 해당 채널에서 최적의 정보 전달 방식을 설계하였다.

한편, L. Hu는 지식 그래프(Knowledge Graph)를 기반으로 시멘틱을 정의하였다 [4]. 그는 시멘틱 디코딩이란 송신된 정보를 보고 이를 생성한 지식 그래프를 추론하는 것이라고 보았으며, 이 과정에서 생기는 오류가 시멘틱 노이즈라고 보았다. 해당 체계 하에서 그는 그래프 임베딩을 통해 시멘틱 디코딩 알고리즘을 설계하였고 이것이 시멘틱 노이즈에 강인함을 검증하였다.

이렇듯 시멘틱의 정의에 대해서는 아직도 활발한 논의가 진행 중이다. 그러나 많은 논의에서 시멘틱은 원래의 정보를 가공하여 얻을 수 있는 이차 랜덤변수로 정의되며, 해당 정의 위에서 시멘틱을 처리하기 위한 여러 방법론들이 논의되고 있다. 또한, 시멘틱 인코딩 및 디코딩을 진행하는 데

있어 사전 지식의 필요성 역시 모든 논의에서 공통적으로 발견되는 사항이다.

III. 시멘틱 통신 연구 동향

앞서 논의했듯이, 신호의 시멘틱이 무엇인지에 대한 정확한 합의점은 아직 존재하지 않는다. 그러나 많은 경우 시멘틱 통신은 신호로부터 의미를 추출하기 위해 머신 러닝 모델을 사용하여 인코더와 디코더를 구성한다. 여기서 시멘틱이란, 원래 신호를 머신 러닝 모델을 통해 처리하였을 때의 그 출력 자체를 의미하는 경우가 많다. 특히 그 중에서도, 대용량의 비트를 더 적은 비트로 임베딩하는 표현 학습(Representation Learning) 분야는 시멘틱 통신에 주로 사용되는 머신 러닝 분야이다 [5].

고전 통신은 그림 1에서 서술되어 있듯이, 정보를 압축하기 위한 소스 인코더 및 디코더, 채널의 영향을 극복하기 위한 채널 인코더 및 디코더로 이루어져 있다. 반면, 시멘틱 통신은 해당 구조에, 신호의 시멘틱 정보를 추출하기 위한 시멘틱 인코더와 시멘틱 정보를 통해 원래의 신호를 추출해 내기 위한 시멘틱 디코더를 도입하는 것이 특징이다. 해당 시멘틱 인코더 및 디코더는 흔히 머신러닝 모델로 구현된다.

H. Xie는 트랜스포머 모델을 사용하여, 연속된 비트에 포함되어 있는 시멘틱을 추출 및 시멘틱을 통해 원래 데이터를 원복하는 구조를 제안하였다 [6]. 저자는 트랜스포머의 특징을 사용하여 시멘틱 코딩뿐만 아니라 소스/채널 코딩까지 통합적으로 수행할 수 있는 구조를 제안하였다는 것 역시 해당 구조의 장점이다.

L. Guo는 Variational Autoencoder(VAE) 및 Diffusion 모델을 사용하여 시멘틱 통신을 수행하였다 [7]. 그는 대역폭의 한계를 VAE를 통한 정보 압축으로 극복하였고, 이를 통해 실시간 이미지 및 동영상 전송 프레임워크를 제시하였다. 해당 연구는 실제로 시멘틱 통신을 통해 한정된 대역폭 문제를 극복할 수 있는 실마리를 야기하였다는 점에서, 시멘틱 통신의 특징을 보여 준다.

최근 급속도로 발전하고 있는 LLM을 활용한 시멘틱 통신 연구 역시 존재한다. S. Salehi는 LLM의 강력한 언어 표현 능력이 통신에 어떻게 적용될 수 있을지를 분석하였다 [8]. 해당 연구에서는 지식 그래프(Knowledge Graph)를 통해 전처리된 정보를 LLM을 통해 토큰화하고, 이를 통해 정보 송신량을 감소시켰다.

특히 LLM과 관련한 연구는 멀티모달 데이터의 시멘틱에 관한 연구로 이어지고 있다. Y. Zhao는 LLM을 사용하여 이미지 정보를 비트로 바꾸고, 이를 전송함으로써 이미지를 전송하는 데 드는 비트의 수를 대폭 감소시켰다 [9]. 이렇듯 시멘틱 통신 연구는 다수의 모달리티가 조합된 데이터의 송신에서도 유용하게 사용될 수 있는 가능성을 가지고 있다.

해당 연구에서 공통적으로 발견되는 시멘틱 통신의 장점은, 낮은 Signal-to-Noise Ratio(SNR)에서도 안정적인 전송이 가능하다는 것이다. 이는 시멘틱 통신에서 전달되는 것이 비트 그 자체가 아니라 신호의 의미이기 때문으로, 고전 통신에서 발생하는 비트와 정보의 강한 결합성이 사라졌음을 의미한다. 또한 시멘틱 통신은 전체 정보를 보내는 것이 아니라 신호의 시멘틱만 보내므로 전송 용량이 급격히 감소하게 되지만, 송수신기에서 해당 시멘틱을 추출 및 시멘틱을 통한 원래 신호의 복원에 보다 더 복잡한 과정이 필요하다는 특징이 생긴다.

IV. 향후 연구 방향

이처럼 시멘틱 통신은 지속적으로 연구되고 있는 머신 러닝 분야와 결합하여 성장하고 있지만, 아직 해결해야 할 문제가 남아 있다. 가장 우선적

인 것은 시멘틱에 대한 정의가 확립되지 않았다는 것이다. Shannon의 정보이론에 근거한 정의가 고전 통신의 한계 및 발전 방향을 설정할 수 있는 지표가 되었듯이, 시멘틱 통신 역시 그에 준하는 정의가 필요할 것이다. 또한 Bit Error Rate(BER)처럼 시멘틱의 전달 양상을 평가할 수 있는 평가 지표의 부재 또한 서로 다른 시멘틱 통신 기술을 비교하고 분석하는데 걸림돌이 되고 있다. 현재 시멘틱 통신 기술은 문자열, 이미지 등의 정보를 전송함에 있어 해당 도메인에 특화된 BLEU(Bilingual Evaluation Understudy Score), Peak Signal-to-Noise Ratio(PSNR) 등의 평가 지표를 사용하고 있으나, 일반적인 정보 전달 방식을 평가할 수 있는 평가 지표가 부재한 상황이다. 따라서, 서로 다른 시멘틱 통신 기술을 비교할 수 있는 정량적인 지표가 논의되어야 한다.

ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원-대학ICT연구센터(ITRC)의 지원을 받아 수행된 연구임 (IITP-2025-RS-2024-00436887).

참 고 문 헌

- [1] T. M. Cover and J. A. Thomas. Elements of Information Theory. Wiley-Interscience, USA. 2006.
- [2] Y. Shao, Q. Cao and D. Gündüz, "A Theory of Semantic Communication," in IEEE Transactions on Mobile Computing, vol. 23, no. 12, pp. 12211-12228, Dec. 2024.
- [3] Beck, Edgar, Carsten Bockelmann, and Armin Dekorsy. "Semantic information recovery in wireless networks." Sensors, no. 23, vol. 14, pp. 6347-6370, Jul. 2023.
- [4] L. Hu, Y. Li, H. Zhang, L. Yuan, F. Zhou and Q. Wu, "Robust Semantic Communication Driven by Knowledge Graph," in Proc. International Conference on Internet of Things: Systems, Management and Security (IOTSMS), Milan, Italy, Nov. 2022, pp. 1-5.
- [5] Y. Bengio, A. Courville and P. Vincent, "Representation Learning: A Review and New Perspectives," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 8, pp. 1798-1828, Aug. 2013.
- [6] 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, Apr. 2021.
- [7] L. Guo, W. Chen, Y. Sun, B. Ai, N. Pappas and T. Quek, "Diffusion-Driven Semantic Communication for Generative Models with Bandwidth Constraints," in IEEE Transactions on Wireless Communications, Early Access, 2025.
- [8] Salehi, Shavbo, Melike Erol-Kantarci, and Dusit Niyato. "LLM-Enabled Data Transmission in End-to-End Semantic Communication." in Proc. International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Lucca, Italy. Sep. 2024.
- [9] Y. Zhao, Y. Yue, S. Hou, B. Cheng and Y. Huang, "LaMoSC: Large Language Model-Driven Semantic Communication System for Visual Transmission," in IEEE Transactions on Cognitive Communications and Networking, vol. 10, no. 6, pp. 2005-2018, Dec. 2024.