

시맨틱 통신을 위한 각종 지표에 관한 연구 동향

김주영, 홍성훈, 김가현, 최성진, 이승찬, 조성래
중앙대학교 컴퓨터공학과

{jykim, shhong, ghkim, sjchoi, sclee}@uclab.re.kr, srcho@cau.ac.kr

A state-of-the-art on Metrics for Semantic Communication

Juyoung Kim, Seonghun Hong, Gahyun Kim, Seongjin Choi, Seungchan Lee, and
Sungrae Cho

Department of Computer Science and Engineering, Chung-Ang University

요약

인코딩 된 비트를 송수신하는 전통적인 통신 메커니즘과 달리 데이터에서 의미를 추출하여 이를 송수신하는 시맨틱 통신이 6G 차세대 통신의 핵심적인 기술로 부상하고 있다. 기존의 통신 메커니즘과 달리 의미를 다루는 기술이므로 각종 평가 또한 기존의 비트 기반이 아닌, 의미를 기반으로 하는 지표를 사용해야 한다. 본 논문은 시맨틱 통신의 성능 평가, 학습 과정의 목적함수 등에 사용되는 시맨틱 유사도, 스펙트럼 효율성, 시맨틱 엔트로피를 정량적으로 표현하기 위한 측정 지표의 필요성과 제안된 각종 지표에 관한 연구 동향을 조사한다.

I. 서론

지난 수십년간 무선 통신은 심볼, 비트를 정확하고 효과적으로 전송하는 데에 초점을 맞추고 발전하였으며 Shannon Limit 에 점진적으로 다가가고 있다. 이는 기존의 패러다임은 상한선에 다가갔다는 것을 말하며 기존의 패러다임으로는 더 이상의 발전이 어렵다는 점을 시사한다.

그러나 최근 메타버스[1], 차량 네트워크[2], 자율 UAV(Unmanned Aerial Vehicle)[3] 등 상당히 많은 데이터 전송을 필요로 하는 다양한 서비스가 등장하며 고성능 통신에 대한 요구가 많아졌다. 시맨틱 통신은 이러한 각종 산업의 높은 통신 요구 사항을 만족시킬 기술로서 많은 주목을 받고 있다. Shannon 기반의 통신 패러다임은 데이터를 비트로 인코딩하여 송신하는 반면 시맨틱 통신은 데이터에서 의미를 추출하고 그 의미를 분석하여 불필요하거나 관련이 없는 데이터를 제거한다. 이러한 과정을 통해 의미를 보존하면서 데이터를 압축하여 초저지연, 스펙트럼 효율성 등을 동시에 만족할 수 있다.

그럼에도 불구하고 시맨틱 통신을 위한 이론적 프레임워크는 아직 확립되지 않았다 [4]. 특히 정보량, 스펙트럼 효율성 등을 정량적으로 측정하는 지표는 가장 기초적인 문제이다. 기존의 Shannon 기반의 통신 패러다임은 소스 심볼의 통계적 지식에 기반하여 비트를 생성하여 시맨틱 정보를 반영하지 않는다. 따라서 기존에 사용하던 각종 지표는 시맨틱 통신에 적용이 불가능하다. 따라서 시맨틱 통신 관점에서 재고해야 할 필요가 있다. 본 논문은 시맨틱 통신을 위한 각종 수학적 이론 및 표현, 시맨틱 정보이론을 소개한다.

II. 본론

시맨틱 통신의 인공지능 모듈을 학습하기 위해 손실함수를 최소화하는 방법을 통해 파라미터 업데이트를 진행한다. 이때 손실함수에 송신한 데이터와 수신한 데이터와의 차이, 즉 두 문장의 유사도를 포함하면 효과적인 학습이 가능하다. BLEU는 단순히 각 단어들의 일치 여부를 기준으로 평가한다. 따라서 같은 의미의 다른 단어를 비교하는 경우, 유사하지 않다고 측정하는 한계가 있다. [5]에서는 BVCS (BERT Vector-Cosine Similarity)라는 지표를 제안하고 유사도를 평가한다. 이는 BERT 모델을 활용하여 앞서 언급한 한계를 해결한다. BERT는 딥러닝 (DL) 기반의 강력한 언어모델로, 주어진 텍스트로부터 문맥 정보를 반영한 임베딩 벡터를 생성한다. 이러한 과정을 통해 문장에서 시맨틱 정보를 추출한다. 시맨틱 정보에 해당하는 벡터를 생성한 뒤에, 각 벡터의 유사도를 구하기 위해 코사인 유사도를 사용한다. BVCS 값은 -1 부터 1의 값을 가지며 1에 가까울수록 높은 유사도를 의미한다.

기존 통신의 스펙트럼 효율은 1 초당 1Hz 대역폭에서 전송되는 비트 수로 측정한다. 그러나 이러한 방식은 의미가 아닌, 통계적 특성에 따라 생성되기 때문에 시맨틱 통신을 평가할 수 없다. 시맨틱 통신을 위한 각종 스펙트럼 효율 지표 또한 제안되었다. [6]는 시맨틱 스펙트럼 효율 S-SE(Semantic Spectral Efficiency)를 정의하기 위해 시맨틱 정보의 기본 단위 sut (Semantic Unit)를 도입한다. 이후 텍스트 데이터에서 평균적인 의미 정보량과 문장 길이를 기반으로, 실제 전송되는 시맨틱 심볼들이 담고 있는 의미량과 전송 경로에서 유지된 의미 유사도를 종합적으로 고려하여 시맨틱 스펙트럼 효율을 산출한다. 이러한 접근은 단순히 데이터의 양이 아닌, 데이터의 본질적인 의미가 얼마나

잘 보존되었는지를 고려한다. 따라서 딥러닝 기반 시맨틱 통신 모델에서 채널 조건과 학습된 구조에 따라 의미 전달 성능이 달라질 수 있는 환경에서 성능 변화를 정량적으로 반영하는 핵심 지표로 활용될 수 있다.

또한 시맨틱 통신에서는 정보량 자체를 정량화 하는 지표로서 시맨틱 엔트로피(Semantic Entropy) 개념이 제안되었다. 이는 전통적인 Shannon 정보 이론에서의 엔트로피 개념을 확장한 것으로 문장이나 메시지가 얼마나 많은 의미 정보를 내포하는지 측정한다. 전통적인 Shannon 엔트로피는 데이터의 통계적 분포에 기반한 확률에 따라 정보량을 계산한다. 하지만 시맨틱 엔트로피는 논리적 확률 또는 의미의 Logical probability 를 기반으로 정의되며, 단순히 문자나 비트의 조합이 아닌 문장의 의미가 주어진 맥락 또는 지식 내에서 얼마나 타당한지에 따라 정보량을 산정한다 [7].

이러한 지표들은 성능 평가뿐만 아니라 학습 과정에도 활용될 수 있다.

III. 결론

시맨틱 통신은 단순한 데이터의 전달을 넘어, 의미 중심의 정보 교환을 통해 통신의 효율성과 지능화를 동시에 실현할 수 있는 차세대 통신 패러다임으로 주목받고 있다. 본 논문에서는 시맨틱 통신의 성능을 정량적으로 평가하고, 인공지능 기반 학습 과정에 적용할 수 있는 다양한 시맨틱 평가 지표에 대해 고찰하였다. 특히 문장 간 의미 유사도를 평가하는 BVCS, 전송 효율을 평가하는 시맨틱 스펙트럼 효율, 정보량을 측정하는 시맨틱 엔트로피 등은 기존 Shannon 기반의 기술로는 설명할 수 없었던 시맨틱 통신의 특성을 반영한 지표들이다. 이러한 지표들은 단순한 문자 또는 비트 단위의 평가를 넘어 문맥, 지식 등을 고려한 정보 평가를 가능하게 하며 향후 시맨틱 통신 기술의 성능 향상과 실용화에 중요한 역할을 할 것이다.

ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로
정보통신기획평가원-대학 ICT 연구센터(ITRC)의 지원을 받아
수행된 연구임(IITP-2025-RS-2022-00156353, 50% / IITP-
2025-RS-2024-00436887, 50%)

- [3] W. J. Yun *et al.*, "Cooperative Multiagent Deep Reinforcement Learning for Reliable Surveillance via Autonomous Multi-UAV Control," in *IEEE Transactions on Industrial Informatics*, vol. 18, no. 10, pp. 7086–7096, Oct. 2022.
- [4] T. M. Getu, G. Kaddoum and M. Bennis, "Making Sense of Meaning: A Survey on Metrics for Semantic and Goal-Oriented Communication," in *IEEE Access*, vol. 11, pp. 45456–45492, 2023S.
- [5] S. Kabir and M. A. Razzaque, "BVCS: An Integrated BERT Vector-Cosine Similarity Metric for Text Semantic Communication," *2024 6th International Conference on Sustainable Technologies for Industry 5.0 (STI)*, Narayanganj, Bangladesh, 2024, pp. 1–6.
- [6] 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.
- [7] G. Xin, P. Fan, and K. B. Letaief, "Semantic Communication: A Survey of Its Theoretical Development" *Entropy* 2024, 26, 102.

참 고 문 현

- [1] S. Park, H. Baek and J. Kim, "Spatio-Temporal Multi-Metaverse Dynamic Streaming for Hybrid Quantum-Classical Systems," in *IEEE/ACM Transactions on Networking*, vol. 32, no. 6, pp. 5279–5294, Dec. 2024.
- [2] W. J. Yun, D. Kwon, M. Choi, J. Kim, G. Caire and A. F. Molisch, "Quality-Aware Deep Reinforcement Learning for Streaming in Infrastructure-Assisted Connected Vehicles," in *IEEE Transactions on Vehicular Technology*, vol. 71, no. 2, pp. 2002–2017, Feb. 2022.