

강인한 심층 그래프 생성 모델 설계

박진덕¹, Cong Tran^{1,2}, 신원용¹

¹연세대학교, ²단국대학교

jindeok6@yonsei.ac.kr, trancong208@gmail.com, wy.shin@yonsei.ac.kr

Design of Robust Deep Graph Generative Models

Jin-Duk Park¹, Cong Tran^{1,2}, Won-Yong Shin¹

¹Yonsei University, ²Dankook University

요약

대부분의 실제 네트워크 데이터는 부분적으로만 관찰 가능한 그래프(이하 불완전한 그래프)로 알려져 있다. 예를 들어, 소셜 네트워크로부터 데이터를 수집 시 비공개 계정 등의 이유로 우리는 완전한 네트워크 구조를 파악할 수 없다. 반면, 그래프에서의 심층 생성 모델에 대한 연구가 활발히 이루어지고 있는데, 네트워크에서의 불완전하게 관찰되는 문제를 해결하고자 강인한 심층 그래프 생성 모델 설계를 제안한다. 제안하는 모델은 한 쌍의 불완전한 그래프를 통해 완전한 그래프의 연결 관계를 추측한 후, 순환 신경망에 기반한 생성 모델을 적용하여 완전한 그래프를 생성한다. 본 논문에서는 제안한 모델의 장건한 결과를 다양한 인공적 네트워크 및 실제 네트워크 데이터셋을 활용해 입증한다.

I. 서론

대부분의 실제 네트워크 데이터는 부분적으로만 관찰 가능한 불완전한 그래프로 이루어진다. 본 논문에서는 불완전한 그래프의 쌍을 training 데이터셋으로 활용해 완전한 그래프의 구조를 추측하고, 추측한 그래프의 구조와 비슷한 다양한 그래프 데이터를 생성하는 강인한 심층 그래프 생성 모델을 제안한다.

II. 본론

불완전한 그래프 쌍을 G_i ($i = 1, 2$)라고 하고, 추측한 완전한 그래프를 \widehat{G}_i , \widehat{G}_i 의 인접행렬을 A_i , A_i 를 1 차원 벡터로 변형한 것을 a_i , 생성 모델이 샘플링하는 그래프의 간선 확률 분포를 φ 라고 정의한다.

제안한 모델은 크게 1) 마스킹 추측 모듈, 2) 교차 학습 모듈, 3) 그래프 생성 모듈 이렇게 세 가지 모듈로 나누어진다. 마스킹 추측 모듈에서는 [1]에서의 방법에 따라 그래프 쌍으로부터 각각의 그래프에서 어느 간선이 마스킹 되었는지를 추측한다.

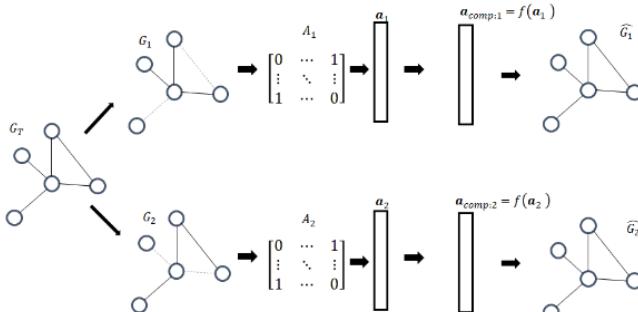


Figure 1. 교차 학습 모듈 다이어그램

Figure 1에서 확인할 수 있듯이 교차 학습 모듈에서는 복원 심층 신경망 $f(\cdot)$ 을 통해, 마스킹 추측 모듈에서 추측된 마스킹된 간선에 대해서 서로의 간선 값을 보상하여 완전한 그래프 구조를 예측한다. 또한, 그래프 생성 모듈에서는 자동 회귀(auto-regressive) 간선 단위 생성 모델 중 하나인 GraphRNN(그래프 순환 신경망) [2] 기반 생성 모델을 이용하여, 예측된 완전한

그래프 \widehat{G}_i 의 간선 분포 φ 를 학습하고, φ 에 기반해서 그래프를 간선 단위로 샘플링한다.

III. 결론

본 논문에서는 불완전한 그래프의 쌍을 이용해서 서로 간의 연결정보를 보상한 뒤, 예측된 완전한 그래프를 기반으로 그 분포를 학습해서 다양한 완전한 그래프들을 강인하게 생성하는 실험을 진행한다. 성능 척도로는 MMD (Maximum Mean Discrepancy)를 사용하였고, Table 1에서와 같이 실제 완전한 그래프(즉, ground truth)와 모델을 통해 생성된 그래프들 간 분포를 비교하여 제안한 모델의 강인한 결과를 확인할 수 있다.

| 데이터셋 | MMD score | | |
|----------------|-------------|-------------|-------------|
| | 10% masking | 20% masking | 30% masking |
| Grid Network | 0.03 | 0.05 | 0.08 |
| Karate Network | 0.12 | 0.17 | 0.20 |

Table 1. 데이터셋에 따른 MMD score

ACKNOWLEDGMENT

This work was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (HI20C0127), and by the Yonsei University Research Fund of 2020 (2020-22-0101).

참고 문헌

- [1] Z. Xia and A. Chakrabarti, "Training image estimators without image ground-truth." In Proceedings of 33rd Conference on Neural Information Processing Systems, Vancouver, Canada, December 2019, pp. 1-11.
- [2] J. You, R. Ying, X. Ren, W. L. Hamilton, and J. Leskovec, "GraphRNN: Generating realistic graphs with deep auto-regressive models," In Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, July 2018, pp. 1-12.