

LSTM-DCCNN GAN을 사용한 5G Traffic 생성

김대겸, 고명진, 문성우, 김성현, 천경열*, 박승근*, 김윤배*, 윤현구**, 최용훈

광운대학교, *한국전자통신연구원, **명지전문대학

0913ktg@gmail.com, audwls970526.dv@gmail.com, msw3322@gmail.com, waverdeep@kw.ac.kr, *kycheon@etri.re.kr, *seungkp@etri.re.kr, *doko9gum@etri.re.kr, **hgyoon@mjc.ac.kr, yhchoi@kw.ac.kr

Generation of 5G Traffic Using LSTM-DCCNN GAN

Dae-gyeom Kim, Myeong-jin Ko, Sung-woo Moon, Sung-hyun Kim,

Kyung-Yul Cheon*, Seungkeun Park*, Yunbae Kim*, Hyungoo Yoon**, Yong-Hoon Choi

Kwangwoon University, *ETRI, **Myongji College

요약

사설 5G망을 구축하기 위해서 정확한 요구 대역폭을 알아야 한다. 요구 대역폭을 계산하려면 5G 트래픽 소요량을 알아야 한다. 기존의 확률기반 생성 모델은 다양한 패턴이 있는 트래픽을 생성하기 힘들다. 본 논문은 신뢰도 높은 5G 트래픽 생성을 위해 generative adversarial networks (GAN) 모델을 사용하여 실제 5G 트래픽과 유사한 트래픽을 생성했다. 본 논문에서는 생성자로 순환신경망 (LSTM: long short-term memory)을 사용했고, 판별자로 인과확장 합성곱 (DCCNN: dilated causal convolution neural network)을 사용했다. 모델의 성능 평가 지표로는 JSD (Jensen-Shannon Divergence)와 MMD (Maximum Mean Discrepancy)를 사용했다. 본 연구는 5G 트래픽 생성을 통해 정확한 사설 5G망 요구 대역폭을 산출할 수 있는 모델 설계를 목표로 하고 있으며, 인공 신경망 기반 5G 트래픽 생성기가 실제 트래픽과 유사한 트래픽을 생성할 수 있다는 것을 확인하였다.

I. 서론

사설 5G망 (Private 5G Network)이란 기업이나 대학 등 특정 주체가 5G 망을 직접 구축하여 사용하는 배타적 목적의 사설망으로 MNO (Mobile Network Operator)들이 구축하는 공중 5G망 (Public 5G Network)과 상반되는 개념이다. 사설 5G망을 구축하면 기존 LAN의 제약 (짧은 커버리지와 망 안정성, 보안 취약성)에서 벗어날 수 있고, 5G의 대용량, 초저지연, 초연결성을 활용하여 다양한 기업용 응용들을 새로이 구축하거나, 기존의 응용들을 최적화시킬 수 있는 큰 장점이 있다. 사설 5G 망을 구축하기 위해 정확한 요구 주파수 소요량 예측이 필요하다. 주파수 소요량 예측에 5G 트래픽 데이터가 사용되는데, 현재 사용되는 확률기반 생성 모델이 생성한 트래픽은 다양한 형태의 실제 트래픽의 모습과 차이가 있다. 본 연구에서는 시계열 데이터인 5G 트래픽을 실제 트래픽의 모양과 유사하게 생성할 수 있는 딥러닝 기반 모델을 설계한다.

II. 본론

가. 데이터 수집

본 연구에서 사용한 5G 트래픽 데이터셋은 모바일 모니터링 앱 (G-NetTrack Pro)를 사용하여 수집한 시계열 데이터로 github에 공개되어 있다 [1]. 데이터셋에 비디오 스트리밍 트래픽이 포함되어 있으며 이중 Netflix와 Amazon Prime 트래픽을 사용했으며 이동하는 차량에서 초당 수집한 데이터 36,000개와 정지한 상태에서 초당 수집한 데이터 36,000개를 합쳐 총 72,000개의 데이터를 사용하였다. 데이터의 표준편차가 크면 모델의 학습 속도가 떨어지는 문제가 발생한다. 이를 해결하기 위해서 데이터 값을 일정 범위 내로 변환하는 정규화 과정이 필요하다. 최대값과 최소값을 이용하여 -1에서 1사이 값으로 정규화 시키는 MIN-MAX Normalization을 사용하였으며 식 (1)과 같다.

$$Y = \frac{2(x - \min(x))}{\max(x) - \min(x)} - 1. \quad (1)$$

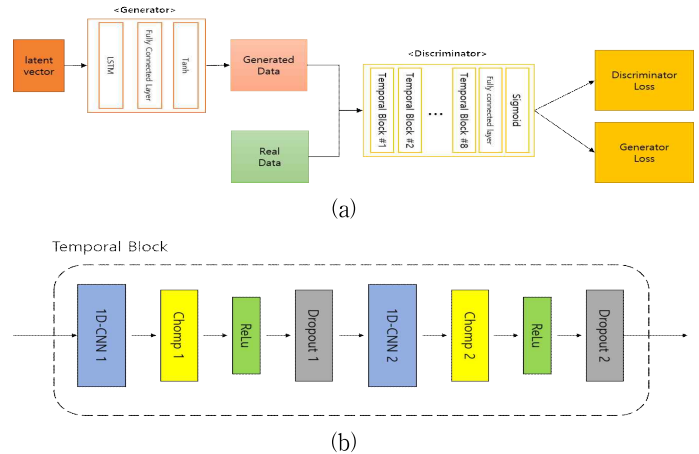


그림 1. (a) 설계한 모델 구조, (b) Temporal Block 내부 구조

나. GAN 구조

설계 모델은 생성자와 판별자 모델을 적대적으로 경쟁시키면서 발전시키는 모델이다. 생성자는 랜덤 정규분포인 latent vector를 입력으로 사용하여 가짜 5G 트래픽을 만들어서 판별자를 속이고 판별자는 생성자가 만들어낸 가짜 5G 트래픽과 실제 5G 트래픽을 구분하면서 학습을 진행한다. 식 (2)는 GAN의 손실 함수이다.

$$\min_G \max_D V(D, G) = E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log (1 - D(G(z)))]. \quad (2)$$

표 1. 각 모델별 JSD, MMD 점수

모델	JSD	MMD
DCCNN-DCCNN (amazon)	0.006753	0.123564
DCCNN-DCCNN (netflix)	0.015946	0.164307
1LSTM-DCCNN (amazon)	0.000499	0.067060
1LSTM-DCCNN (netflix)	0.000163	0.030687
2LSTM-DCCNN (amazon)	0.006547	0.028672
2LSTM-DCCNN (netflix)	0.002708	0.012222
WGAN-GP (amazon)	0.000086	0.169899
WGAN-GP (netflix)	0.000046	0.015154

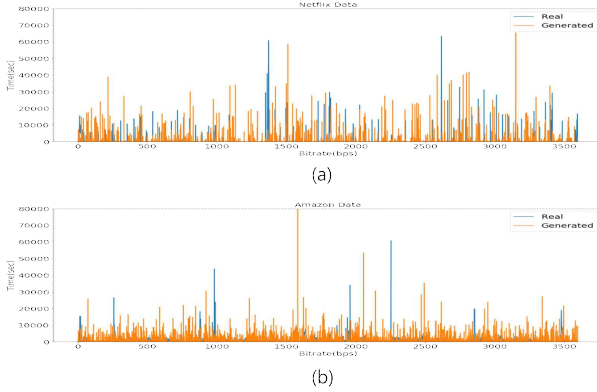


그림 2. (a) Netflix 트래픽 생성 결과, (b) Amazon Prime 트래픽 생성 결과

$E_x \sim p_{data}(x)$ 는 실제 트래픽을 판별자의 입력으로 한 손실이고 $E_z \sim p_z(z)$ 는 생성자가 만들어낸 트래픽을 판별자의 입력으로 한 손실이다. 생성자는 $E_z \sim p_z(z)$ 을 최소화 할 수 있도록 파라미터를 수정하고 판별자는 $E_x \sim p_{data}(x)$ 을 최대화 할 수 있도록 파라미터를 수정한다.

설계한 모델의 생성자는 시계열 데이터 예측에 유리한 LSTM (long short-term memory) [2]을 사용하여 구성했고 판별자는 시계열 데이터 특징 추출에 유리한 DCCNN (dilated causal convolution neural network) [3]을 사용하여 구성했다. 생성자는 256개의 은닉층을 가진 LSTM 층을 통과한 후 선형층과 tanh 활성화 함수를 통과하여 실제 데이터와 마찬가지로 -1에서 1 사이값으로 정규화시킨다. 판별자는 총 8개의 Temporal Block을 사용하였고 판별자의 Temporal Block이 1층 늘어날수록 dilation 이 2의 제곱으로 늘어나며 수용장을 넓힌다. dilation의 범위는 1부터 128 까지이다. Temporal Block을 구성하고 있는 1D-CNN 층은 합성곱 연산을 통해 시계열 데이터의 특징을 추출해 내고 Chomp 층은 합성곱 연산과 정에서 패딩에 의해 늘어난 데이터의 길이를 일정하게 유지시켜주는 역할을 한다. 모든 블록을 통과하고 선형층과 sigmoid 활성화 함수를 통과해서 손실을 구한다.

다. 실험 결과

본 연구에서는 4가지 모델을 사용하여 5G 트래픽 생성 실험을 진행하였다. 모델 성능 평가지표로 JSD와 MMD [4]를 사용했다. JSD는 KL-Divergence [5]를 사용해 거리를 계산할 수 없어 새로 고안된 지표이며 식 (3)과 같다.

$$JSD(P||Q) = \frac{1}{2}(KL(P||M) + KL(Q||M)). \quad (3)$$

여기서 $M = 0.5(P + Q)$ 이다. 분포 P와 Q의 평균값인 M을 사용하여 KL-Divergence와 달리 대칭적인 특징을 가진다. MMD는 두 분포 사이의 차이를 계산하는 방법이다. 본 연구에서는 고정된 time step (3,600초)에서 MMD²를 사용했고 식 (4)와 같다.

$$\widehat{MMD}_u^2 = \frac{1}{(n(n-1))} \sum_{i=1}^n \sum_{j \neq i}^m K(x_i, x_j) - \frac{2}{mn} \sum_{i=1}^n \sum_{j=1}^m K(x_i, x_j) + \frac{1}{m(m-1)} \sum_{i=1}^n \sum_{j \neq i}^m K(y_i, y_j), \quad (4)$$

$$K(x, y) = \exp(-\|x - y\|^2 / (2\sigma^2)). \quad (5)$$

식 (4)의 K 는 가우시안 커널을 의미하고 식 (5)와 같다. LSTM 기반 생성자와 DCCNN 기반 판별자로 설계한 모델은 데이터의 분포, 시계열 특징과 패턴을 효과적으로 학습하였다. LSTM 층을 늘려가며 실험을 진행하였을 때 데이터의 패턴을 더욱 잘 학습하는 모습을 보였지만, 연산에 사용되는 파라미터 개수가 많아져 학습 시간이 증가하고 데이터의 분포를 잘 학습하지 못하는 문제가 발생하였다. 생성자와 판별자를 DCCNN으로 구성한 모델은 데이터 패턴 학습이 이루어진 것에 반해 시계열 특징은 학습하지 못하였다. WGAN-GP [6,7]를 사용한 모델은 데이터의 시간축 특징과 패턴을 잘 학습하지만 분포학습에서 LSTM-DCCNN 모델에 비해 상대적으로 낮은 성능을 보였다. 그림 2는 630 epoch 동안 학습된 LSTM-DCCNN모델의 5G 트래픽 생성 결과이다.

III. 결론

본 논문에서는 LSTM과 DCCNN으로 구성된 GAN을 사용하여 5G 트래픽 생성 모델을 제안했고 실제 트래픽과 생성된 트래픽 비교를 진행하였다. 또한 시계열 데이터 생성에 관련 있는 다른 모델들과 성능 비교를 진행하였다. 실험 결과 분석을 통해 CNN 보다 LSTM 기반 생성기가 시계열 정보의 유지가 뛰어나기 때문에 실제 5G 트래픽과 유사한 트래픽을 생성했고, 평가지표에서도 좋은 성능을 나타내었다. WGAN-GP 모델이 생성한 트래픽보다 LSTM을 이용해 생성한 트래픽의 평가 결과가 실제 5G 트래픽과 더 비슷하다는 것을 확인하였다.

ACKNOWLEDGMENT

이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(No.2021-0-00092, 주파수 확보 및 공급 기반 기술개발).

이 성과는 정부 (과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2021R1F1A1064080).

참 고 문 헌

- [1] Github 5G dataset, <https://github.com/uccmis/5Gdataset>
- [2] Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." Neural computation 9.8 (1997): 1735-1780.
- [3] Wang, Jih-Huang, et al. "Real-time water-level forecasting using dilated causal convolutional neural networks." Water Resources Management 33.11 (2019): 3759-3780.
- [4] Gretton, Arthur, et al. "A kernel two-sample test." The Journal of Machine Learning Research 13.1 (2012): 723-773.
- [5] Fang, Song, and Quanyan Zhu. "Independent Gaussian Distributions Minimize the Kullback-Leibler (KL) Divergence from Independent Gaussian Distributions." arXiv preprint arXiv:2011.02560 (2020).
- [6] Arjovsky, Martin, Soumith Chintala, and Léon Bottou. "Wasserstein generative adversarial networks." International conference on machine learning. PMLR, 2017.
- [7] Gulrajani, Ishaan, et al. "Improved training of wasserstein gans." arXiv preprint arXiv:1704.00028 (2017).