

LSTM 기반 멀티스텝 트래픽 예측 기법 평가

염성웅*, 김형태*, 콜레카르 산자이 시바니*, 김경백^o

Assessing LSTM-based multi-step traffic prediction methods

Sungwoong Yeom*, Hyungtae Kim*, Shivani Sanjay Kolekar* Kyungbaek Kim^o

요약

최근 IoT 기기들의 활성화에 의해 네트워크가 복잡해짐에 따라, 네트워크의 혼잡을 예측하고 미리 대비하기 위해 단기 트래픽 예측을 넘어 장기 트래픽 예측 연구가 활성화되고 있다. 단기 트래픽 예측 결과를 입력으로 재사용하는 재귀 전략은 멀티 스텝 트래픽 예측으로 확장되었지만, 재귀 단계가 진행될수록 오류가 축적되어 예측 성능 저하를 일으킨다. 이 논문에서는 다중 출력 전략을 사용한 LSTM 기반 멀티스텝 트래픽 예측 기법을 소개하고 그 성능을 평가한다. 실제 DNS 요청 트래픽을 기반으로 실험한 결과, 제안된 LSTM기반 다중출력 전략 기법은 재귀 전략 기법에 비해 비정상성 트래픽에 대한 트래픽 예측 성능의 MAPE를 약 6% 줄일 수 있음을 확인하였다.

Key Words : Software-Defined Networking (SDN), LSTM, Multi-step Traffic Prediction

ABSTRACT

Recently, as networks become more complex due to the activation of IoT devices, research on long-term traffic prediction beyond short-term traffic prediction is being activated to predict and prepare for network congestion in advance. The recursive strategy, which reuses short-term traffic prediction results as an input, has been extended to multi-step traffic prediction, but as the steps progress, errors accumulate and cause deterioration in prediction performance. In this paper, an LSTM-based multi-step traffic prediction method using a multi-output strategy is introduced and its performance is evaluated. As a result of experiments based on actual DNS request traffic, it was confirmed that the proposed LSTM-based multiple output strategy technique can reduce MAPE of traffic prediction performance for non-stationary traffic by 6% than the recursive strategy technique.

“본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터육성지원사업의 연구결과로 수행되었음” (IITP-2021-2016-0-00314)

• First Author : Chonnam National University Department of Artificial Intelligence Convergence, yeomsw0421@gmail.com

o Corresponding Author : Chonnam National University Department of Artificial Intelligence Convergence, kyungbaekkim@jnu.ac.kr

* Chonnam National University Department of Artificial Intelligence Convergence, {akkain0513, shivanikolekar}@gmail.com

논문번호 : KNOM2021-02-01, Received October 15, 2021; Revised November 16, 2021; Accepted December 1, 2021

I. 서론

최근 IoT 기기들의 활성화로 인해 인터넷 서비스 제공업체는 인터넷 운영 효율성을 향상시키기 위해 시기적절한 트래픽 흐름을 예측하여 대역폭 할당, 이상 탐지 및 혼잡 제어기법에 대한 연구가 활성화되고 있다 [1][15]. 특히, 인터넷 서비스 관리자가 트래픽 흐름을 정밀하게 파악하고 장단기 트래픽 스케줄링을 수행하기 위해 트래픽에 대한 정확한 예측 기법이 필요해지고 있다 [16]. 네트워크 트래픽은 일정한 주기로 반복되는 계절성, 트래픽 변화가 나타나는 증가 및 감소 추세, 그리고 외부적 요인으로 인해 발생하는 불규칙성(잔차)과 같은 특징을 가질 수 있다. 실생활 트래픽에서는 네트워크 특성에 따른 이러한 계절성, 추세, 잔차의 유무에 따라 다양한 패턴이 나타날 수 있다.

이러한 네트워크 트래픽을 예측하기 위해 자기 회귀 통합 이동 평균(ARIMA) 모델과 같은 통계적 접근 기법이 연구되었다 [2]. 이 기법은 트래픽의 비정상성이 관찰되는 트래픽의 자기상관을 고려하여 트래픽을 예측한다. 하지만, 이러한 통계적 분석 기반 트래픽 예측 기법은 증가하는 IoT 기기들에 의해 불규칙적인 변화가 드러나는 비선형적 트래픽에서는 낮은 성능을 보여준다.

최근 시계열 딥러닝 신경망의 활성화로 인해 시계열 신경망 기반 트래픽 볼륨 예측 기법은 단기 트래픽 예측 측면에서 높은 성능을 보이고 있다 [3]. 하지만, 이러한 단기 트래픽 예측 기법은 네트워크 혼잡을 예측하고 미리 대비하기 위한 충분한 시간을 제공하기 어려울 수 있다. 또한, 단기 트래픽 예측결과를 이용한 잦은 네트워크 설정은 높은 비용으로 인해 네트워크 자원 관리가 비효율적일 수 있다.

제한된 트래픽 자원 관리의 효율성을 높이고 네트워크 혼잡 대비 계획을 위해 단일 단계를 넘어 향후 단계에 대한 트래픽을 볼륨을 예측할 수 있는 멀티 스텝 트래픽 예측 기법이 연구되었다 [4]. 멀티 스텝 시계열 예측을 위한 두 가지 일반적인 전략은 직접적 예측 기법과 재귀적 예측 기법이다. 직접 접근 방식에서는 서로 다른 모델이 병렬로 학습되어 미리 한 단계 앞서 예측한다. 하지만, 각 단계에 대한 예측을 수행하는 다수의 모델을 직접적으로 학습하여 관리하기 때문에 다소 비효율적일 수 있다. 반면에 재귀적 접근 방식은 연속적인 단계의

출력이 다음 예측 단계의 입력으로 사용되며, 반복적인 한 단계 앞선 예측을 수행하여 여러 단계의 시계열 예측을 수행한다. 이 기법은 단일 예측 모델로부터 예측된 결과를 다시 입력으로 사용하여 필요한 시간 범위까지 반복함으로써 트래픽의 동적 특성을 고려한다. 하지만, 반복적 접근 방식은 반복할 때마다 오차가 누적되어, 단계의 수가 증가함에 따라 예측 결과가 실제 관찰에서 점차적으로 벗어나게 된다.

이 논문에서는 장단기 메모리(LSTM) 기반으로 멀티스텝 트래픽을 예측하는 기법을 제안한다. 제안하는 기법은 단계적 예측이 진행됨에 따라 오류가 누적하는 것을 줄이기 위해 인코더 디코더 형태의 다중 출력 모델을 통해 장기적 트래픽 예측을 수행한다. 또한, 제안된 모델이 관찰하는 시간 범위에 따라 변하는 네트워크 특성을 고려하여 장기적인 안목에서 상대적으로 시간이 변해도 통계적 특성이 일정한 정상성 트래픽[17] 및 단기적인 안목에서 상대적으로 시간에 따라 통계적 특성이 변하는 비정상성 트래픽[18]에 대해 고려하여 예측을 수행한다. 이때, 네트워크 유저의 불규칙한 활동으로 인해 트래픽의 정상성이 상대적으로 약해지거나 강해질 수 있다. 우리는 트래픽으로부터 관찰되는 비선형적인 특징을 이동 회귀를 사용한 계절적 추세 분해(STL)를 통해 정의하고 비선형적 특성 정도에 따라 제안된 모델에 미치는 영향을 분석한다. 제안된 기법의 유효성을 검증하기 위해 다양한 네트워크 트래픽 특성이 관찰되는 실제 DNS 요청을 기반으로 실험을 수행하여 평균 절대 백분율 편차(MAPE)를 평가한다.

이 논문은 다음과 같이 구성된다. 2장에서는 기존의 트래픽 예측 연구들에 대해 소개한다. 3장에서는 제안하는 LSTM 기반 멀티스텝 트래픽 예측 기법을 소개한다. 실험 및 해당 평가 결과는 4장에 나와 있다. 5장은 본 논문의 결론 및 향후 연구 내용에 대해 기술 한다.

II. 연구 배경 및 관련 연구

1. 순환 신경망 아키텍처

LSTM은 RNN의 변형 중 하나로서 RNN의 메모리를 확장하여 입력의 장기적인 종속성을 유지하고 학습할 수 있다. LSTM 유닛의 구조는 그림 1과 같

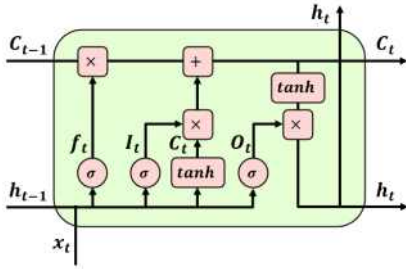


그림 1. LSTM 셀 구조

Fig. 1. The LSTM cell structure

다. LSTM은 기본 RNN 구조에 메모리 정보를 전달하는 셀 상태와 정보 흐름을 조정하는 게이트 메커니즘이 추가된다. LSTM 모델의 게이트에는 메모리에 필요하지 않은 정보를 제거하는 망각 게이트, 메모리에 새로운 정보를 추가여부를 결정하는 입력 게이트, 그리고 출력에 기여하는 메모리를 결정하는 출력 게이트가 있다. LSTM 모델은 이러한 게이트들을 통해서 보존하거나 망각해야 할 가치가 있는 정보를 학습하고 이는 정보의 중요성을 나타내는 가중치에 의해 결정된다. LSTM의 상태 업데이트는 다음과 같다.

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (1)$$

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (2)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (3)$$

$$\tilde{C}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (4)$$

$$C_t = i_t \odot \tilde{C}_t + f_t \odot C_{t-1} \quad (5)$$

$$h_t = o_t \odot \tanh(C_t) \quad (6)$$

여기서 W_i , W_f , W_o , W_c 는 x_t 를 입력 게이트 i_t , 망각 게이트 f_t , 출력 게이트 o_t 및 셀 입력 \tilde{C}_t 에 각각 연결하는 가중치 행렬이다. U_i , U_f , U_o , U_c 는 h_{t-1} 을 3개의 게이트 및 셀 입력에 각각 연결하는 가중치 행렬이다. b_i , b_f , b_o , b_c 는 편향, σ 는 시그모이드 함수, 그리고 \odot 는 Hadamard 곱을 나타낸다. LSTM은 단기 및 장기 기억을 보존하여 모델의 성능을 향상시킬 수 있다. 하지만, 가용 미래 정보를 포착하기 위해 LSTM 모델을 크게 만든다면 새로운 정보를 결합하지 못하고 과거의 정보

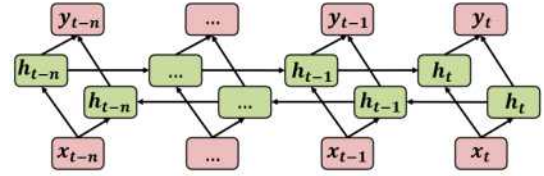


그림 2. Bi-LSTM 모델 아키텍처

Fig. 2. The Bi-LSTM model architecture

를 기억하는데 집중되어 장시간에 대한 예측 성능이 떨어질 수 있다. 장시간 예측에 대한 효율을 높이기 위해 하나의 모델에서 과거 및 미래 정보를 서로 다른 시간 방향으로 학습함으로써 지연을 최소화할 필요가 있다. Bi-LSTM은 LSTM 모델의 확장으로서 두 개의 LSTM을 사용하여 양의 시간 방향(순방향)과 음의 시간 방향(역방향)을 각각 담당하여 학습한 후 그 결과를 병합하는 구조이다 [13]. Bi-LSTM의 구조는 그림2와 같다. 첫 번째 라운드에서는 입력 시퀀스가 순방향 레이어의 LSTM 모델에 적용된다. 두 번째 라운드에서는 입력 시퀀스의 역 형태가 역방향 레이어의 LSTM 모델에 입력된다. LSTM을 두 번 적용하면 학습 장기 종속성이 향상되어 결과적으로 모델의 정확도가 향상될 수 있다 [3].

2. 네트워크 트래픽 예측 기술

최근 네트워크상의 자원 관리 기술이 요구됨에 따라 트래픽 흐름을 정밀하게 파악하고 트래픽 예측을 정확하게 하는 기법이 필요해지고 있다 [14]. 이 트래픽 예측 기법은 네트워크에서 관찰되는 특징에 따라 영향을 받는다. 일반적으로, 네트워크 트래픽은 네트워크 유저의 영향을 받아 오전에는 증가하고 오후에는 변동적이며 저녁에는 감소하는 패턴이 주기적으로 반복되는 계절성이 관찰된다. 이러한 네트워크 트래픽은 장기적인 안목에서 계절적 패턴이 반복됨에 따라 주기의 시작이나 끝을 예측할 수 없어 정상성이 나타나게 된다. 하지만, 단기적인 안목에서 추세에 영향을 받아 비정상성이 나타나게 된다.

이러한 정상성과 비정상성을 고려하여 실시간으로 트래픽을 예측하기 위해 지수 평활 모델 [6], ARIMA 모델 [5] 등을 활용하여 단기 트래픽 예측 기법들이 제안되었다. 이 중 ARIMA 모델은 트래픽 예측 모델을 구축하기 위한 프레임워크로 널리

알려져 있다 [7]. 하지만, 최근 네트워크 유저의 수가 증가함에 따라 네트워크 트래픽에서 무작위 및 비선형적 패턴이 관찰되고 있다. 이러한 무작위 및 비선형적인 패턴이 관찰되는 트래픽에 의해 기존의 통계 기반 트래픽 예측 기법들의 성능이 전반적으로 낮아지고 있다 [8].

최근 시계열 딥러닝 신경망이 활성화됨에 따라 지수 평활 모델과 ARIMA 등과 같은 통계적 기법 대신 다양한 시계열 딥러닝 신경망이 트래픽 볼륨 기법에 적용되고 있다. 다양한 시계열 딥러닝 신경망 중 LSTM은 장기 메모리를 확보하고 있어 오래 전 정보도 기억할 수 있기 때문에 장기 트래픽 예측 측면에서 더 효율적이다. 논문 [9]는 소스 측에서 관찰된 트래픽 계절성을 학습하여 LSTM 기반 네트워크 트래픽 볼륨 추정 방법을 제안했다. 논문 [10]은 네트워크 트래픽 폭주와 불확실성에 대처하기 위해 LSTM 기반의 실시간 네트워크 트래픽 예측 모델을 제안했다. 논문 [11]는 GEANT 백본 네트워크 데이터를 사용하여 LSTM 신경망을 통해 트래픽의 비선형 특성을 학습하고 예측했다. 이러한 단기 트래픽 예측 기법은 실시간성이 요구되는 시스템에 활발하게 사용되고 있다.

제한된 트래픽 자원 관리의 효율성을 높이고 네트워크 혼잡 대비 계획을 위해 단일 단계를 넘어 향후 단계에 대한 트래픽을 볼륨을 예측할 수 있는 멀티 스텝 트래픽 예측 연구가 활성화되고 있다. 논문[19]는 하나의 비선형 자기회귀 외생 모델 (NARX) 신경망의 예측 결과를 다시 입력으로 사용하여 재귀적 반복을 통해 멀티 스텝 예측을 한다. 하지만, 이 기법은 반복할 때마다 오류가 누적되어, 단계의 수가 증가할수록 오차가 점차적으로 커지게 된다. 이 논문에서는 단계가 진행될수록 오차가 누적되는 문제를 해결하기 위해 다중 출력 전략을 채택하여 LSTM 기반 멀티스텝 트래픽 예측 기법을 제안한다.

III. LSTM 기반 멀티 스텝 트래픽 예측 기법

이 논문에서는 네트워크의 혼잡을 대비하고 네트워크 자원의 관리 효율성을 높이기 위해 LSTM 기반 멀티 스텝 트래픽 예측 기법을 제안한다. 제안된 모델의 입출력을 위해서 트래픽을 수집하는 트래픽 샘플러로부터 입력 트래픽 세트 X 와 트래픽 실측값 세트 Y 를 구성하여 관리한다. 예를 들어, 멀티

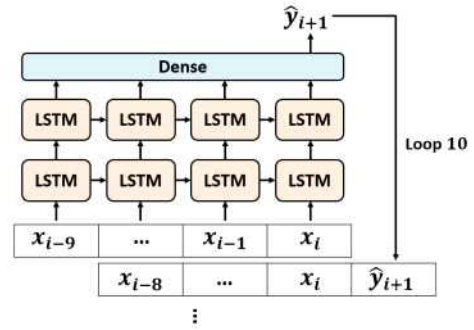


그림 3. 재귀적 전략을 사용한 LSTM 기반 멀티 스텝 트래픽 예측 모델 아키텍처

Fig. 3. The LSTM-based multi-step traffic prediction model architecture using the recursive strategy

스텝 트래픽 예측 모델이 n 개의 입력 트래픽 세트를 이용하여 m 단계 후까지의 트래픽 예측한다면 i 번째 입력 트래픽 세트 X_i 와 트래픽 실측값 세트 Y_i 는 각각 n , m 차원으로 구성하며 수식 (7), (8)와 같이 표현한다.

$$X_i = [x_{i-n+1}, x_{i-n+2}, \dots, x_i] \quad (7)$$

$$Y_i = [y_{i+1}, y_{i+2}, \dots, y_{i+m}] \quad (8)$$

제안된 모델의 예측 결과는 예측 결과 \hat{Y} 를 통해 구성하여 관리한다. i 번째 예측 결과 \hat{Y}_i 는 트래픽 실측값 세트와 같이 m 차원으로 구성하며 수식 (9)와 같이 표현한다.

$$\hat{Y}_i = [\hat{y}_{i+1}, \hat{y}_{i+2}, \dots, \hat{y}_{i+m}] \quad (9)$$

이 예측 결과는 추후에 모델의 오차를 계산하기 위해 사용한다.

1. 멀티스텝 트래픽 예측 전략

제안하는 LSTM기반 멀티 스텝 트래픽 예측 기법은 재귀 전략과 다중 출력 전략을 사용할 수 있다. 그림 3은 재귀 전략을 사용하는 LSTM 기반 멀티 스텝 트래픽 예측 모델의 구조를 보여준다. 재귀적 모델은 단일 단계 트래픽 예측 모델에서 예측된 결과를 다음 예측의 입력으로 사용하는 과정을 반복함으로써 멀티 스텝 트래픽 예측을 수행한다. 재귀 전략의 핵심인 단일 단계 예측 모델은 2개의 LSTM 계층과 1개의 Dense 계층으로 구성된다. 각 LSTM 계층에는 10개의 노드로 구성하고 Dense 계

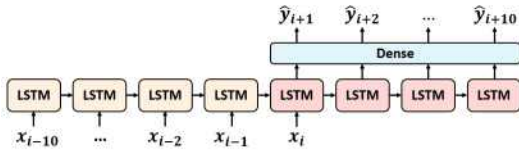


그림 4. 다중 출력 전략을 사용한 LSTM 기반 멀티 스텝 트래픽 예측 모델 아키텍처

Fig. 4. The LSTM-based multi-step traffic prediction model architecture using the multi-output strategy

층의 활성화 함수는 선형 함수를 사용한다. 이 단일 단계 예측 모델은 i 번째 입력 트래픽 세트에서 10개의 과거 트래픽 $x_{i-9+1}, x_{i-n+2}, \dots, x_i$ 을 입력으로 사용하여 $i+1$ 번째 트래픽 \widehat{y}_{i+1} 을 예측하며 이를 10번 반복하여 최종적으로 10 단계에 대한 트래픽 예측 $\widehat{y}_{i+1}, \widehat{y}_{i+2}, \dots, \widehat{y}_{i+10}$ 을 출력한다. 하지만, 모델을 학습할 때 예측된 결과를 입력으로 재사용하는 과정을 반복함에 따라, 이전 단계에서 발생할 수 있는 오류가 다음 단계에서 축적되어 모델의 전반적인 성능이 저하될 수 있다. 예측된 결과를 다음 단계의 입력으로 사용하여 발생하는 오류 축적 문제를 해결하기 위해 다수의 예측 결과를 산출하는 형태의 모델이 필요하다.

그림 4는 인코더 디코더 형태의 다중 출력 전략을 사용하는 LSTM 기반 멀티 스텝 트래픽 예측 모델 구조를 보여준다. 다중 출력 전략은 LSTM을 인코더 디코더 형태로 구성하여 멀티 스텝 트래픽 예측을 수행한다. 제안된 다중 출력 전략 모델은 1개의 LSTM 계층으로 구성된 인코더 및 디코더와 1개의 Dense 레이어로 구성된다. 인코더 및 디코더의 LSTM 계층은 각각 10개의 노드로 구성하고 Dense 레이어의 활성화 함수는 선형 함수를 사용한다. 인코더는 입력 트래픽 $x_{i-9+1}, x_{i-n+2}, \dots, x_i$ 을 순차적으로 입력받은 뒤 모든 정보를 하나의 컨텍스트 벡터로 압축하고 디코더로 전송한다. 디코더는 컨텍스트 벡터를 받아서 10 단계에 대한 트래픽 예측 $\widehat{y}_{i+1}, \widehat{y}_{i+2}, \dots, \widehat{y}_{i+10}$ 을 출력한다. 제안된 모델은 초기에 더 빠르게 수렴하고 오류가 누적하는 것을 방지하기 위해 디코더의 입력으로 트래픽 실측값 세트를 사용하는 교사 강제 기술(Teacher Forcing)을 통해 학습한다.

제안된 멀티스텝 트래픽 예측 모델들은 관찰되는 네트워크 특징에 따라 예측 성능이 달라질 수 있다. 네트워크 트래픽은 계절적 성분에 따라 반복되는

계절성분이 존재하며 이 계절성분에 의해 네트워크 트래픽에서는 장기적인 안목에서 정상성과 단기적인 안목에서 비정상성이 관찰될 수 있다. 이러한 정상성과 비정상성이 관찰되는 경우에 따라 멀티스텝 트래픽 예측 모델은 적절한 입력벡터 표현이 필요하다. 먼저, 정상성 트래픽의 경우, 입력벡터의 첫 번째 도메인은 트래픽 추세가 시간과 상관없이 통계적 특성이 일정하도록 트래픽 변화율 ch_i 으로 설정한다. 입력 벡터의 두 번째 도메인은 트래픽 인덱스 t_i 으로 설정한다. 트래픽 인덱스 t_i 의 간격을 5분으로, 최대 길이를 288으로 설정하여 값이 1부터 288을 반복하게 한다. i 번째 입력 트래픽 세트에서 구성된 입력벡터는 $x_i = (ch_i, t_i)$ 로 표현한다. 비정상성 트래픽의 경우, 시계열 딥러닝 신경망에서 기울기 폭발 또는 손실이 발생하지 않도록 하기 위해 입력벡터의 첫 번째 도메인을 트래픽 정규화 \ddot{s}_i 로 설정한다. 입력 벡터의 두 번째 도메인은 이전과 마찬가지로 트래픽 인덱스 t_i 으로 설정한다. i 번째 입력 트래픽 세트에서 구성된 입력벡터는 $x_i = (\ddot{s}_i, t_i)$ 로 표현한다.

하지만, 네트워크 트래픽에서 관찰되는 비선형적인 추세가 잦아진다면 제안된 모델의 성능이 저하될 수 있다. 이러한 비선형적 트래픽 추세를 면밀하게 학습하기 위해 두 개의 LSTM 모델을 사용하여 순방향과 역방향으로 학습하는 Bi-LSTM을 사용해야 할 필요가 있다.

2. STL 분해 기반 비선형 특징 추출 기법

최근 네트워크 유저의 불규칙한 행동에 의해 트래픽에서 비선형적 특징의 정도에 따라 성능이 달라질 수 있기 때문에 버스트 비율과 지터와 같은 비선형적 특징을 추출하는 기법이 필요하다. 정상성 트래픽의 경우 이러한 비선형적 특징의 정도에 따라 정상성에 더 가까운 강정상성과 정상성에 덜 가까운 약정상성으로 구분할 수 있다. 제안된 모델의 성능과 추출한 비선형적 특징들 간의 관계를 상세히 분석하기 위해 수식으로 정의할 필요가 있다.

이 논문에서는 STL 분해를 기반으로 트래픽 볼륨으로부터 비선형 특징을 추출한다 [12]. STL 분해 기법은 트래픽 볼륨을 계절성 S , 추세 T , 그리고 잔차 R 을 가산분해하고 이를 사용하여 버스트 비율과 트래픽 지터 J 로 표현한다. 버스트 비율 BR 은 일정한 기간 N 동안 i 번째 트래픽에서 버스트

표 1. DNS 요청 트래픽 데이터 셋

Table 1. The DNS request traffic dataset

Dataset	Bucket Size	5 min		10 min		15 min	
	Non-linearity Feature Related City	Avg. Jitter	Burst Ratio	Avg. Jitter	Burst Ratio	Avg. Jitter	Burst Ratio
		LB	Rio de Janeiro	8.24	5.84	7.48	5.53
HB	Uberlandia	8.76	7.12	7.97	7.11	7.6	6.47
LJ	Reston	6.93	5.46	6.74	5.87	6.58	6.29
HJ	Wilmington	14.32	5.62	13.68	5.27	13.21	5.17

트 B_i 가 발생할 확률이며 버스트 B_i 와 버스트 비율 BR 은 수식 (10), (11)과 같이 표현한다.

$$B_i = \begin{cases} 1, & R_i \geq 2 * \sum_{i=1}^N \frac{|R_i|}{N} \\ 0, & R_i < 2 * \sum_{i=1}^N \frac{|R_i|}{N} \end{cases} \quad (10)$$

$$BR = \sum_{i=1}^N \frac{B_i}{N} \times 100 \quad (11)$$

트래픽 지터 J 는 트래픽에서 일정한 패턴이 주기적으로 반복되는 양상이 변위되는 특성을 뜻하며 트래픽 볼륨에서 계절성과 추세를 제외한 잔차의 비율을 나타낸다.

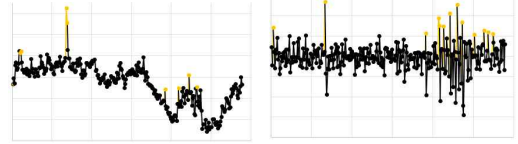
$$J = \sum_{i=1}^N \left| \frac{R_i}{(S_i + T_i) \times N} \right| \times 100 \quad (12)$$

이 버스트 비율 BR 및 지터 J 는 정상성 트래픽과 비정상성 트래픽으로부터 추출되고 정도에 따라 제안된 기법의 성능을 비교할 수 있다.

IV. 결과 및 검증

1. 평가 개요 및 데이터 세트

정상성 트래픽과 비정상성 트래픽에 대해 제안된 멀티스텝 트래픽 예측 기법을 평가하기 위해 재귀적 전략(RE)과 다중 출력 전략(S2S)을 이용하는 멀티스텝 트래픽 예측 기법에 대해 성능을 비교한다. 멀티스텝 트래픽 예측 기법은 사용하는 모델과 전략에 따라 LSTM_RE, Bi-LSTM_RE, LSTM_S2S 그리고 Bi-LSTM_S2S로 표시한다.



(a) 비정상성 트래픽

(b) 정상성 트래픽

그림 5. 비정상성 트래픽과 정상성 트래픽의 예시

Fig. 5. The examples of non-stationary traffic and stationary traffic

보다 사실적인 평가를 위해, ICANN(Internet Corporation for Assigned Names and Numbers)에서 운영하는 DNS-STAT:Hedgehog에서 수집된 레스톤, 윌밍턴, 리우데자네이루, 우베를란디아 등 4개 사이트의 실제 DNS 요청 트래픽으로 각 기법을 평가한다. 각 데이터 세트에는 2020년의 9월 한달 동안의 DNS 요청 트래픽을 나타낸다. 그림 5는 수집된 데이터 세트들의 비정상성 트래픽과 정상성 트래픽의 예시를 보여준다. 그림 5 (a)는 하루 동안 수집된 비정상성 트래픽의 형상을 보여준다. 이 비정상성 트래픽은 단위 시간에 따라 통계적 특성이 변한다. 그림 5 (b)는 수집된 트래픽에 차분을 적용하여 만든 정상성 트래픽의 형상을 보여준다. 이 정상성 트래픽은 단위 시간에 따라 통계적 특성이 변하지 않는 정상성 트래픽을 보여준다. 하지만, 정상성 트래픽과 비정상성 트래픽은 불규칙한 네트워크 유저들의 행동에 의해 높은 지터와 높은 버스트 비율과 같은 비선형적 특성이 관찰될 수 있다.

표 1은 제안된 STL 분해 기반 비선형 트래픽 추출 기법을 통해 서로 다른 지역들에서 수집된 버킷 사이즈별 DNS 요청 트래픽으로부터 추출된 지터 평균 및 버스트 비율을 보여준다. 데이터 세트 중 레스톤 및 윌밍턴은 비슷한 버스트 비율과 지터의 평균을 보여주고 리우데자네이루 및 우베를란디아는

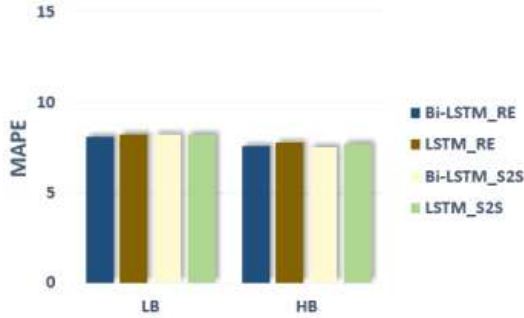


그림 6. 버스트 비율이 다른 정상성 트래픽 환경에서 모델 성능 비교

Fig. 6. The comparison of model performance in stationary traffic environments with different burst rates

비슷한 지터의 평균과 버스트 비율을 보여준다. 레스토랑과 리우데자네이루에서 관찰된 트래픽은 비교적 낮은 버스트 비율과 지터를 보여주기 때문에 선형적 및 계절 방식으로 동작할 가능성이 높고 월밍턴과 우베를란지아 트래픽은 비교적 높은 버스트 비율과 지터를 보여주기 때문에 비선형적, 랜덤 및 버스트 방식으로 동작할 가능성이 높다. 이때, 트래픽 지터 및 버스트 비율은 버킷 사이즈가 커질수록 관찰하는 간격이 늘어나면서 변동적인 트래픽 추세가 다소 줄어들게 되고 전반적인 지터의 평균이 감소할 수 있다. 각 데이터 세트의 처음 20일간의 트래픽은 제안된 모델들을 학습하는데 사용하였으며, 각 데이터 세트의 마지막 10일은 테스트 세트에 사용하였다.

제안된 모델들은 딥러닝 모델의 효율적인 리소스 사용과 관련된 배치 크기를 64로 설정하고 학습 데이터가 신경망을 통과한 횟수와 관련된 에포크(Epoch)를 128로 설정한다. 또한, 제안된 모델들은 학습 프로세스에서 실제 파라미터 갱신과 관련된 최적화 기법으로 Adam Optimizer를 사용한다. 제안된 모델들의 예측에 대한 오차율을 측정하기 위해 가장 보편적으로 사용되는 MAPE를 사용하였다.

2. 성능 비교

제안된 멀티스텝 트래픽 예측 모델들은 정상성 트래픽과 비정상성 트래픽에 따라 다른 영향을 받을 수 있다. 또한, 제안된 모델들은 정상성 트래픽과 비정상성 트래픽에서 관찰되는 지터와 버스트 비율의 정도에 따라 영향을 받을 수 있다. 이때, 정상성 트래픽은 지터와 버스트 비율 정도에 따라 정

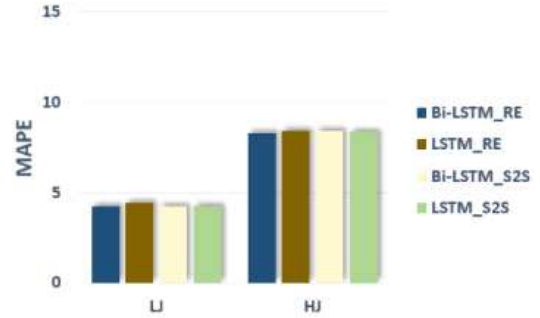


그림 7. 지터 평균이 다른 정상성 트래픽 환경에서 모델 성능 비교

Fig 7. The comparison of model performance in stationary traffic environments with different jitter averages

성능에 더 가까운 강정상성과 정상성에 덜 가까운 약정상성이 나타날 수 있다. 따라서 여러 특징들이 관찰되는 트래픽에서 각 특징이 멀티스텝 트래픽 예측 모델에 미치는 영향을 알아볼 필요가 있다. 이를 위해 버스트 비율이 낮은 레스토랑의 트래픽은 LB, 버스트 비율이 높은 월밍턴의 트래픽은 HB로 표현한다. 또한, 지터가 낮은 리우데자네이루의 트래픽은 LJ, 지터가 높은 우베를란지아의 트래픽은 HJ로 표현한다.

2.1 정상성 환경에서 제안 모델 성능 비교

그림 6은 정상성이 관찰되는 트래픽에서 버스트 비율에 따른 멀티 스텝 트래픽 예측 모델의 성능 비교를 보여준다. 불규칙하고 급격하게 변하는 트래픽 버스트의 비율을 고려하여 비교적 버스트 비율이 낮은 LB 정상성 트래픽을 강정상성 트래픽, 비교적 버스트 비율이 높은 HB 정상성 트래픽을 약정상성 트래픽으로 정의한다. 정상성 트래픽은 시간의 흐름에 따라 통계적 특성이 변하지 않기 때문에 불규칙하게 관찰되는 트래픽 버스트는 시계열 딥러닝 기반 멀티 스텝 트래픽 모델들에게 하나의 잡음으로 인지된다. 따라서 LSTM 기반 멀티 스텝 트래픽 예측 모델들은 LB와 HB 모두에서 약 8%로 유사한 MAPE를 보여준다.

그림 7은 정상성이 관찰되는 트래픽에서 지터 정도에 따른 멀티스텝 트래픽 예측 모델의 성능 비교를 보여준다. 전반적으로 트래픽 변화율이 변동적이고 불규칙적인 트래픽 지터를 고려하여 비교적 지터가 낮은 LJ 정상성 트래픽을 강정상성 트래픽, 비교적 지터가 높은 HJ 정상성 트래픽을 약정상성 트래픽으로 정의한다. 트래픽 추세가 더 변동적인 HJ보다 LJ에서 비교적 더 낮은 MAPE를 보여준다. 이때, 제안된

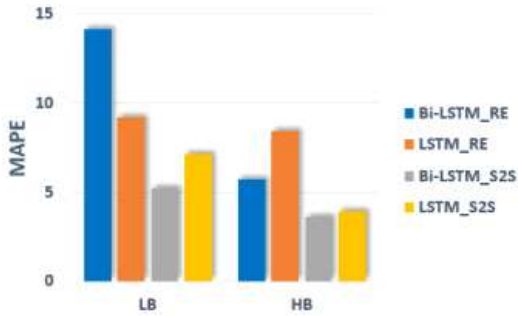


그림 8. 버스트 비율이 다른 비정상성 트래픽 환경에서 모델 성능 비교

Fig. 8. The comparison of model performance in non-stationary traffic environments with different burst rates

모델들은 LJ에서 약 4%로 유사한 MAPE를 보여준다. 그림 6와 7에 대한 평가 결과를 통해 LSTM 기반의 멀티 스텝 트래픽 예측 기법은 채택한 전략과 LSTM 종류에 관계없이 정상성 트래픽에서 비슷한 성능을 보여주고 지터가 낮을수록 높은 성능을 얻을 수 있음을 확인하였다.

그림 6 과 7에 대한 평가 결과를 통해 LSTM 모델들은 정상성이 관찰되는 트래픽에서 특징에 따라 비슷한 MAPE를 보여준다. 특히, LSTM은 버스트 비율보다 지터 정도에 영향을 받는 것을 확인하였다.

2.2 비정상성 환경에서 제안 모델 성능 비교

그림 8은 비정상성이 관찰되는 트래픽에서 버스트 비율에 따른 멀티스텝 트래픽 예측 모델의 성능 비교를 나타낸다. 다중 출력 전략인 LSTM_S2S 및 Bi-LSTM_S2S는 이전 단계의 오류가 섞인 예측 값을 다시 입력으로 사용하는 재귀적 기법인 LSTM_RE 및 Bi-LSTM_RE에 비해 LB와 HB 경우 모두에 대해서 낮은 MAPE를 가진다. LSTM_S2S 및 Bi-LSTM_S2S는 비정상성 트래픽 볼륨의 두드러지는 특징이 되는 버스트에 의해 더 많은 버스트가 존재하는 HB에서 LB에 비해 전반적으로 더 낮은 MAPE를 가진다. 또한, Bi-LSTM_S2S는 기존의 LSTM 셀에 순방향에 역방향을 추가하여 모델의 길이가 길어질수록 과거의 정보가 손실되는 문제를 극복함으로써 LSTM_S2S에 비해 버스트 비율과 관계없이 전반적으로 더 낮은 MAPE를 가진다. Bi-LSTM_S2S는 LB에서 약 5%, HB에서 약 4%의 MAPE를 가진다.

그림 9은 비정상성이 관찰되는 트래픽에서 지터에

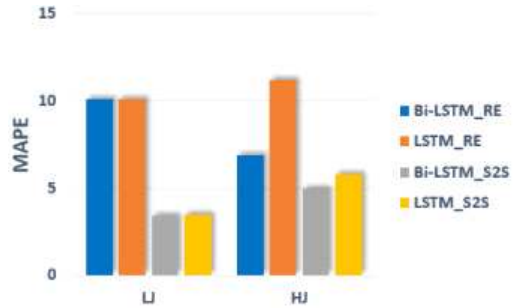


그림 9. 지터 평균이 다른 비정상성 트래픽 환경에서 모델 성능 비교

Fig. 9. The comparison of model performance in non-stationary traffic environments with different jitter averages

따른 멀티스텝 트래픽 예측 모델의 성능 비교를 보여준다. 다중출력 전략인 LSTM_S2S 및 Bi-LSTM_S2S는 재귀적 기법인 LSTM_RE 및 Bi-LSTM_RE에 비해 LJ와 HJ 모두에서 낮은 MAPE를 가진다. LSTM_S2S 및 Bi-LSTM_S2S는 트래픽 추세가 더 변동적인 HJ에 비해 LJ에서 더 낮은 MAPE를 가진다. Bi-LSTM_S2S는 LSTM_S2S에 비해 지터와 관계없이 전반적으로 더 낮은 MAPE를 가진다. Bi-LSTM_S2S는 LJ에서 약 4%, HJ에서 약 5%의 MAPE를 가진다.

그림 8 과 9에 대한 평가 결과를 통해 Bi-LSTM_S2S가 다른 모델들에 비해 비정상성 트래픽에서 버스트 비율과 지터와 관계없이 가장 낮은 MAPE를 가지고, 비정상성 트래픽의 버스트가 높아지고 지터가 낮아질수록 낮은 MAPE를 가지는 것을 확인할 수 있었다.

종합적으로 Bi-LSTM_S2S 기법이 비정상성 트래픽과 정상성 트래픽 모두에서 지터와 버스트 비율과 관계없이 낮은 MAPE를 가질 수 있음을 확인하였다.

2.3 버킷 사이즈에 따른 제안 모델 성능 비교

제안된 Bi-LSTM_S2S 모델은 장기적인 트래픽을 예측하기 위해 스텝의 수를 조절할 수 있다. 하지만, Bi-LSTM_S2S 모델의 입출력 길이를 늘어난다면 기울기 폭발 또는 소실과 같은 장기 의존성 문제가 발생할 수 있다. 이러한 장기 의존성 문제를 해결하기 위해서 입출력 길이 조절이 아닌 트래픽 볼륨의 단위를 조절하는 트래픽의 허용 버킷 사이즈(Bucket Size)를 늘려 장기적인 트래픽을 예측할 수 있다. 기본 버킷 사이즈는 5분으로 설정한다. 하지만, 장기 트래픽 예측을 위해 트래픽 볼륨 단위인

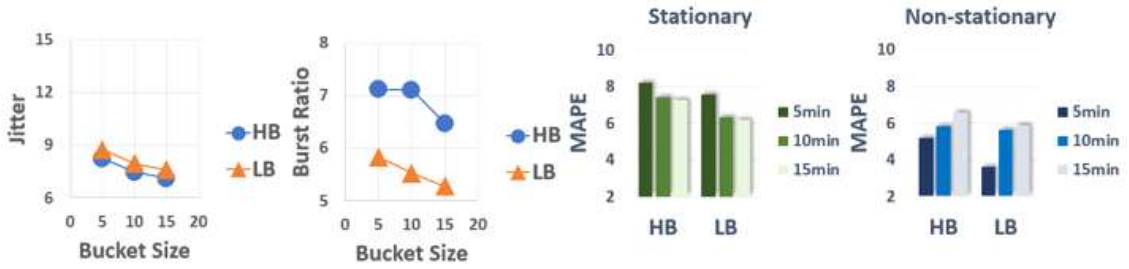


그림 10. 버스트 비율 비교 환경에서 버킷 사이즈에 따른 Bi-LSTM_S2S 성능 평가

Fig. 10. The Bi-LSTM_S2S performance evaluation according to bucket size in burst ratio comparison environment

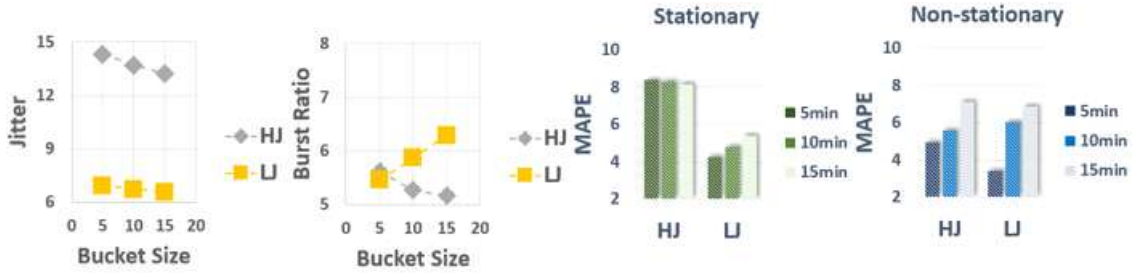


그림 11. 지터 비교 환경에서 버킷 사이즈에 따른 Bi-LSTM_S2S 성능 평가

Fig. 11. The Bi-LSTM_S2S performance evaluation according to bucket size in jitter comparison environment

버킷 사이즈를 변경한다면 해당 트래픽이 가지고 있는 고유 특성과 버스트 비율 및 지터와 같은 비선형 특성이 변하여 성능이 달라질 수 있다. 예를 들어, 트래픽 버킷 사이즈를 증가시킬수록 네트워크 트래픽 볼륨의 변화가 감소하기 때문에 지터가 낮아지게 된다. 버킷 사이즈에 따른 네트워크 특성 변화와 Bi-LSTM_S2S 모델의 성능을 평가하기 위해 버킷 사이즈를 변화시켜가며 예측 모델의 성능을 평가할 필요가 있다. 실험에 사용한 Bucket Size는 5분, 10분, 그리고 15분이다.

그림 10은 버스트 비율 비교 환경에서 버킷 사이즈에 따른 Bi-LSTM_S2S 모델의 성능을 보여준다. 버스트 비율 비교 환경에서 버킷 사이즈가 증가함에 따라 지터와 버스트 비율 모두 감소하게 된다. 이에 따라 정상성 트래픽에서 Bi-LSTM_S2S는 버스트 비율과 지터의 영향을 받아 MAPE가 감소하게 된다. 또한, Bi-LSTM_S2S는 지터가 비교적 낮은 트래픽에서 버스트가 두드러지기 때문에 HB에 비해 버스트 비율이 비교적 낮은 LB에서 전반적으로 낮은 MAPE를 가진다. 즉, Bucket Size 15분 및 LB의 정상성 트래픽에서 Bi-LSTM_S2S는 6%의 MAPE를 보여준다. 반면에 비정상성 트래픽에서 Bi-LSTM_S2S는 버킷 사이즈가 증가함에 따라 상

세한 정보가 사라져 MAPE가 증가하게 된다. 또한, Bi-LSTM_S2S는 비정상성 트래픽에서도 지터가 낮을수록 버스트가 두드러지기 때문에 비교적 낮은 LB에서 전반적으로 낮은 MAPE를 가진다. Bucket Size 15분 및 LB의 정상성 트래픽에서 Bi-LSTM_S2S는 3%의 MAPE를 보여준다.

그림 11은 지터 비교 환경에서 버킷 사이즈에 따른 Bi-LSTM_S2S 모델의 성능을 보여준다. 지터 비교 환경에서 버킷 사이즈를 증가시킨다면 지터는 일관되게 감소한다. 반면에 버스트 비율은 LJ에서 증가하고 HJ에서 감소한다. 이에 따라 HJ의 정상성 트래픽에서 Bi-LSTM_S2S는 버스트 비율과 지터의 영향을 받아 MAPE가 감소하게 되지만 지터가 상대적으로 높아 MAPE 변화는 미묘하다. LJ의 정상성 트래픽에서 Bi-LSTM_S2S는 버스트 비율의 증가율이 지터의 증가율에 비해 비교적 크기 때문에 버킷 사이즈에 비례하여 MAPE가 증가한다. 또한, Bi-LSTM_S2S는 HJ에 비해 지터가 비교적 낮은 LJ에서 전반적으로 낮은 MAPE를 가진다. Bucket Size 5분 및 LJ의 정상성 트래픽에서 Bi-LSTM_S2S는 4%의 MAPE를 보여준다. 반면에 비정상성 트래픽에서 Bi-LSTM_S2S는 버킷 사이즈가 증가함에 따라 MAPE가 지터와 반비례하게 증

가한다. 또한, Bi-LSTM_S2S는 Bucket Size 5분 및 LJ의 정상적 트래픽에서 Bi-LSTM_S2S는 3%의 MAPE를 보여준다.

V. 결 론

이 논문에서는 트래픽 자원 관리의 고 효율성 및 네트워크 혼잡 대비 계획을 위해 LSTM 기반 멀티 스텝 트래픽 예측을 제안한다. 제안된 LSTM 기반 멀티 스텝 트래픽 예측 모델은 재귀적 전략모델과 다중 출력 전략 모델 두 가지로 구분하였다. 또한, 네트워크 유저의 불규칙한 패턴으로 발생하는 버스트 비율과 지터와 같은 비선형적 특성을 트래픽으로부터 STL 분해를 통해 수식으로 정의하였다. 제안하는 기법들을 성능을 비교하기 위해 정상성 트래픽과 비정상성 트래픽으로부터 관측된 지터와 버스트 비율에 따라 평가하였다. 비선형성 정도에 따른 멀티스텝 트래픽 예측 모델들의 성능을 비교한다. 먼저, 다중 출력 전략 모델 Bi-LSTM_S2S는 시간의 흐름에 따라 통계적 특성이 비슷한 정상성 트래픽에서 재귀적 전략 모델과 비슷한 오차율을 달성하고 비정상성 트래픽에서 오류 누적 문제를 해결함으로써 재귀적 전략 모델에 비해 약 6% 낮은 오차율을 달성할 수 있음을 검증하였다. 또한, 제안된 기법들은 버스트 비율과 지터 중에 지터에 더 큰 영향을 받았으며 버스트 비율과 지터가 커질수록 오차율이 높아지는 것을 확인할 수 있었다. 향후, 제안된 기법을 적용하여 실제 네트워크 시스템에서 효율적인 네트워크 자원 관리하는 방법에 대한 연구를 진행하고자 한다.

References

- [1] Bhuyan, Monowar H., Dhruva Kumar Bhattacharyya, and Jugal K. Kalita. "Network anomaly detection: methods, systems and tools." *Ieee communications surveys & tutorials* 16.1 (2013): 303-336.
- [2] T. Alghamdi, K. Elgazzar, M. Bayoumi, T. Sharaf and S. Shah, "Forecasting Traffic Congestion Using ARIMA Modeling," 2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC), Tangier, Morocco, 2019, pp. 1227-1232, doi:10.1109/IWCMC.2019.8766698.
- [3] R Vinayakumar, KP Soman, and Prabaharan Poornachandran. 2017. Applying deep learning approaches for network traffic prediction. In 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI). IEEE, 2353-2358.
- [4] Zang, Chuanyun. "Deep Learning in Multiple Multistep Time Series Prediction." arXiv preprint arXiv:1710.04373 (2017).
- [5] Haviluddin, Haviluddin, and Rayner Alfred. "Forecasting network activities using ARIMA method." (2014).
- [6] Chan, Kit Yan, et al. "Traffic flow forecasting neural networks based on exponential smoothing method." 2011 6th IEEE Conference on Industrial Electronics and Applications. IEEE, 2011.
- [7] Rutka, G. "Network traffic prediction using ARIMA and neural networks models." *Elektronika ir Elektrotechnika* 84.4 (2008): 53-58.
- [8] Somenath Mukherjee, Rajdeep Ray, Rajkumar Samanta, Mofazzal H Khondekar, and Goutam Sanyal. 2017. Nonlinearity and chaos in wireless network traffic. *Chaos, Solitons & Fractals* 96 (2017), 23-29.
- [9] Giang-Truong Nguyen, Van-Quyet Nguyen, Huu-Duy Nguyen, and Kyungbaek Kim. 2018. LSTM based Network Traffic Volume Prediction. In *Proceedings of 2018 KIPS Spring Conference*.
- [10] Giang-Truong Nguyen, Van-Quyet Nguyen, Huu-Duy Nguyen, and Kyungbaek Kim. 2018. LSTM based Network Traffic Volume Prediction. In *Proceedings of 2018 KIPS Spring Conference*.
- [11] H. Lu and F. Yang. 2018. Research on Network Traffic Prediction Based on Long Short-Term Memory Neural Network. In 2018 IEEE 4th International Conference on Computer and Communications (ICCC). 1109-1113.
- [12] Cleveland, Robert B., et al. "STL: A

seasonal-trend decomposition." J. Off. Stat 6.1 (1990): 3-73.

- [13] Siami-Namini, Sima, Neda Tavakoli, and Akbar Siami Namin. "The performance of LSTM and BiLSTM in forecasting time series." 2019 IEEE International Conference on Big Data (Big Data). IEEE, 2019.
- [14] Waleed Akbar, Afaq Muhammad, and Wang-Cheol Song, "Analysis of time-series user request pattern dataset for MEC-based video caching scenario", KNOM Review, Vol. 24, No. 1, Aug. 2021, pp. 20-28.
- [15] Min-A Kim and Seung-Joon Seok, "A Study of Virtual IoT System using Edge Computing", KNOM Review, Vol. 23, No. 1, Aug. 2020, pp. 51-62.
- [16] Doyung Lee, Seyon Jeon, Jonghwan Hyun, Jian Li, and James Won-Ki Hong, "Application-aware Traffic Engineering in SDN", KNOM Review, Vol. 19, No. 2, Dec. 2016, pp. 1-12.
- [17] Ntlangu, Mbulelo Brenwen, and Alireza Baghai-Wadji. "Modelling network traffic using time series analysis: A review." Proceedings of the International Conference on Big Data and Internet of Thing. 2017.
- [18] Cheng, Changqing, et al. "Time series forecasting for nonlinear and non-stationary processes: a review and comparative study." Iie Transactions 47.10 (2015): 1053-1071.
- [19] F. Pilka and M. Oravec, "Multi-step ahead prediction using neural networks," Proceedings ELMAR-2011, 2011, pp. 269-272.

염성웅 (Sungwoong Yeom)



2019년 : 전남대학교 전자컴퓨터공학과 학사 졸업
 2021년 : 전남대학교 인공지능융합학과 석사 졸업
 2021년 ~ 현재 : 전남대학교 전자컴퓨터공학과 박사과정
 <관심 분야> 소프트웨어 정의 인프라스트럭처, 네트워크

이상 감지, 흐름 제어, 딥러닝

김형태 (Hyungtae Kim)



2019 조선대학교 정보통신공학과 졸업
 2021년 ~ 현재 전남대학교 정보보안협동과정 석사과정
 <관심분야> 네트워크 이상 탐지, 딥러닝, 네트워크 보안, 인공지능 보안

클레카르 산자이 시바니 (Shivani Sanjay Kolekar)



2020.03: Shivaji University (Bachelor of Engineering)
 2021.03- present: School of Artificial Intelligence Convergence, Chonnam National University, (M.S.-Ph.D. Integrated Student)

<관심분야> Medical Image Processing, Distributed Network Systems, Big data Platform

김경백 (Kyungbaek Kim)



1999년 : 한국과학기술원 전기 및 전자공학과 학사 졸업
 2001년 : 한국과학기술원 전기 및 전자공학과 석사 졸업
 2007년 : 한국과학기술원 전기 및 전자공학과 박사 졸업

2007년~2011년 : University of California Irvine, 박사 후 연구원

2012년~2015년 : 전남대학교 전자컴퓨터공학부 조교수

2016년~2020년 : 전남대학교 전자컴퓨터공학부 부교수

2021년~현재 : 전남대학교 인공지능융합학과 교수

<관심분야> 분산시스템, 소프트웨어 정의 인프라스트럭처, 빅데이터 플랫폼, 소셜 네트워킹 시스템, 블락체인, AI기반 CPS