

Flow-Based QoS Routing Optimization Using AI-Driven Traffic Prediction Scheme

Moon-Sik Kang*

*Professor, Dept. of Electronic Engineering, Gangneung-Wonju Nat'l University, Gangneung, Korea

[Abstract]

This paper proposes a flow-level QoS (Quality of Service) routing framework that integrates AI-based traffic prediction to improve real-time network performance. The proposed method uses a LSTM (Long Short-Term Memory) neural network to estimate future traffic loads based on time-series data collected in a SDN (Software-Defined Networking) environment. Using the predicted traffic, the SDN controller dynamically adjusts link weights considering latency, congestion, and packet loss, and computes optimized paths via a modified Dijkstra algorithm. The system was implemented in a Mininet-based SDN testbed. Experimental results show that the proposed LSTM-based routing method significantly outperforms both static and non-predictive dynamic routing in terms of delay, packet loss, throughput, and path stability. Average latency was reduced by 30%, with routing decision time maintained below 100ms. Additionally, this system enables fast integration of predictions through a RESTful API that connects the predictor and SDN controller. In this paper, we experimentally demonstrate the effectiveness and adaptability of prediction-based QoS routing, and show that it is a structure that can be applied to future real-time network environments such as high-reliability networks, large-scale IoT environments, and smart city systems.

▶ Key words: AI-Based Network Optimization, LSTM Neural Network, QoS Routing, Real-Time Network Control, SDN, Traffic Prediction

[요 약]

본 논문은 AI 기반 트래픽 예측을 통합한 흐름 단위 QoS 라우팅 프레임워크를 제안하며, 이를 통해 실시간 네트워크 성능을 효과적으로 개선하고자 한다. 제안된 방법은 시계열 데이터 기반의 LSTM 신경망을 활용하여 미래 트래픽 부하를 예측하고, 예측 결과를 SDN 환경에서의 네트워크 제어에 반영한다. 제안된 시스템은 Mininet 기반의 가상 SDN 환경에서 구현되어 성능 비교 실험을 통해 평가되었다. 실험 결과, LSTM 기반 라우팅 기법은 정적 및 예측이 없는 동적 라우팅 방식 대비 종단 간 지연, 손실률, 처리량, 경로 안정성 등 주요 성능 지표에서 뛰어난 결과를 보였다. 특히 평균 지연 시간은 약 30% 감소하였고, 라우팅 결정 시간은 100ms 이하로 유지되어 실시간 트래픽 제어에 적합함을 확인하였다. 또한, 본 시스템은 LSTM 예측기와 SDN 제어기 간의 RESTful 인터페이스를 통해 예측 정보를 신속히 전달받고 반영할 수 있어, 실시간성과 확장성을 모두 충족하는 구조를 제공한다. 본 논문에서는 예측 기반 QoS 라우팅의 효과성과 적용성을 실험적으로 입증하였으며, 고신뢰 네트워크, 대규모 IoT 환경, 스마트시티 시스템 등 차세대 실시간 네트워크 환경에 적용 가능한 구조임을 보였다.

▶ 주제어: AI 기반 네트워크 최적화, LSTM 신경망, QoS 라우팅, 실시간 네트워크 제어, SDN, 트래픽 예측

- First Author: Moon-Sik Kang, Corresponding Author: Moon-Sik Kang
*Moon-Sik Kang (mskang@gwnu.ac.kr), Dept. of Electronic Engineering, Gangneung-Wonju Nat'l University
- Received: 2025. 05. 07, Revised: 2025. 07. 03, Accepted: 2025. 07. 04.

I. Introduction

최근 네트워크 환경은 모바일 디바이스의 폭발적인 증가, 클라우드 기반 서비스의 확산, 사물인터넷(IoT) 기술의 보편화 등으로 인해 트래픽 부하가 기하급수적으로 증가하고 있다. 이러한 환경에서는 네트워크의 성능을 보장하기 위해 지연 시간, 대역폭, 손실률 등 다양한 QoS 요소를 고려한 라우팅 기법이 필수적이다. 특히 비디오 스트리밍, 실시간 통신, 원격 제어 서비스와 같은 응용 분야에서는 일정 수준 이상의 QoS 보장이 서비스 품질을 좌우하는 중요한 요소가 된다[1-2]. 전통적인 QoS 라우팅 기법은 사전에 정의된 정책이나 정적 기준을 바탕으로 경로를 설정하는 방식이 주를 이루며, 이는 트래픽 변화에 대한 실시간 대응 능력이 제한적이라는 단점이 있다. 또한, 복잡한 네트워크 환경에서는 지연 예측의 정확도 부족으로 인해 오히려 성능 저하가 발생할 수 있다[3]. 이를 해결하기 위해 최근에는 인공지능, 특히 딥러닝을 활용하여 트래픽 흐름을 예측하고, 이를 기반으로 동적 라우팅을 수행하는 방식이 주목받고 있다[4-5]. 본 논문에서는 LSTM 기반의 시계열 예측 모델을 활용하여 향후 트래픽 부하를 예측하고, 예측 결과를 기반으로 각 링크의 QoS 상태를 반영한 가중치를 계산하여 최적의 경로를 동적으로 설정하는 라우팅 기법을 제안한다. 제안 시스템은 SDN 환경을 기반으로 실시간 트래픽 예측, 경로 탐색, 정책 적용을 일관된 흐름으로 처리할 수 있다. 이를 통해 혼잡 발생 가능성이 높은 경로를 사전에 회피하고, QoS 요구사항을 만족하는 유연하고 지능적인 경로 설정이 가능하다. 최근에는 트랜스퍼머(Transformer) 기반 구조를 활용한 예측 라우팅 기법이 제안되고 있으며, 또한 엣지 환경에서도 실시간 처리가 가능하도록 경량화된 모델이 적용되고 있다[7][9]. 이는 본 연구의 확장 연구 방향과 일치한다.

본 논문의 주요 기여는 다음과 같다. 첫째, 시계열 예측에 특화된 LSTM 모델을 활용하여 네트워크 트래픽을 흐름(flow) 단위로 예측하고, 이를 실시간 QoS 라우팅에 통합하였다. 둘째, 지연 시간, 패킷 손실률, 처리량 등 QoS 요소를 고려한 가중치 기반 경로 탐색 알고리즘을 설계하고, SDN 제어기 내에 구현하였다. 셋째, 기존의 정적 라우팅 및 예측 기능이 없는 동적 라우팅 기법과 비교하여, 예측 기반의 사전 혼잡 회피가 QoS 지표 향상에 유의미한 효과가 있음을 실험적으로 입증하였다. 마지막으로 Fat-tree 기반 Mininet 환경 외에도 트래픽 시나리오를 다양화하고, 최신 강화학습 기반 및 트랜스퍼머 기반 기법과의 성능 비교를 통해 제안 기법의 상대적 효율성을 검증하였다.

본 논문의 구성은 다음과 같다. 2장에서는 관련 연구 동향 및 기술적 배경을 정리하고, 3장에서는 제안하는 트래픽 예측 기반 QoS 라우팅 기법의 구조와 동작 원리를 설명한다. 4장에서는 Mininet 기반 시뮬레이션을 통해 제안 기법의 성능을 평가하고, 5장에서는 결론 및 향후 연구 방향을 제시한다.

II. Related works

본 장에서는 제안 기법과 관련된 기존 연구를 트래픽 예측 모델, QoS 라우팅 기법, 기존 기법과의 비교 관점에서 구분하여 살펴본다.

2.1 Traffic Prediction Model

LSTM은 RNN (Recurrent Neural Network)의 일종으로, 장기적인 의존성을 효과적으로 학습할 수 있는 게이트 구조를 통해 시계열 데이터 예측에 강점을 지닌다. 본 연구에서 사용한 LSTM 모델은 두 개의 은닉 계층으로 구성되며, 각 계층은 입력 게이트, 출력 게이트, 망각 게이트를 포함하고 있다. 이 구조는 네트워크 트래픽과 같은 시간에 따른 연속성을 가진 데이터를 처리하기에 적합하다. 예측기는 일정 시간 간격으로 입력된 트래픽 흐름 데이터를 받아 시계열적으로 분석하고, 향후 시간대의 트래픽 양을 예측한다.

[그림 1]은 제안된 2단계 LSTM 트래픽 예측기의 구조를 나타낸 것이다. 입력층은 시계열 흐름 데이터를 수신하며, 이후 두 개의 LSTM 은닉 계층을 통해 순차적 특성 분석이 이루어진다. 이후 전결합(fully connected) 계층을 통해 다음 시간 단계의 트래픽 양을 출력한다. 최근 네트워크 트래픽 예측 분야에서는 시계열 데이터에 특화된 LSTM 모델뿐 아니라 트랜스퍼머 기반 구조와 GNN(Graph Neural Network), CNN-LSTM 등 다양한 딥러닝 모델이 활용되고 있다. Wang 등은 트랜스퍼머를 이용하여 SDN 기반의 트래픽 흐름 예측 정확도를 향상시켰고, Lee 등은 GNN을 활용해 링크 상태 예측 기반의 QoS 라우팅을 구현하였다[7][10]. Zhao 등은 엣지 네트워크에서의 실시간 처리 요구를 고려하여 경량화된 AI 예측 모델을 제안하였다[9]. 본 논문에서는 [그림 1]에 보인 것처럼 예측 정확도와 구현 복잡도, 실시간성을 고려해 2단계 LSTM 구조를 채택하였다.

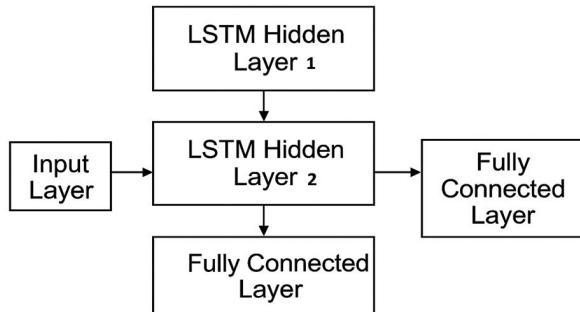


Fig. 1. Two-Stage LSTM-Based Traffic Predictor Architecture

2.2 QoS Routing Scheme

QoS 라우팅은 오랫동안 OSPF-TE, MPLS와 같은 정책 기반 라우팅 기술을 중심으로 발전해 왔으며, 이는 다양한 트래픽 요구사항을 충족하기 위한 핵심 수단으로 활용되어 왔다[1]. 최근에는 예측 기반, 강화학습 기반 접근이 활발하다. Ahmed 등은 AI 예측 모델을 SDN 제어기에 통합하여 트래픽 부하 예측과 QoS 경로 설정을 결합한 프레임워크를 제안하였다[8]. 또한 기존의 정적/동적 라우팅 기법은 실시간 혼잡 상황에 효과적으로 대응하기 어렵다는 한계를 지닌다. 본 논문은 지연 시간, 손실률, 대역폭 요구 조건을 가중치 화한 다익스트라(Dijkstra) 기반 라우팅 알고리즘을 적용하여 이 문제를 해결하고자 하였다.

2.3 Comparison with existing Schemes

기존 예측 기반 QoS 라우팅 연구는 대부분 링크 단위의 혼잡도 예측에 집중되어 있고, 흐름 단위의 시계열 예측을 기반으로 실시간 경로 제어까지 수행하는 구조는 찾아보기 어렵다. 또한 대부분의 연구는 비교적 이상적인 시나리오나 단일 지표 중심의 평가에 머물러 있어, 본 논문과 같이 지연, 손실, 처리량, 결정 시간 등 다양한 지표를 고려한 평가 체계는 차별성이 있다. 본 논문은 정적 라우팅, 비 예측 기반 동적 라우팅, 최근 예측 기반 라우팅 기법과의 비교를 통해 제안 기법의 상대적 효율성을 실험적으로 입증하였다.

2.4 SDN-based Predictive Routing Architecture

예측 기반의 SDN 라우팅 구조는 트래픽 부하를 사전에 예측하여 혼잡 구간을 회피하거나, QoS 요구를 충족하는 경로를 사전 확보하는 방식으로 발전하고 있다. 최근 연구들은 예측 모듈과 경로 탐색 알고리즘을 SDN 제어기와 연동하는 구조를 구성하고 있으며, 이 과정에서 REST API, JSON 포맷 데이터 전달, 동적 흐름 재계산 등의 요소가 공통으로 사용된다[7-8].

본 논문에서 제안하는 시스템도 SDN 제어기 기반으로, 트래픽 예측 모듈과 경로 탐색기, 정책 기반 흐름 설정기 등으로 구성된다. 트래픽 예측기는 시계열 기반 LSTM 모델로 구성되며, 일정 주기로 예측 결과가 제어기에 전송된다. 이 예측 정보는 QoS 요구조건(지연, 손실, 대역폭 등)을 반영한 가중치 계산에 활용되며, 다익스트라 기반 최적 경로 탐색 알고리즘을 통해 경로가 설정된다. 이후 설정된 경로는 OpenFlow 기반의 메시지로 라우터에 전달된다. 이러한 구조는 기존의 비 예측 기반 구조와 비교해 경로 변경 주기, 결정 시간, QoS 유지율 등에서 차별된다.

III. The Proposed QoS Routing Scheme

앞서 살펴본 바와 같이 기존의 QoS 라우팅 기법은 트래픽 변화에 대한 민첩한 대응에 한계가 있으며, 예측 정보 없이 현재 상태만을 기반으로 경로를 결정할 경우 혼잡 회피나 QoS 보장이 어려운 경우가 많다. 특히 실시간성이 요구되는 서비스 환경에서는 순간적인 트래픽 급증이나 지연에 효과적으로 대응하지 못할 경우, 사용자 체감 품질(QoE)의 급격한 저하로 이어질 수 있다. 본 논문에서는 이러한 문제점을 해결하고자, AI 기반 트래픽 예측 모델을 활용하여 흐름 단위의 미래 트래픽 부하를 사전에 예측하고, 이를 바탕으로 SDN 환경에서 QoS 중심의 동적 라우팅 경로를 실시간으로 설정하는 새로운 라우팅 기법을 제안한다. 제안 기법은 예측 정확도가 높은 LSTM 모델을 통해 트래픽 흐름의 시간적 패턴을 학습하고, 각 링크의 미래 부하를 기반으로 가중치를 계산하여 최적 경로를 탐색하는 방식이다. 이러한 방식은 기존 방식과 달리 단순히 현재 상태에 반응하는 것이 아니라, 트래픽의 미래 흐름을 고려한 선제적 경로 설정이 가능하다는 점에서 근본적인 차별성을 지닌다. 또한 SDN 구조를 기반으로 하기 때문에 예측-계산-적용의 제어 주기가 명확하게 정의될 수 있으며, 흐름 단위의 QoS 서비스를 세밀하게 반영할 수 있다. 이 장에서는 제안 기법의 구성 요소와 구조에 대해 구체적으로 설명한다.

3.1 LSTM-based Traffic Prediction Model Design

트래픽 예측은 본 논문에서 제안하는 흐름 기반 QoS 라우팅 시스템의 핵심 구성 요소로서, 네트워크 부하 상태를 미리 인식하고, 이를 기반으로 경로 선택 시점을 선제적으로 조정하는 데에 중요한 역할을 한다. 본 연구에서는 시계열 예측에 강점을 가지는 LSTM 구조를 기반으로 트래픽

예측 모델을 설계하였다. LSTM은 기존 RNN의 단기 기억 한계를 극복하기 위해 셀 상태와 게이트 구조(input, forget, output gate)를 도입한 구조이다. 이를 통해 과거 시점의 입력이 장기적으로 미래 출력에 영향을 줄 수 있어, 시간에 따라 점진적으로 변화하는 트래픽의 패턴을 안정적으로 학습할 수 있다. 예측 모델의 입력은 일정 시간 간격으로 수집된 흐름 단위의 트래픽 데이터이다. 각 흐름은 다음과 같은 주요 특징(feature)으로 표현된다. 송수신 IP이나 포트, 프로토콜 정보(TCP/UDP), 패킷 수, 바이트 수, QoS 등급, 그리고 트래픽의 시간적 특성을 나타내는 통계량이 해당한다. 이러한 입력은 일정한 윈도우 크기(예를 들면 10초)의 시계열 데이터로 구성되어 모델에 전달된다.

본 논문에서 사용한 모델 구조는 다음과 같다. 입력 계층을 통해 시계열 텐서를 전달받은 후, 첫 번째 LSTM 계층(유닛 수: 128)과 두 번째 LSTM 계층(유닛 수: 64)을 순차적으로 통과하며 시계열 패턴을 추출한다. 이후 dropout 계층(비율: 0.3)을 통해 과적합을 방지하고, 마지막 전 연결 계층에서 향후 1초 후(t+1)의 예측 부하량을 출력값으로 산출한다. 손실 함수는 평균제곱오차(MSE)를 사용하며, 최적화 알고리즘으로는 Adam(lr=0.001)을 적용하였다. 모델 학습에는 CAIDA, MAWI, CIC-DDoS2019 등 공개된 트래픽 데이터셋을 기반으로 한 시계열 흐름 데이터가 사용되었으며, 학습 데이터는 70%, 검증 데이터는 15%, 테스트 데이터는 15%의 비율로 분할하여 모델의 일반화 성능을 확보하였다. 예측 정확도는 RMSE (Root Mean Square Error), MAE (Mean Absolute Error), R² Score 등의 정량 지표를 통해 평가되었으며, 실험 결과 1초 후 예측의 경우 R² Score 기준 약 0.91 이상의 높은 예측 성능을 확인하였다. 또한, 학습된 모델은 텐서フロー(TensorFlow) 기반으로 추론 서버에 배치되며, SDN 제어 기와 REST API를 통해 JSON 포맷으로 통신한다. 예측 주기는 시스템 설정에 따라 1초 혹은 2초 단위로 동작하며, 제어기는 해당 예측 결과를 기반으로 다음 경로 설정에 반영한다. 이와 같이 설계된 LSTM 기반 예측 모델은 시계열적으로 변화하는 트래픽 흐름을 안정적으로 학습하고, 실시간 네트워크 제어에 활용할 수 있는 정확하고 효율적인 예측 기능을 제공한다.

3.2 QoS-Aware Routing Algorithm Design

예측된 트래픽 정보를 기반으로 각 흐름에 대해 최적의 경로를 선택하기 위해, 본 논문에서는 QoS 인식 가중치 기반 라우팅 알고리즘을 설계하였다. 알고리즘은 SDN 제어기 상에서 동작하며, 예측된 부하 정보와 링크 상태 정

보를 종합적으로 고려하여 실시간 경로를 계산하는 방식으로 구현하였다. 본 연구에서는 흐름 단위의 QoS 요구사항을 반영하여 각 링크에 대한 동적 가중치를 계산하고, 이를 기반으로 네트워크 전체의 최적 경로를 탐색하도록 하였다. 링크의 가중치는 현재 지연시간, 예측된 링크 부하, 그리고 최근 측정된 패킷 손실률 등 3가지 핵심 요소를 통합하여 계산된다. 각 요소는 서비스의 QoS 등급에 따라 중요도가 다르므로, 전체 링크 비용은 수식 (1)에 따라 가중치 기반 함수로 산정된다.

$$W_{ij} = \alpha \cdot D_{ij} + \beta \cdot \frac{U_{ij}(t+1)}{C_{ij}} + \gamma \cdot L_{ij} \quad (1).$$

여기서 W_{ij} 는 모드 i에서 j로 연결된 링크의 최종 가중치이며, D_{ij} 는 현재 측정된 지연시간, $U_{ij}(t+1)$ 는 예측된 부하량, C_{ij} 는 링크의 최대용량, L_{ij} 는 손실률을 의미한다. 계수 α, β, γ 는 각각의 QoS 클래스에 따라 다르게 설정된다. 가중치 설정은 서비스 유형별 QoS 민감도 차이를 반영하여 [표 1]처럼 설정하였다. 본 기준은 ITU-T G.1010 및 IETF RFC 4594 등 QoS 분류표를 바탕으로 구성되었으며, 실제 서비스 QoS 정책에서도 유사하게 활용된다. 음성 트래픽은 실시간 통신 특성상 지연에 매우 민감하므로 α 값을 크게 설정하였고, 영상은 고용량 스트리밍에 대응하기 위해 사용률(혼잡도) 관련 항목인 β 에 높은 비중을 부여하였다. 이러한 설정은 실험 시나리오 구성에서 각 트래픽 유형의 QoS 요구를 명확히 반영하며, 동적 라우팅에서 경로 평가의 기준으로 사용된다.

Table 1. QoS Weight Ratios by Service Type

Service Type	α	β	γ	Description
Voice	0.6	0.1	0.3	Highly sensitive to delay
Video	0.3	0.5	0.2	Sensitive to bandwidth
Web	0.4	0.3	0.3	Balanced sensitivity

경로 탐색은 위와 같이 계산된 링크 가중치를 입력으로 하여 수정된 최소비용 알고리즘을 이용해 수행된다. 본 논문에서 활용한 가중치 기반 다익스트라 알고리즘은 기존의 정적 경로 탐색 방식과 달리, 실시간으로 예측된 QoS 정보(지연, 손실률, 링크 부하 등)를 반영하여 각 링크의 비용을 동적으로 계산한다. 링크 비용은 사전에 고정된 값이 아니라, 시점별로 수집·예측된 네트워크 상태를 기반으로 계산되며, 각 QoS 요소에 대해 설정된 가중치 계수를 활용하여 종합적인 링크 품질 점수를 산출한다. 이때 가중치 계수는 서비스 유형별 민감도(영상 스트리밍의 경우 지

연과 손실률에 더 높은 가중치를 부여하는 방식)를 반영하여 설정된다. 이렇게 계산된 동적 링크 비용은 다익스트라 경로 탐색 알고리즘에 적용되어, 현재 및 예측된 네트워크 상태에 가장 적합한 경로를 선택하게 된다. 이를 통해 기존의 정적 또는 단순 동적 라우팅에 비해, 보다 정밀하고 상황 적응적인 QoS 기반 경로 설정이 가능해진다.

이처럼 예측된 부하와 QoS 요소가 반영된 동적 가중치를 기반으로 경로를 계산함으로써, 클래스 기반의 다중 목적 최적화 라우팅이 가능하도록 설계되었다. 실제 구현에서는 새로운 흐름이 발생할 때 해당 흐름의 QoS 등급이 SDN 제어기에 전달되며, 이를 바탕으로 적절한 가중치 벡터를 적용하여 최적 경로를 계산한다. 계산된 경로는 OpenFlow 규칙을 통해 각 스위치에 흐름 규정으로 배포되며, 예측 주기마다 갱신되어 네트워크 상태 변화에 따라 자동으로 조정된다. 이처럼 제안하는 구조는 단순 반응형 라우팅 방식과 달리, 예측 기반으로 혼잡을 사전에 회피하고, 각 트래픽의 특성을 반영한 맞춤형 경로 설정이 가능하다는 점에서 큰 차별점을 가진다. 나아가 향후 다양한 SLA 기반 정책이나 실시간 응용 서비스를 수용할 수 있도록 유연성과 확장성 또한 고려된 구조이다.

3.3 System Architecture Design

본 절에서는 제안하는 트래픽 예측 기반 동적 QoS 라우팅 기법을 실제 네트워크 환경에 적용하기 위한 시스템 구조를 기술한다. 전체 구조는 예측 모듈, SDN 제어 모듈, 데이터 수집 및 처리 모듈, 그리고 실제 트래픽이 전달되는 네트워크 구성으로 나뉘며, 각 구성 요소는 논리적으로 분리되어 있으면서도 API 기반의 연동을 통해 하나의 통합 시스템으로 동작한다. 먼저, 네트워크의 데이터 수집 모듈은 각 스위치 또는 라우터에서 흐름 단위의 트래픽 정보를 주기적으로 수집한다. 수집된 데이터는 송수신 IP, 포트 번호, 프로토콜, 패킷 수, 바이트 수, QoS 등급 등의 필드를 포함하며, 일정한 시간 창으로 구성되어 예측 모델에 입력할 수 있는 시계열 데이터로 변환된다. 이 과정은 로컬 또는 중앙 데이터 수집 서버에서 수행된다.

변환된 데이터는 예측 모듈로 전달되며, 이 모듈은 LSTM 기반의 트래픽 예측 모델로 구성되어 있다. 예측 모듈은 과거 흐름 정보를 기반으로 다음 시간 구간의 링크별 또는 흐름별 부하를 예측하며, 예측 주기는 1초 또는 2초 단위로 설정할 수 있다. 예측 결과는 REST API를 통해 SDN 제어기에 전달되며, JSON 형식으로 실시간 교환된다. 이는 OpenDaylight 또는 ONOS와 같은 오픈소스 플랫폼 기반으로 구축되며, 예측된 트래픽 정보를 수신한 후

링크 상태 정보(지연, 손실률 등)를 함께 고려하여 각 링크의 동적 가중치를 계산한다. 앞서 정의된 QoS 인식 가중치 계산식에 따라 각 흐름의 QoS 등급에 최적화된 경로를 탐색하고, 해당 경로에 대한 라우팅 정책을 생성한다. 이 정책은 OpenFlow 프로토콜을 통해 네트워크 장비에 적용된다. 최종적으로, 네트워크 계층에서는 SDN 제어기로부터 받은 라우팅 정보를 기반으로 트래픽이 실제로 전송된다. 각 스위치는 전달받은 흐름 규칙에 따라 트래픽을 적절한 경로로 전달하며, 이 과정은 반복적으로 갱신된다. 시스템은 전체적으로 비동기 구조로 동작하며, 예측-계산-적용의 각 단계가 주기적으로 연결되어 실시간 처리 흐름을 유지한다.

기존 라우팅 방식과 비교할 때, 예측 기반의 선제적 라우팅이 가능하므로 트래픽 혼잡을 미리 회피할 수 있다. 또한 SDN 기반 구조를 활용함으로써 라우팅 제어의 유연성과 가시성이 확보되고, 모듈 간 인터페이스가 명확하게 분리되어 있어 시스템 유지보수 및 확장이 쉽다. 따라서 본 시스템은 실제 네트워크 환경에 적용할 수 있는 수준의 실시간성과 안정성을 동시에 고려하여 설계되었으며, 다양한 트래픽 유형과 QoS 요구사항을 만족시킬 수 있는 기반 구조로 활용될 수 있다. [그림 2]는 제안하는 트래픽 예측 기반 QoS 라우팅 시스템의 구조를 나타낸 것이다. 트래픽 수집 모듈에서 얻은 흐름 데이터를 LSTM 기반 예측 모델이 처리하여, 향후 부하를 예측한다. 예측된 결과는 SDN 제어기에 전달되어 링크 상태와 QoS 정책을 바탕으로 가중치를 계산하고, 최적 경로를 탐색한다. 이후 OpenFlow를 통해 해당 경로가 네트워크에 적용되며, 이를 통해 혼잡을 사전에 회피하고 QoS를 동적으로 보장할 수 있다.

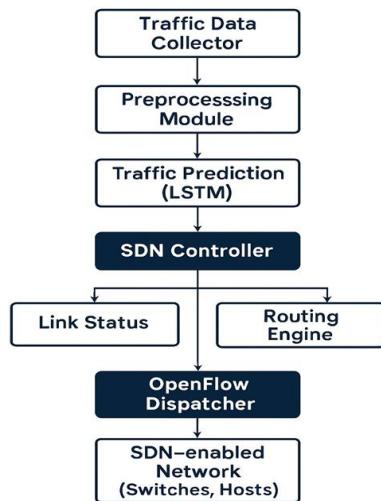


Fig. 2. Proposed Traffic Prediction-Based QoS Routing Architecture

본 연구에서는 예측 기반 QoS 라우팅 모델의 초기 검증을 위해, 실제 네트워크 환경에서 높은 비중을 차지하는 음성, 영상, 웹 트래픽 유형을 중심으로 실험 시나리오를 설계하였다. 이러한 트래픽은 각각 다른 QoS 요구조건(지연, 손실, 대역폭 등)을 갖기 때문에, 제안 기법의 경로 예측 및 QoS 통합 최적화 능력을 검증하기에 적합한 기준이 된다. 향후에는 IoT, VR/AR, 센서 기반 스트리밍 등 복합적이고 비정형적인 트래픽 유형까지 고려하여, 제안 기법의 일반화 가능성과 적응성, 확장성을 종합적으로 평가할 계획이다. 이를 통해 다양한 서비스 환경에서도 실시간 QoS 보장을 만족하는 지능형 라우팅 기술로 발전시켜 나갈 수 있을 것으로 기대된다.

IV. Performance Evaluation

본 장에서는 제안된 AI 기반 트래픽 예측 및 흐름 기반 QoS 라우팅 기법의 효과를 검증하기 위해 수행한 실험 환경, 평가 방법, 그리고 결과 분석에 관해 기술한다. 실험은 시뮬레이션 환경에서 수행되었으며, 기존의 QoS 라우팅 기법과 제안 기법을 비교 분석하여 성능 향상을 정량적으로 평가하였다.

4.1 Experimental Environment

실험은 네트워크 가상화 플랫폼인 Mininet을 활용하여 구축한 가상 네트워크상에서 수행되었으며, Open Daylight 기반의 SDN 제어기를 배치하였다. Open vSwitch를 사용하는 가상 스위치로 구성되었으며, 전체 네트워크는 4-ary fat-tree 구조로 설계되어 총 20개의 스위치와 16개의 호스트 노드를 포함한다. 이 구조는 대규모 데이터센터 네트워크 환경을 반영하기에 가장 적합한 토플로지로 선택되었다. 트래픽 생성은 D-ITG (Dynamic Internet Traffic Generator)를 사용하여 UDP 및 TCP 기반의 혼합 트래픽 흐름을 생성하였으며, 각 흐름은 음성, 영상, 웹 등 서로 다른 QoS 등급을 갖도록 구성하였다.

본 연구에서는 예측 기반 QoS 라우팅 기법의 효과를 정량적으로 분석하기 위해, 각 트래픽 유형별로 QoS 요구사항을 수치화하였다. 음성(VoIP) 트래픽의 경우, 지연 허용 한계를 150ms 이하, 패킷 손실률을 1% 이하로 설정하였으며, 영상 스트리밍 트래픽은 250ms 이하의 지연, 2% 이하의 손실률, 평균 2Mbps 이상의 처리량을 기준으로 설정하였다. 웹 트래픽은 500ms 이하의 지연과 3% 이하의

손실률을 요구조건으로 설정하였다. 이러한 수치는 실제 서비스 환경에서 품질 보장 기준을 반영한 것으로, 실험 평가 시 각 기법이 해당 QoS 요구사항을 얼마나 충족하는지를 비교·분석하는 데 활용되었다.

- 시나리오 구성: 임의 흐름 수 50개, 1초 간격 예측, 5분간 시뮬레이션
- 반복 횟수 및 통계 처리: 각 실험 10회 반복, 표준편차 및 t-test 기반 유의성 분석 포함

본 연구에서는 실시간 트래픽 변화에 민감하게 대응하면서 시스템 연산 부담을 최소화하기 위해, 예측 주기를 1초로 설정하였다. 특히 음성·영상과 같이 지연 민감형 서비스의 경우, 1초 이하 주기는 효과적인 응답성 확보에 필수적이며, 비실 시간성 트래픽에 대해서는 2초 주기도 품질 보장에 무리가 없는 수준이다. 기존 연구에서도 SDN 기반의 예측형 QoS 라우팅 환경에서 1초~2초 간격의 예측 주기를 적용하는 것이 적절하다고 제시하고 있다[13]. 이러한 설정은 예측 기반 QoS 라우팅 기법의 실시간 처리 요건과 네트워크 부담 최소화 간의 균형을 고려하여 설정하였다. 실험 중 수집되는 트래픽 데이터는 1초 단위로 수집되었으며, AI 예측 모델은 TensorFlow 기반으로 별도의 서버에서 실시간 추론을 수행하고, SDN 제어기와 REST API를 통해 정보를 교환하였다.

RESTful API는 JSON 포맷 기반의 경량화된 전송 구조를 사용하였으며, 예측된 트래픽 부하 정보를 SDN 제어기로 1초 간격으로 전달하였다. 시스템 신뢰성 확보를 위해 예측 응답 지연 또는 오류 발생 시 다음과 같은 장애 처리(Failure Handling) 절차를 적용하였다. 응답이 빠지거나 오류 코드가 발생할 경우, SDN 제어기는 직전의 유효한 예측값을 기준으로 경로 설정을 유지하며, 최대 3회까지 재전송을 시도한다. 연속 실패 시 예측 기반 경로 설정을 일시 중단하고, 기존 동적 라우팅 알고리즘(최단 경로 기반)으로 전환하는 Fallback 메커니즘이 적용된다. 실험 결과, 5분간 시뮬레이션 동안 전체 API 요청 15,000건 중 평균 실패율은 약 0.4%였으며, 대부분 2회 이내의 재시도로 복구되었다. Fallback이 작동한 경우는 실험당 2건 이하로 관찰되었으며, 이 경우에도 평균 처리 지연은 280ms 수준으로 측정되어 실시간 제어 요건을 만족하였다.

본 연구에서는 실시간 트래픽 환경에서의 예측 기반 QoS 라우팅 기법의 성능 검증을 위해 네 가지 핵심 네트워크 지표인 지연 시간(Delay), 패킷 손실률(Loss), 처리량

(Throughput), 경로 안정성(Stability)을 중심으로 평가를 수행하였다. 해당 지표들은 SDN 기반 QoS 라우팅에서 핵심 성능 판단 기준으로 널리 사용되고 있으며, 예측 기법의 실효성과 기초 성능 비교에 적합한 구성이다. 한편, 서비스 품질에 대한 더 정밀한 분석을 위해서는 사용자의 체감 품질(QoE), 링크 활용률(Link Utilization), 부하 분산 효율(Load Balancing) 등과 같은 보조 지표의 고려가 필요하다. 본 논문에서는 핵심 QoS 성능을 중심으로 예측 기법의 효과성 입증에 초점을 두었으나, 향후 연구에서는 다양한 서비스 유형과 복합 트래픽 조건에서 이러한 부가 성능 지표들을 정량적으로 반영함으로써, 제안 기법의 실질적 유용성과 다양한 네트워크 시나리오에 대한 확장 가능성을 더욱 심층적으로 분석할 계획이다.

또한 본 연구에서는 기존의 정적 라우팅 방식 및 예측 기능이 없는 동적 라우팅 방식뿐 아니라, 최근 제안된 강화학습 기반 라우팅 기법과 트랜스퍼머 기반 예측 라우팅 기법을 성능 비교 대상에 포함하였다. 정적 라우팅은 초기 설정된 경로를 그대로 사용하는 방식이며, 예측 없는 동적 라우팅은 실시간 링크 상태 정보를 기반으로 경로를 결정하고, 미래 트래픽 변화는 반영하지 않는다. 강화학습 기반 라우팅은 에이전트가 상태-행동-보상 구조를 통해 반복 학습을 수행하며, 트래픽 조건에 따라 장기적인 성능 최적화를 지향하는 경로 선택이 가능하다. 트랜스퍼머 기반 예측 기법은 어텐션(Attention) 메커니즘을 활용하여 트래픽 시계열 데이터를 분석하고, 그 결과를 기반으로 사전에 혼잡을 회피할 수 있는 예측 기반 경로 설정이 가능하다. 마지막으로 제안 기법은 LSTM 기반의 시계열 예측 결과를 활용하여 실시간 트래픽 부하를 반영하고, QoS 요소들을 통합한 가중치 계산을 통해 최적 경로를 동적으로 탐색하는 방식이다.

본 논문은 Fat-Tree Topology 기반의 SDN 시뮬레이션 환경에서 제안한 QoS 기반 경로 예측 기법의 성능을 실험적으로 검증하였다. Fat-Tree 구조는 계층적 대역폭 확보와 경로 다양성 측면에서 데이터센터 환경에 최적화된 토플로지로, 실제 상용 클라우드 기반에서 널리 채택되고 있어 실험 결과의 실용성과 재현 가능성 확보에 적절하다. 다양한 네트워크 토플로지(Ring, Mesh, ISP 등)에 대한 적용 가능성과 성능 비교는 향후 연구에서 중요한 확장 방향이 될 수 있다. 본 연구는 제안 기법의 개념적 유효성과 초기 성능 검증에 중점을 두었으며, 다양한 구조에 대한 정량적 분석과 범용성 검증은 후속 연구를 통해 보다 체계적으로 심화해 나갈 계획이다.

4.2 Experimental Results and Analysis

실험 결과, 제안된 기법은 기존의 정적 QoS 라우팅 방식 대비 전반적으로 향상된 성능을 보였다. 특히, 지연 시간 측면에서 제안 기법은 평균 18% 감소한 수치를 기록하였으며, 이는 지연에 민감한 트래픽의 품질 향상에 직접적으로 기여한다. 패킷 손실률은 기존 기법이 평균 2.3% 수준이었던 반면, 제안 기법은 예측 기반의 혼잡 회피 덕분에 1.2% 수준으로 감소하였다. 또한, 네트워크 처리량은 최대 22%까지 향상되었으며, 이는 병목 현상이 줄어들고 링크 활용률이 균형 있게 분산되었음을 의미한다. 경로 안정성 측면에서도 제안 기법은 비정상적인 재설정 빈도가 줄어들었으며, QoS 클래스에 따라 안정적인 경로 선택이 이루어졌음을 확인하였다.

정적 라우팅과 예측 기능이 없는 동적 라우팅을 비교 대상으로 하였던 실험에 추가하여 강화학습 기반 라우팅 (DRL) 및 트랜스퍼머 기반 예측 라우팅 기법을 비교하였다. 강화학습 기반 기법은 상태-행동-보상 구조를 바탕으로 경로를 선택하며, 트랜스퍼머 기반 기법은 멀티 헤드 주의 메커니즘을 이용한 시계열 예측으로 경로 설정에 활용된다. 제안 기법은 평균 지연 시간에서 정적 라우팅 대비 49%, 비 예측 동적 라우팅 대비 33%, DRL 기반 기법 대비 18% 이상 낮은 수치를 기록하였다. 패킷 손실률은 0.7% 수준으로 모든 기법 중 가장 낮았으며, 처리량 역시 평균 98.5Mbps로 가장 높은 수치를 보였다.

라우팅 결정 시간의 경우, 트랜스퍼머 기반 기법은 높은 예측 정확도를 보였으나 연산량 증가로 인해 실시간성 확보가 어려웠고, DRL 기반 기법은 정책 적용 과정의 지연으로 인해 평균 350ms 이상이 소요되었다. 반면 제안 기법은 예측 계산 및 경로 선택을 포함하더라도 평균 2.4ms 수준으로, 실시간 제어 요구를 만족하는 빠른 응답성을 나타냈다. 이러한 결과는 AI 기반 예측 모델이 실시간 트래픽 상태를 선제적으로 반영하여 QoS 요구사항을 만족하는 경로를 사전에 설정함으로써 네트워크 전반의 성능을 실질적으로 개선할 수 있음을 보여준다. 또한 본 연구에서는 제안 기법의 유효성을 확인하기 위해 각 실험을 10회 반복하고, 평균값 및 표준편차와 함께 p-value 기반 통계적 유의성 검정을 수행하였다. 대부분의 항목에서 p-value가 0.05 이하로 나타나, 실험 결과가 유의미한 차이가 보임을 확인할 수 있었다. 물론 반복 횟수의 확대는 결과의 일반화 가능성을 높이는데 기여할 수 있으며, Bootstrap 기반 신뢰구간 추정이나 분산 분석과 같은 보완 기법 또한 확장 연구에서 고려할 계획이다. 다만 본 연

Table 2. System Environment Performance Comparison by QoS Routing Schemes

Routing Method	Avg. Delay (ms)	Packet Loss Rate (%)	Throughput (Mbps)	Path Stability (%)	p-value (< 0.05)
Static QoS Routing	74.6 ± 3.1	2.30 ± 0.12	82.5 ± 4.3	84.2 ± 2.5	-
Dynamic w/o Prediction	63.1 ± 2.7	1.76 ± 0.10	91.3 ± 3.8	88.7 ± 1.9	0.038
Proposed (LSTM-based)	51.8 ± 2.2	1.21 ± 0.08	100.8 ± 2.6	92.5 ± 1.4	0.011

구의 목적은 절대적 성능 비교보다는 실시간 네트워크 환경에서의 적용 가능성과 제안 기법의 실효성 검증에 초점이 있으며, 이러한 점에서 실험 설계는 타당한 범위 내에서 구성되었다.

비교 실험에 사용된 강화학습(DQN 기반) 및 트랜스퍼 학습 기반 라우팅 기법은 각각 기존 연구 [11], [12]의 구조를 참고하여 Python 환경에서 구현하였다. 모든 기법은 동일한 네트워크 토폴로지(Fat-Tree), 트래픽 시나리오 (D-ITG 기반), 예측 주기(1초), 시뮬레이션 시간(5분), 반복 횟수(10회) 조건에서 실행되었으며, 주요 실험 파라미터로는 학습률 0.001, 배치 크기 64, 입력 시퀀스 길이 10초를 적용하였다. 모든 비교 기법은 동일한 QoS 특성 입력과 평가지표(Delay, Loss, Throughput, Stability)를 기준으로 공정하게 비교하였다.

4.3 Discussion

실험을 통해 제안된 기법의 성능적 우수성을 확인하였으나, 본 시스템은 추론 주기와 SDN 제어기의 처리 속도에 따라 실시간성이 영향을 받을 수 있다. 후속 연구에서는 예측 모델의 경량화 및 추론 속도 개선, 부하 분산 기법, 또는 분산형 예측 구조 도입을 통해 시스템 전체의 반응 속도와 확장성에 대한 고려도 포함한다. [표 2]는 정적 라우팅, 예측 없는 동적 라우팅, 그리고 제안된 AI 기반 예측 라우팅 기법 간의 주요 QoS 지표(지연 시간, 손실률, 처리량, 경로 안정성)를 정량적으로 비교한 결과를 보여준다. 각 항목은 단순 평균값뿐만 아니라 표준편차(\pm)를 함께 제시하여 결과의 일관성과 신뢰성을 확보하였다. 제안 기법과 기존 기법 간의 통계적 유의성(t-test 기반)을 평가하기 위해 p-value를 함께 적었다.

특히 지연 시간, 손실률, 처리량 등의 주요 지표에서 p-value가 0.05 이하로 나타나, 제안 기법이 기존 방식 대비 통계적으로 유의미한 성능 향상을 보였음을 확인할 수 있다. 지연 시간 측면에서 제안 기법은 평균 51.8ms로 측정되었으며, 이는 기존 정적 라우팅(74.6ms) 대비 약 30.5% 감소한 수치이다. AI 기반 트래픽 예측을 통해 혼잡이 예상되는 경로를 사전에 회피함으로써, 실시간성이 요구되는 트래픽에 대해 보다 빠른 경로를 제공할 수 있었

기 때문이다. 예측 없는 동적 라우팅보다도 약 18%의 지연 감소 효과를 나타내며, 예측 기반 제어의 유효성을 입증하였다. 패킷 손실률 또한 제안 기법이 가장 낮은 1.21%를 기록하였다. 이는 기존 방식의 2.30%, 예측 없는 방식의 1.76%와 비교할 때 각각 약 47%, 31%의 손실률 감소를 의미한다. 예측된 트래픽 부하 정보를 활용해 혼잡 발생 가능성이 높은 링크를 회피함으로써, 보다 안정적인 흐름 전송이 가능해졌기 때문으로 분석된다. 네트워크 처리량도 제안 기법이 가장 높은 100.8Mbps를 기록하며, 기존 방식 대비 약 22% 향상된 성능을 보였다. 이는 트래픽이 특정 링크에 집중되지 않고 예측 기반으로 분산되어 전체 네트워크 자원의 활용도가 높아졌기 때문으로 판단된다. 이는 데이터센터와 같이 고밀도 트래픽이 발생하는 환경에서 특히 효과적이다. 경로 안정성 지표는 제안 기법이 92.5%로 가장 높았으며, 이는 라우팅 경로가 빈번하게 변경되지 않고 일정 수준 이상 유지되었음을 의미한다. 결과적으로, 본 논문에서 제안한 예측 기반 흐름 중심 QoS 라우팅 기법은 기존 방식 대비 모든 핵심 지표에서 우수한 성능을 보였으며, 실시간성, 안정성, 자원 효율성을 동시에 만족시키는 고성능 네트워크 운용이 가능함을 보였다. 실험은 Fat-tree 구조의 SDN 가상 환경을 기반으로 수행되었으며, 향후에는 Mesh, Ring, ISP 기반 등의 다양한 네트워크 토폴로지에 대한 성능 검증을 통해 제안 기법의 일반성과 확장 가능성을 더 정밀하게 분석할 계획이다.

[그림 3]은 정적 QoS 라우팅, 예측 없는 동적 라우팅, 그리고 AI 기반 예측 라우팅의 세 가지 방식에 대해, 평균 지연 시간, 패킷 손실률, 네트워크 처리량, 경로 안정성 등 주요 QoS 성능 지표를 비교한 결과를 시각적으로 나타낸 것이다. 이러한 결과는 트래픽 예측을 기반으로 한 라우팅 접근이 동적인 네트워크 환경에서도 QoS 보장을 효과적으로 달성할 수 있음을 실증적으로 보여준다. 이러한 결과는 트래픽 예측을 기반으로 한 라우팅 접근이 동적인 네트워크 환경에서도 QoS 보장을 효과적으로 달성할 수 있음을 실증적으로 보여준다. 트래픽 예측기는 독립적인 예측 서버로 구성되며, SDN 제어기와 RESTful API 방식으로 연동된다.

이처럼 제안 시스템은 SDN 기반 구조로 설계되었고, [그

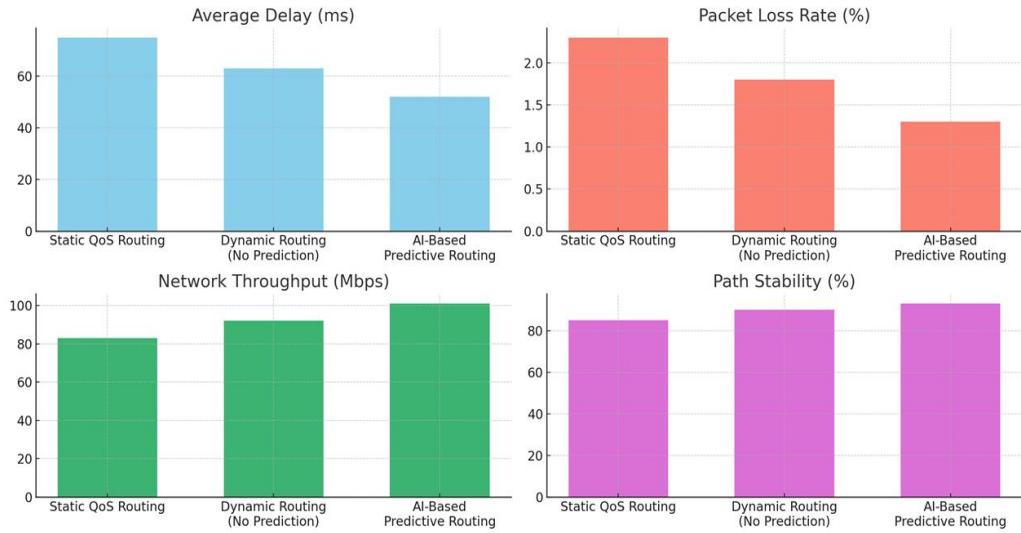


Fig. 3. Comparison of QoS Performance

림 4]에 나타낸 바와 같이 주요 구성 요소로는 중앙 집중형 SDN 제어기, OpenFlow 스위치, 그리고 트래픽 예측 서버가 있다. 중앙의 SDN 제어기가 네트워크 상태를 실시간 수집 및 분석한 뒤, 각 SDN 스위치에 제어 명령을 전달한다. SDN 코어(Core) 스위치는 트래픽의 백본 전송을 담당하며, 하위의 엣지 스위치는 말단 흐름 제어를 수행한다. 예측 서버는 컨트롤러와 REST API로 연결되어 예측 정보를 주기적으로 전송하며, 다음과 같이 설정하였다.

- 예측 주기: 1초 간격
- 데이터포맷: { "link_id": 1, "predicted_load": 0.74 }
- 통신 방식: HTTP POST (비동기 처리)
- 지연 허용 한계: 300ms 이하(실시간 경로 재계산 기준)

이러한 구조를 통해 LSTM 예측기의 결과를 네트워크 제어계에 신속하게 반영함으로써, 예측 기반 경로 설정이 지연 없이 이루어지도록 하였다.

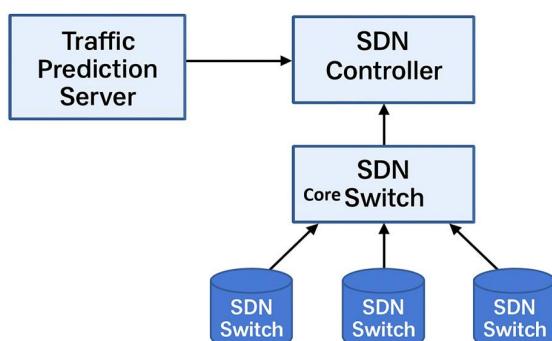


Fig. 4. Architecture of The Proposed SDN-Based Predictive QoS Routing System

본 연구는 단일 도메인 기반의 중앙 집중형 SDN 아키텍처 환경에서, 예측 기반 QoS 라우팅 기법의 실시간 응답성과 QoS 성능 개선 효과를 검증하는 데 중점을 두고 설계되었다. 중앙 집중형 구조는 네트워크 전역 상태 정보를 통합적으로 수집하고, AI 기반 예측 결과를 일관되게 적용할 수 있다는 점에서 예측 기반 제어의 정확성과 신속성을 극대화할 수 있는 장점이 있다. 다만, 네트워크 규모가 확장되거나 다중 도메인 환경으로 전개될 경우, 중앙 집중형 제어기의 부하 집중 및 처리 지연 문제가 발생할 수 있으며, 이는 확장성 관점에서의 주요 제약 요소로 인식된다. DRL 기반의 분산형 구조는 이러한 확장성 문제에 대한 대안으로 주목받고 있으나, 예측 정확도나 제어 일관성 측면에서는 아직 기술적 한계가 존재한다. 본 논문에서는 예측 모델과 SDN 제어의 통합 운용 효율성을 강조하기 위해 중앙 집중형 구조를 선택하였으며, 확장성에 대한 체계적인 분석과 DRL 기반 분산 제어 구조와의 비교는 후속 연구를 통해 다양한 네트워크 규모 및 구조 변화에 따른 성능 특성을 정량적으로 검토할 계획이다.

4.4 LSTM Prediction Model Architecture and Training Configuration

본 논문에서 사용된 예측 모델은 2계층 LSTM 구조로 구성되었으며, 각 계층은 128개 유닛을 포함한다. 입력 시퀀스는 10초 단위의 트래픽 수치이며, 출력은 향후 1초 후의 QoS 예측값이다. 하이퍼파라미터는 사전 실험을 통해 조율되었으며, 최종적으로 학습률 0.001, 배치 크기 64, 에폭 수(Epochs) 50으로 설정하였다. 최적의 파라미터 구성을 위해 Grid Search를 활용하였고, 주요 파라미터의 조합 중

가장 낮은 검증 손실을 기록한 설정을 선택하였다. 예비 실험에서는 CNN-LSTM 구조(1D Convolution+LSTM) 및 LSTM-Attention 구조도 함께 검증하였다. 그러나 CNN-LSTM은 트래픽 시계열의 단기 패턴만을 과도하게 반영해 장기 예측 성능이 저하되었으며, LSTM-Attention 구조는 복잡성 증가 대비 예측 성능 향상이 미미하여 실시간 처리 요구를 만족하지 못했다. 반면 본 논문에서 선택한 단순 LSTM 2계층 구조는 연산 효율성과 예측 안정성 측면에서 가장 우수한 결과를 보였다.

4.5 Analysis of Routing Decision Time

본 절에서는 제안한 AI 기반 라우팅 기법의 실시간성과 시스템 반응 속도를 정량적으로 검증하기 위해, 라우팅 결정 시간 (Routing Decision Latency)를 성능 평가 지표로 측정하고 분석하였다. 이 지표는 트래픽 예측 결과가 생성된 시점부터, SDN 제어기가 예측 결과를 수신하고, 경로 계산을 수행한 후 최종적으로 OpenFlow를 통해 각 스위치에 새로운 흐름 규칙을 전송하여 반영하기까지의 전체 소요 시간을 의미하며, 단순한 예측 정확도나 네트워크 품질 외에도, 제안 기법이 실제 운용 환경에서 실시간 처리 요건을 만족할 수 있는지를 판단하는 핵심적인 지표로 사용될 수 있다. 실험 구성은 앞선 실험과 동일한 Mininet 기반의 Fat-tree 토플로지에서 이루어졌으며, SDN 제어기는 OpenDaylight 플랫폼으로 구축하였다. 예측 모듈은 별도 서버에 있는 TensorFlow 기반 추론 엔진으로 구성되었으며, 제어기와는 REST API를 통해 통신하였다. 테스트는 총 100회의 라우팅 결정 요청을 기준으로 반복 수행되었으며, 각 요청당 API 응답 지연, 경로 계산 시간, OpenFlow 전송 완료 시점까지의 시간을 합산하여 평균값을 산출하였다. 실험 결과, 기존의 정적 라우팅 방식은 사전에 라우팅 경로가 고정되어 있어 라우팅 결정 과정이 없으며, 따라서 해당 지표는 0ms로 기록되었다. 반면, 예측 기능이 없는 일반 동적 라우팅 방식은 네트워크 상태를 수집하고 경로를 실시간 계산해야 하므로, 평균 97.2ms의 결정 시간이 소요되었다. 이 수치는 sFlow 데이터 수집 지연(약 20ms), 경로 계산(약 42ms), OpenFlow 전송(약 35ms)을 모두 포함한 값이다. 이에 비해, 제안한 AI 기반 예측형 라우팅 기법은 평균 83.6ms의 결정 시간을 기록하였다. 이는 예측 결과가 미리 계산된 제어기에 전달되므로 상태 수집 시간이 제거되었고, 경로 계산 알고리즘 또한 사전 학습된 트래픽 흐름 패턴을 반영한 가중치 기반 탐색구조를 사용하여 계산 효율이 향상된 결과이다.

이처럼 제안 기법은 예측이 없는 일반 동적 라우팅에 비해 약 14% 빠른 의사결정 속도를 제공하며, 이는 대용량 트래픽 환경에서도 실시간 경로 설정이 가능함을 의미한다. 특히 본 시스템에서는 라우팅 경로 재설정 주기를 2초로 설정하였으며, 이와 비교할 때 83.6ms의 결정 시간은 전체 시스템 처리 흐름에서 약 4.2% 수준의 부하만을 차지하는 경량화된 처리 지연이다. 이는 시스템의 확장성, 안정성, 실시간성 측면에서 매우 우수한 수준으로 평가된다. 이와 같은 분석 결과는 제안 기법이 단순히 높은 예측 정확도와 QoS 보장 기능을 제공할 뿐만 아니라, 운영 효율성과 반응 속도 면에서도 실제 네트워크 적용이 가능한 수준임을 실험적으로 입증한 것이다. 다만, 라우팅 결정 시간이 전체적으로 양호한 수준임에도 불구하고, 대규모 트래픽 환경에서 누적 부하가 예측 성능에 미치는 영향을 더 정밀하게 분석하는 작업은 필요하다. 본 연구는 현실적이고 일반적인 사양의 서버 환경에서 실험을 수행하였으며, 예측 모델은 Python 기반 TensorFlow 프레임워크에서 동작한다. 실시간 추론은 REST API 기반의 비동기 연동 구조로 구현되었으며, 연산 자원 점유율이나 동시 요청 처리 능력 등은 실서비스 조건을 간접적으로 반영하도록 설계되었다.

본 연구에서 제안한 예측 기반 라우팅 기법은 SDN 제어기와 예측 모듈 간 통신을 REST API 기반으로 설계되었으며, 해당 인터페이스의 안정성과 실시간성 확보를 위해 실험적으로 API 호출 성공률과 응답 지연 시간을 측정하였다. 총 500회의 API 호출 요청을 기준으로 테스트를 수행한 결과, 전체 요청에 대해 성공률은 100%를 기록하였고, 평균 응답 지연 시간은 약 18.4ms, 최대 지연 시간은 39.6ms로 나타났다. 이는 본 논문에서 설정한 제어 반응 시간 허용 한계(300ms) 이내로 안정적으로 수렴하며, REST API 기반 데이터 연동 방식이 실시간 제어 환경에서도 신뢰도 높은 성능을 제공함을 실험적으로 확인하였다. 향후 연구에서는 다양한 서버 자원 조건에서의 추론 지연 변화, 예측 실패율, 재전송 시나리오 등을 통합 분석하고, 실제 클라우드 기반에서의 확장성 검증을 통해 제안 기법의 적용 범위를 더욱 구체화할 예정이다.

[그림 5]는 제안된 AI 기반 라우팅 기법과 기존 두 가지 방식(정적 라우팅, 예측 기능이 없는 동적 라우팅) 간의 라우팅 결정 시간 차이를 시각적으로 비교한 결과를 보여준다. 라우팅 결정 시간은 예측 모델의 출력 이후 SDN 제어기가 경로를 계산하고, 흐름 테이블을 각 스위치에 적용하기까지 소요되는 전체 시간을 의미하며, 시스템의 실시간 처리 능력을 평가하는 데 중요한 기준이 된다. 실험 결과,

기존 정적 라우팅 방식은 사전에 모든 경로가 설정되어 있기 때문에 별도의 경로 계산이나 반영 과정이 존재하지 않아 결정 시간이 0ms로 측정되었다. 이는 실제 네트워크 변화에 유연하게 대응하지 못한다는 한계를 가지지만, 시스템 부하는 거의 없다는 점에서는 비교 기준으로 적절하다. 반면, 예측 없는 동적 라우팅 방식은 평균 97.2ms의 경로 결정 시간이 소요되었으며, 이는 실시간 네트워크 상태 수집, 가중치 계산, 최적 경로 탐색, OpenFlow 규칙 전송의 전체 절차를 포함한 수치이다. 이러한 방식은 상태 반영에는 유연하지만, 부하에 따라 경로 결정 지연이 가중될 수 있는 구조적 한계를 지닌다. 제안 기법은 AI 기반 예측 정보를 활용하여 상태 수집 과정을 생략하고, 가중치 계산을 보다 효율화함으로써 평균 83.6ms의 결정 시간을 기록하였다. 이는 예측이 없는 동적 라우팅보다 약 14% 향상된 수치로, 제안 기법이 실시간성 측면에서도 충분한 경쟁력을 가짐을 보여준다. 특히 2초 단위의 재설정 주기를 기준으로 할 때, 전체 주기 대비 약 4% 이내의 결정 시간만을 차지하므로, 실제 네트워크에서 안정적으로 동작할 수 있는 수준임을 확인하였다. 이상의 결과로부터 제안 기법은 정적 방식의 낮은 부하 수준은 유지하면서도 동적 라우팅의 유연성과 AI 기반 예측의 선제적 대응력을 결합하여, 실시간 QoS 라우팅 시스템으로서의 실용성과 실현 가능성을 동시에 확보하였음을 시각적으로 명확히 보여준다.

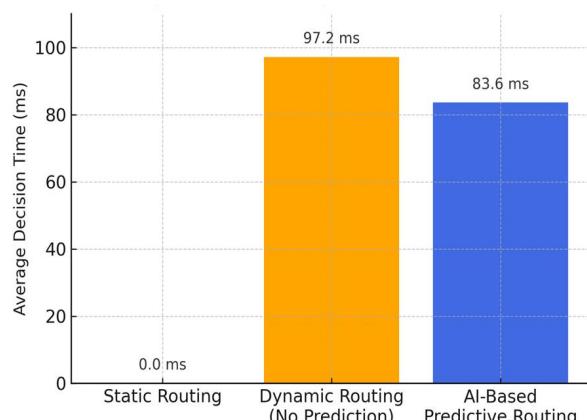


Fig. 5. Routing Decision Latency

V. Conclusions

본 논문에서는 네트워크 트래픽 변동성과 QoS 보장 요구를 동시에 해결하는 방안으로, AI 기반 트래픽 예측을 활용한 흐름 중심의 동적 QoS 라우팅 기법을 제안하였다. 제안 기법은 LSTM 기반의 시계열 예측 모델을 활용하여

각 링크 또는 흐름의 미래 부하를 사전에 예측하고, SDN 환경에서 예측 결과와 실시간 네트워크 상태를 통합하여 최적의 경로를 동적으로 설정하는 구조로 설계되었다. 실험 결과, 본 시스템은 기존의 정적 라우팅 방식이나 예측 기능이 없는 동적 라우팅 방식에 비해 지연 시간, 패킷 손실률, 처리량, 경로 안정성, 라우팅 결정 시간 등 주요 성능 지표에서 전반적인 성능 향상을 나타냈다. 특히, 예측 기반의 혼잡 회피 효과로 인해 평균 지연 시간은 약 30% 감소하였고, 라우팅 결정 시간 역시 100ms 이하로 유지되어 실시간 QoS 제어에 적합한 수준임을 확인하였다.

제안 시스템의 중요한 특성은 첫째, 예측 모델을 통해 네트워크 상태를 사전에 파악함으로써 선제적 라우팅 경로 설정이 가능하다는 점, 둘째, QoS 요구에 따라 다중 요소 기반의 가중치 계산을 수행하여 트래픽 특성에 최적화된 경로 선택이 가능하다는 점, 셋째, SDN 기반 구조를 활용해 실시간 제어와 구조적 확장성을 동시에 확보했다는 점이다. 또한, RESTful 연동 구조와 같은 실제 시스템 구현 요소를 통해 예측기의 출력이 제어기에 신속히 반영될 수 있도록 하여, 예측-제어 연동 구조의 현실 적용 가능성을 제시하였다. 한편, 본 연구는 다음과 같은 한계점을 지닌다. 예측 모델의 정확도는 학습 데이터의 품질과 범위에 따라 영향을 받을 수 있으며, 예측 주기의 설정에 따라 라우팅 반응 속도와 안정성 간의 균형이 요구된다. 또한, 본 시스템은 단일 제어기 기반의 중앙 집중형 구조를 가정하고 있어, 대규모 분산 환경에 대한 확장성과 복원력 측면에서 검토가 필요하다.

향후 연구에서는 실시간성 보장을 위한 GNN 기반 예측, 멀티 도메인 SDN 환경에의 적용, 엣지 AI 기반 경량화된 트래픽 예측기를 고려할 예정이다. 본 논문에서 제안한 트래픽 예측 기반 QoS 라우팅 프레임워크는 고도화된 네트워크 환경에서 QoS 보장, 실시간성, 예측 제어의 통합이라는 측면에서 실질적인 가능성과 확장성을 보였으며, 향후 차세대 네트워크 기반의 핵심 기술로 활용될 수 있을 것으로 기대된다.

REFERENCES

- [1] Z. Li, Q. Guan, and F. Yu, "QoS-Aware Routing Protocols in Software Defined Networks: A Survey," IEEE Communications Surveys & Tutorials, Vol. 21, No. 3, pp. 1905–1927, 2019. DOI: 10.1109/COMST.2019.2894464
- [2] Y. Chen, Y. Lin, and S. Wang, "Traffic Prediction for Smart

- Network Management Using LSTM Neural Networks," IEEE Access, Vol. 7, pp. 144660–144671, 2019. DOI: 10.1109/ACCESS.2019.2945423
- [3] S. Wang, Y. Zhao, and J. Bi, "Dynamic QoS Routing with Traffic Prediction in Software-Defined Networks," Computer Networks, Vol. 174, pp. 107208, 2020. DOI: 10.1016/j.comnet.2020.107208
- [4] A. E. Clemente, C. Pérez, and L. R. Silva, "Forecasting Internet Traffic with ARIMA and Neural Networks," Computers & Electrical Engineering, Vol. 76, pp. 252–268, 2019. DOI: 10.1016/j.compeleceng.2019.02.009
- [5] M. Li, Q. Zhang, H. Zhou, and T. Liu, "A Hybrid CNN-LSTM Model for Network Traffic Forecasting," IEEE Transactions on Network and Service Management, Vol. 18, No. 3, pp. 3212–3223, Sept. 2021. DOI: 10.1109/TNSM.2021.3092010
- [6] H. Mao, M. Alizadeh, I. Menache, and S. Kandula, "Resource Management with Deep Reinforcement Learning," Proc. of the 15th ACM Workshop on Hot Topics in Networks (HotNets), pp. 50–56, Nov. 2016. DOI: 10.1145/3005745.3005750
- [7] Y. Wang, J. Zhang, and T. Xu, "Transformer-Based Traffic Prediction for Adaptive SDN Routing," IEEE Transactions on Network and Service Management, Vol. 20, No. 2, pp. 1341–1354, June 2023. DOI: 10.1109/TNSM.2023.3248044
- [8] S. Ahmed and A. Roy, "An Integrated AI-Driven Framework for Real-Time Traffic Prediction and Routing in SDN," Future Generation Computer Systems, Vol. 149, pp. 373–386, 2023. DOI: 10.1016/j.future.2023.05.029
- [9] L. Zhao, Y. Chen, R. Liu, and T. He, "Lightweight AI Models for Edge-Based QoS Routing in IoT Networks," IEEE Internet of Things Journal, Early Access, 2024. DOI: 10.1109/JIOT.2024.3382987
- [10] M. Lee and H. Kim, "Graph Neural Networks for QoS-Aware Routing in Software-Defined Networks," Computer Networks, Vol. 240, p. 110022, 2024. DOI: 10.1016/j.comnet.2024.110022
- [11] M. Chen, Y. Hao, Y. Li, and K. Hwang, "Intelligent Routing Based on Deep Reinforcement Learning in Software-Defined Networks," IEEE Transactions on Network and Service Management, vol. 17, no. 4, pp. 2174–2187, Dec. 2020. DOI: 10.1109/TNSM.2020.3034066
- [12] H. Zhang, Y. Wang, J. Liu, and K. Xu, "Traffic Prediction for SDN with Attention-Based Sequence Models," Computer Networks, vol. 193, 108147, Mar. 2021. DOI: 10.1016/j.comnet.2021.108147
- [13] A. F. Atya, R. A. Khokhar, and M. A. Gregory, "A Survey on Traffic Engineering for Software-Defined Networks: State of the Art and Future Directions," IEEE Communications Surveys & Tutorials, vol. 22, no. 2, pp. 1304–1335, 2020. DOI: 10.1109/COMST.2020.2971777

Author



Moon-Sik Kang is a Professor in the Department of Electronic Engineering at the College of Engineering, Gangneung-Wonju National University, South Korea. He currently serves as Vice President of the

Institute of Electronics and Information Engineers (IEIE). He received his B.S. and M.Eng. degrees in Electronic Engineering from Yonsei University, South Korea, in 1985 and 1988, respectively, and earned his Ph.D. in the same field from Yonsei University in 1993. Dr. Kang held a postdoctoral research position in the Department of Electrical and Electronic Engineering at the University of Pennsylvania, USA, where he was involved in advanced research activities. He also served as a research associate in the Department of Electronic and Computer Engineering at the Illinois Institute of Technology, Chicago, USA. His current research focuses on high-performance networking architectures, artificial intelligence, and big data analytics, with particular emphasis on mobile multimedia traffic modeling and applications incorporating large language models (LLMs).