

지능형사이버보안연구실

# Network Slicing Meets Artificial Intelligence: an AI-based Framework for Slice Management

논문 세미나

발표자

정유진

발표일

2025.7.22

CONTENTS

# 목차

01

Overview

02

Introduction

03

Method

04

Experiment

05

Conclusion

1

## Overview

1-1

논문 소개

- Title
  - Network Slicing Meets Artificial Intelligence : an AI-base Framework for Slice Management
- Proceeding Conference
  - IEEE Communication Magazine
- Authors
  - Dario Bega, Marco Gramaglia, Andres Garcia-Saavedra, Marco Fiore, Albert Banchs, Xavier Costa-Perez
- Date of Publication
  - 2020.7.15
- Citation
  - 59
- 논문 선정 사유
  - 네트워크 슬라이싱과 AI를 다루고 있어서
  - 잘 모르는 것을 알고 싶어서

## 2

## Introduction

2-1

개념 / 도전 과제

2-2

AI 필요성 / 논문 구성 안내

## 개념

## ■ 네트워크 슬라이싱

- 물리적 인프라 -> 논리적 인스턴스로 나누기
- QoS 맞춤 네트워크

## ■ 슬라이스

- 가상 머신, 컴퓨팅 자원, 백홀 및 무선 통신 자원 등을 포함한 독립적인 서비스 단위

## 도전 과제

## ■ 효율적, 격리된 슬라이스

## ■ 맞춤형 기능

## ■ 쟁점

- 효율적인 실시간 동적 할당
- 독립적인 운용
- 공존
- 동적인 물리적 자원 공유

## AI 필요성

## ■ 인간

- 비효율적, 부적합

## ■ 자동화

- 보다 큰 용량
- 예측

## 논문 구성 안내

## ■ Section 2

- AI의 활용 가능성

## ■ Section 3

- AI를 이용한 네트워크 슬라이싱 3가지

## ■ Section 4

- 실험 결과

## ■ Section 5

- Conclusion

## 3

## Methodology

3-1

AI-BASED SLICE MANAGEMENT FRAMEWORK

3-2

AI-BASED SLICE MANAGEMENT FUNCTIONS



## Management functions in network slicing

### ■ Tenants

- 슬라이스 요청자

### ■ Admission Control

- 슬라이스 생성 전 (Preparation)
- 새 슬라이스 요청 수락 여부 결정 -> 출력

### ■ Resource Orchestration

- 네트워크 자원을 슬라이스에 동적 할당
- 슬라이스 활성화된 후 실시간 최적화

### ■ Radio Scheduling

- 무선 자원 (RAN 등) 세밀하게 배분

## Bringing AI into the picture

### ■ 강화 학습 Reinforcement learning

- 행동 선택 -> 보상 -> 최대화 -> 학습

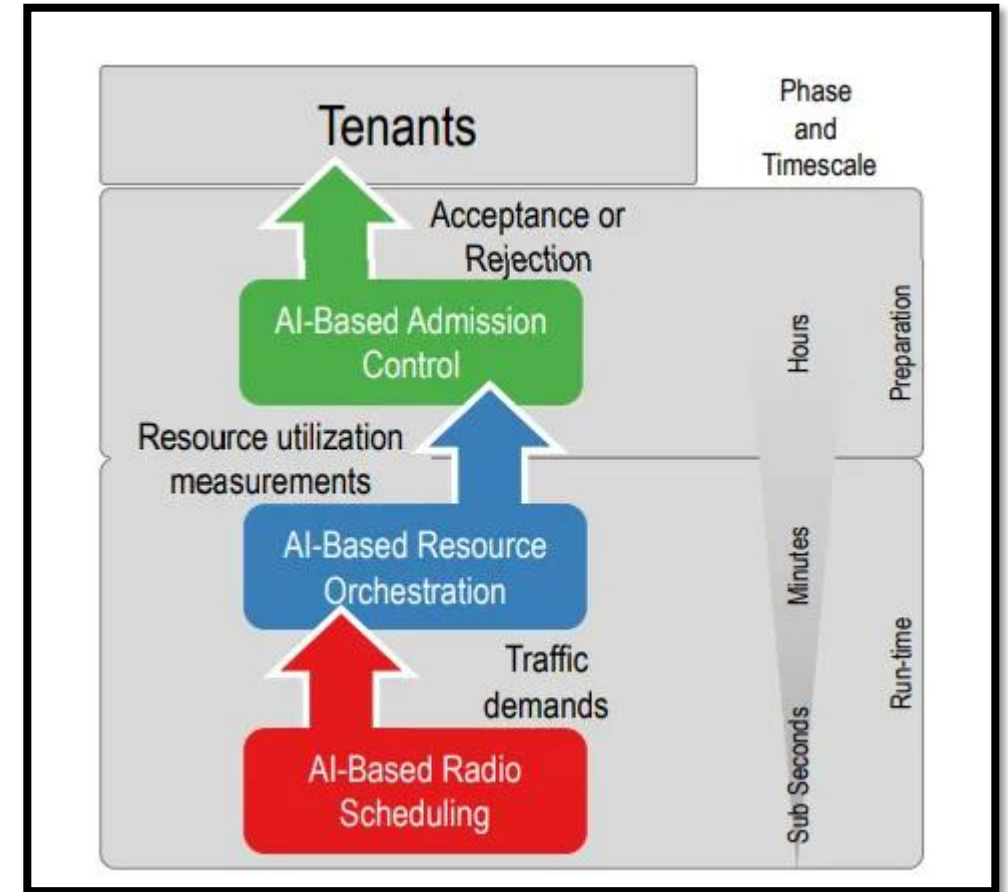


그림 1. Network slicing framework

## 3-2 AI-BASED SLICE MANAGEMENT FUNCTIONS

### AI for admission control of slices

- 한정적인 자원 -> 슬라이스 요청 수락 여부 판단 Admission Control
  - 자원 공유 <-> KPI 보장 간의 균형
    - Key Performance Indicator 성능 측정 핵심 지표: 지연 시간, 처리량, 가용성 등등
  - 해결책: AI (Deep Reinforcement Learning) 활용
    - 자원 상태와 슬라이스 도착 패턴 학습
    - 장기 수익 극대화
1. 입력 -> accept, reject 장기 수익 예측
  2. 예측 비교 -> 더 높은 행동 선택
    1.  $\epsilon$ -greedy 전략: 덜 좋은 행동도 시도
  3. 시스템 적용
  4. 예측 vs 실제 비교 -> 손실 계산
  5. NN(인공 신경망) 업데이트 -> 모델 발전

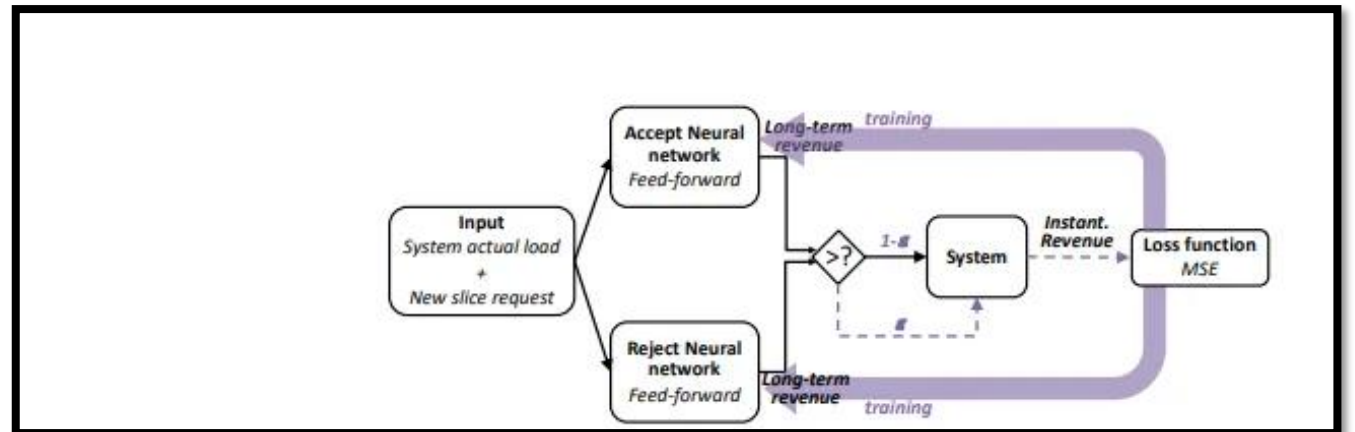


그림 2. AI for admission control of slices

## 3-2 AI-BASED SLICE MANAGEMENT FUNCTIONS

### AI for network resource orchestration

- 운영 비용을 최소화하면서 SLA(정해진 서비스 품질) 위반 X
- 과소 할당, 과대 할당 주의
- **기존 방식:** 정확한 수요 예측 - MAE(단순 예측 오차) 최소화
  - 과소 할당
  - SLA 위반 가능성 높음
  - Error가 음수
- **해결책:** 3D CNN 기반 Capacity Forecasting
  - 3차원 데이터를 처리하는 딥러닝 모델: 시계열 데이터 분석
  - 과잉 할당
  - SLA 위반 X
  - Error가 양수

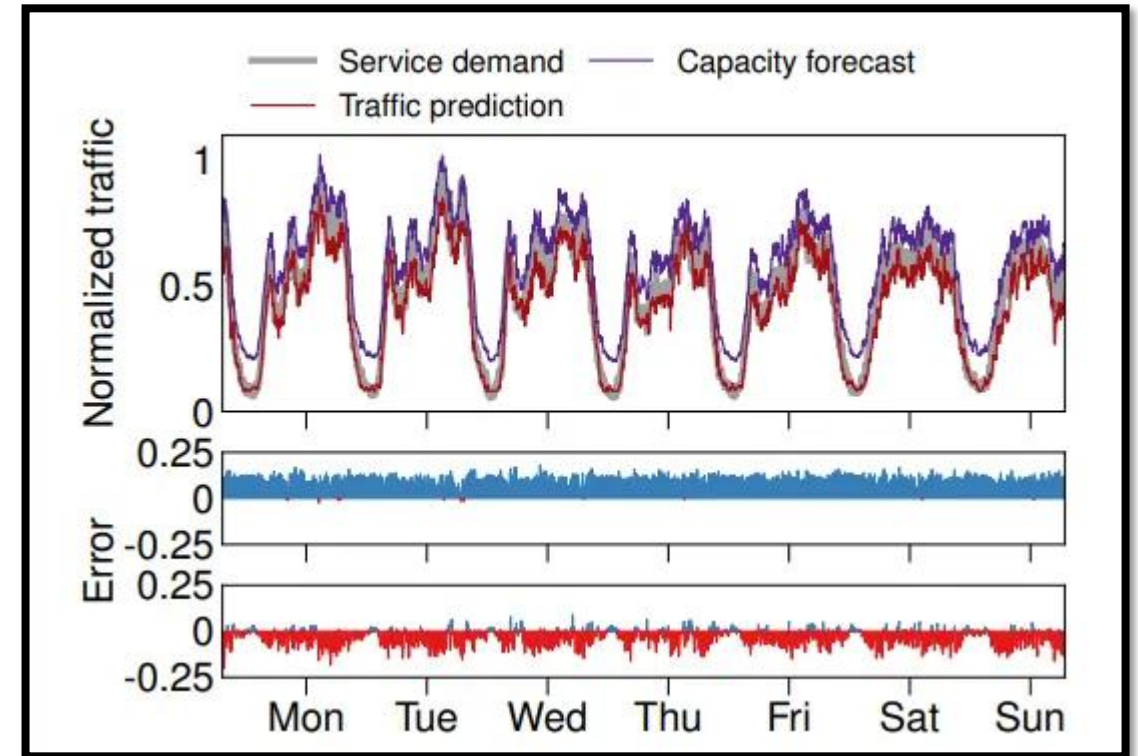


그림 3. AI for network resource orchestration

## 3-2 AI-BASED SLICE MANAGEMENT FUNCTIONS

### AI for slice scheduling at radio access

- vRAN -> 슬라이스간 isolation 유지 + dynamic하게 자원을 슬라이스 별로 배분
- 해결책: AI 기반 자원 스케줄링
- DDPG (Deep Deterministic Policy Gradient)
  - 강화 학습: 행동 선택 -> 보상 -> 학습
  - 연속적인 행동 -> 최적화
- Autoencoder
  - 비지도 학습 모델: 입력 데이터 압축 및 복원 목표
- 빠른 반응성, 높은 성능, 독립성
- 그림
  - X축: CPU 시간 할당 비율
  - Y축: 데이터 전송 효율성
  - 색: Throughput
  - Load 높을수록 CPU 자원 할당 사용자 버퍼링, 디코딩 오류 증가

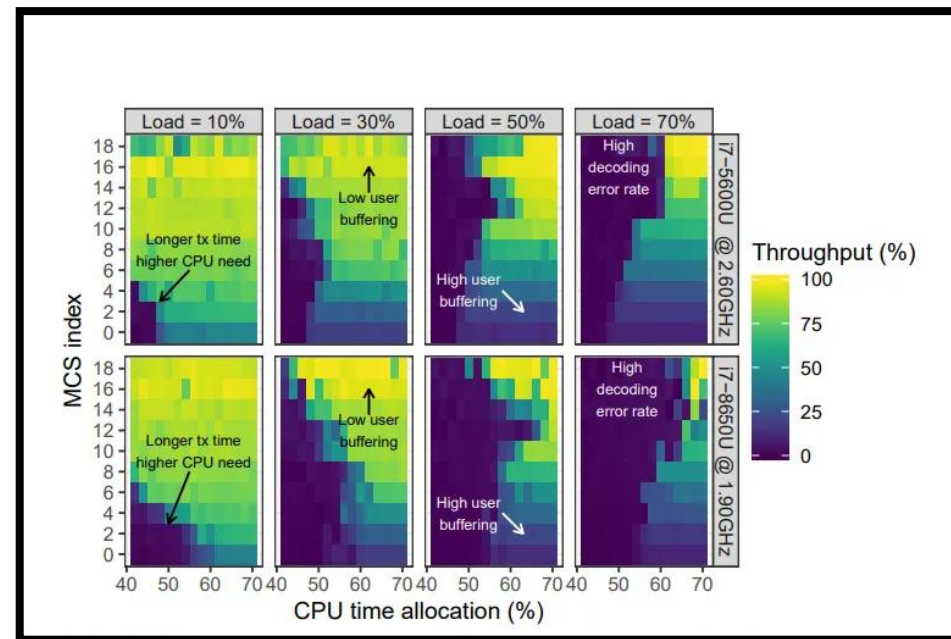


그림 4. AI for slice scheduling at radio access

## 4

## Experiment

4-1

Increasing the infrastructure monetization with optimal admission policies

4-2

Real world and data-driven capacity forecasting

4-3

AI-based CPU optimization at the edge

- 서로 다른 유형의 슬라이스 (eMBB 저가, URLLC 고가)를 수락할 때 수익이 극대화 되는가
- 기준
  - 최적화, 랜덤
  - AI for admission control of slices 사용
- 결과
  - +33.3% 수익 개선
  - 수익 손실 AI -0.23%, 최적 정책 -3.77%
  - 계산 비용이 최적에 가까운 성능

| Function                        | Performance metric  | Percent improvement – Benchmark use case   |
|---------------------------------|---------------------|--|
| Network Slice Admission Control | Revenue improvement | -0.23% – Optimal, ratio 1<br>-3.77% – Optimal, ratio 20<br>33.3% – Random Policies, ratio 15 |

표 1. 실험 1

## ■ 사례

- Facebook 트래픽
- Snapchat 트래픽
- YouTube 트래픽

■ 기존 방식: MAE 기반 예측 + 5% 과잉 할당 으로 SLA 위반 방지

■ AI 방식: AI for network resource orchestration 사용

## ■ 결과

■ 기존 방식보다 AI 방식으로 예측 + 최적 자원 할당이 비용 절감에 성공

|                           |                            |  |
|---------------------------|----------------------------|--|
| Cloud resource allocation | Reduction of monetary cost | 81.6% – Facebook, Core datacenter<br>59.2% – Snapchat, MEC datacenter<br>64.3% – YouTube, C-RAN datacenter |
|---------------------------|----------------------------|--|

표 2. 실험 2



- 사용 방법: AI for slice scheduling at radio access
- 시나리오 1 - 무제한 자원
  - QoS 성능 vs CPU 사용량
  - QoS 성능 25% 향상
  - CPU 30% 절약
- 시나리오 2 - 자원 제약
  - eNB 2개가 CPU를 공유
  - Throughput 25% 증가
  - 디코딩 오류 거의 없음

|                                 |                                |   |
|---------------------------------|--------------------------------|---|
| Virtual RAN resource allocation | CPU savings, delay, throughput | <b>30%</b> – CPU savings over CPU-blind schedulers<br><b>25%</b> – Delay-based QoS over CPU-blind schedulers<br><b>25%</b> – Throughput upon computational capacity deficit |
|---------------------------------|--------------------------------|---|

표 3. 실험 3



## 4-3 AI-based CPU optimization at the edge

- 주기마다 RAP(무선 접속 지점) 부하, SNR 수집
- vrAIIn(AI 기반 제어 시스템)
  - Encoder: 입력된 context를 embedding 사용
  - Controller: AI 모델 이용 -> RAP별 최적의 CPU 및 무선 자원 할당
- CPU 정책 - CPU scheduler, 라디오 정책 - Radio scheduler로 동적 반영
- KPI 측정 - 보상 집 계
- 주기 별로 반복 -> AI가 자원 동적 관리 -> 최적화

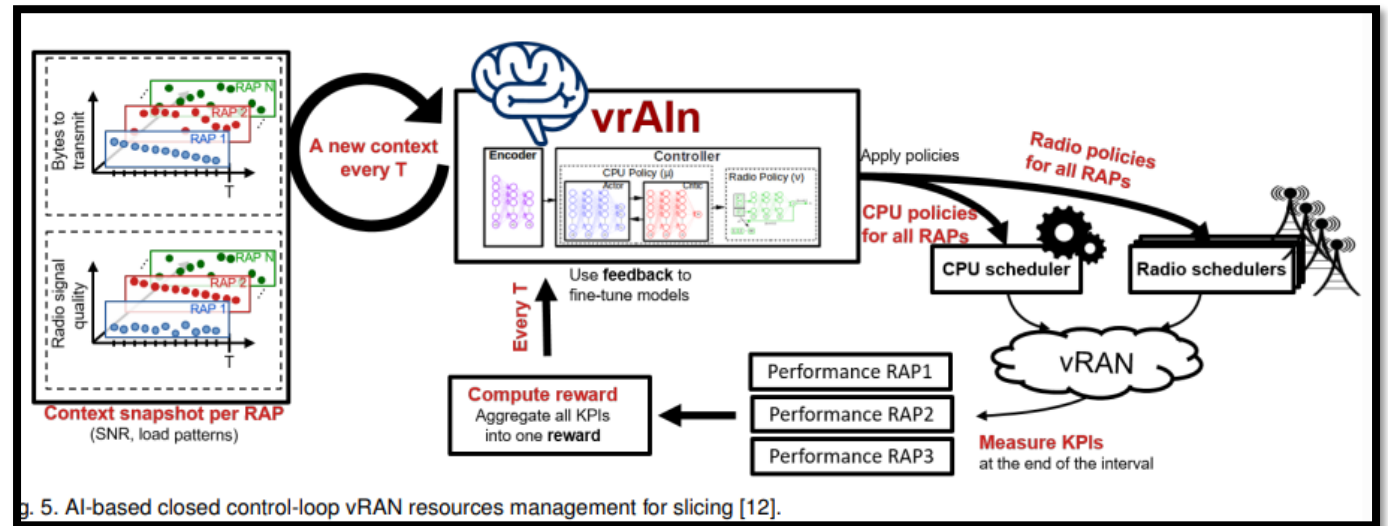


Fig. 5. AI-based closed control-loop vRAN resources management for slicing [12].

그림 5. AI based closed control-loop vRAN resources management for slicing

5

## Conclusions

5-1

결론

- AI: 효율적인 네트워크 슬라이스 관리를 위한 핵심 기술
  - 복잡하고 동적인 슬라이싱 환경에서 기존 방식의 한계 극복
- 성능 개선
  - AI 기반 자원 관리로 QoS 개선, 자원 최적화, 수익 극대화
  - 기존 방식 대비 최대 80% 향상
- 적용 방법들
  - Admission Control, Resource Orchestration, Radio Scheduling
- 앞으로의 과제
  - 신뢰성(Trustworthiness), 해석 가능성(Interpretability) 확보 필요

# CONTENTS

## Q&A