

지능형 소프트웨어 정의 인지무선 네트워크의 M&S를 위한 DDPG 기반 트래픽 분산 컨트롤러 설계 및 구현

이철웅[°], 이호진^{*}, 이규민^{*}, 이병규^{*}

Design and Implementation of DDPG-based Traffic Distribution Controller for M&S of Intelligent Software Defined Cognitive Radio Network

Cheol-woong Lee[°], Hojin Lee^{*}, Gyu-min Lee^{*}, Byungkyu Lee^{*}

요약

컴퓨터의 연산 능력과 무선통신 기술이 향상되면서 다양한 소형 무선 장치들이 등장하였으며, 사물인터넷 기술을 기반으로 하는 다양한 서비스가 출시되었다. 따라서 주파수 자원에 대한 수요가 증가하여, 기존 방식으로는 안정적으로 통신을 수행하기 어려울 만큼 주파수 대역이 포화되었다. 이러한 문제를 극복하기 위해 주파수 자원의 사용 효율 향상을 위한 CR 네트워크 구성 기술들이 제안되었으나, 무선환경 인지를 위해 고려해야 할 요소가 많아 CR 네트워크를 더욱 효과적으로 구성하기 위한 아키텍처가 적용될 필요가 있다. 본 논문에서는 SDN 기술을 지원하는 CR 네트워크 아키텍처를 제안한다. 또한, 제안 구조를 갖는 CR 네트워크에 적용 가능한 DDPG 기반 부하 분산 컨트롤러 아키텍처를 제안한다. 본 논문의 실험결과는 제안하는 컨트롤러가 CR 네트워크의 트래픽을 효과적으로 분산시켜서 주파수 자원이 효율적으로 사용될 수 있음을 보인다.

Key Words : Software Defined Networking, Cognitive Radio Network, Deep Deterministic Policy Gradient, Traffic Distribution Controller, Modeling and Simulation

ABSTRACT

As computer processing power and wireless communication technology have improved, various small wireless devices have appeared, and various services based on Internet of Things technology have been released. Therefore, the demand for frequency resources has increased, and the frequency band has become saturated to the point that it is difficult to perform stable communication using existing schemes. Many configuration technologies for the CR network have been proposed to improve the efficiency of using frequency resources, but since there are many factors to consider for recognizing the wireless environment, an architecture needs to be applied to configure the CR network more effectively. In this paper, we propose a CR network architecture that supports SDN technology. In addition, we propose a DDPG-based load balancing controller architecture applicable to CR networks with the proposed structure. The experimental results in this paper show that the proposed controller effectively distributes traffic in the CR network so that frequency resources can be used efficiently.

※ 본 연구는 LIG넥스원의 지원을 받아 수행되었습니다.

° First and Corresponding Author : LIG Nex1 Co. Ltd., cheolwoong.lee@lignex1.com, 정회원

* LIG Nex1 Co. Ltd., {hojin.lee, gyumin.lee, bklee.4}@lignex1.com

논문번호 : 202311-143-B-RN, Received November 15, 2023; Revised April 28, 2024; Accepted May 10, 2024

I. 서 론

컴퓨터의 연산 능력과 무선통신 기술이 향상되면서, 다양한 소형 무선 장치들이 등장하였다. 또한, 사물인터넷 기술을 기반으로 하는 다양한 서비스가 출시되어, 주파수 자원에 대한 수요가 증가하였다¹⁾. 이에 따라 기존 방식으로는 안정적으로 통신을 수행하기 어려울 만큼 주파수 대역이 포화되어, 이러한 문제를 극복하기 위한 기술이 요구된다²⁾.

인지무선(CR: Cognitive Radio)은 주파수 포화 문제 해결을 위한 기술 중 하나로, 주사용자(PU: Primary User)가 점유하지 않는 주파수 자원을 부사용자(SU: Secondary User) 임시로 사용해서 주파수 사용 효율을 향상시킬 수 있도록 한다^{3,4)}. CR 네트워크(CRN: Cognitive Network)에서는 무선 환경을 측정하여 얻은 정보들을 메트릭화하여 네트워크 알고리즘에 활용한다. 기존 알고리즘과는 달리, CRN에서의 네트워킹 알고리즘은 PU 및 SU의 채널 사용량 등 추가적으로 고려해야 할 요소가 많아 구성 방식이 복잡하다⁸⁻¹⁰⁾.

CRN에서 PU의 주파수 자원 점유 및 다른 SU의 자원 사용이 SU 트래픽 전달 성능에 큰 영향을 미친다. 따라서, CRN의 성능 향상을 위해서는 노드별 채널 사용량에 따른 트래픽 별 전달 경로 제어 방식이 요구된다. 그러나, 기존의 CRN 대상 라우팅 기술은 애드혹 기반의 무선 환경 정보 교환을 통해 분산적으로 라우팅 테이블을 구성한 후에 트래픽을 전달하는 방식을 채용하기 때문에, 트래픽 별 경로 선택 방식을 반영하기 어렵다¹⁰⁻¹²⁾. CRN을 대상으로 제안된 일부 라우팅 기술들은 다양한 측면에서 무선환경을 고려하면서 네트워크의 성능을 향상시키기 위해 인공지능(AI: Artificial Intelligence) 및 기계학습(ML: Machine Learning)을 기반으로 동작하기도 한다¹³⁻¹⁷⁾. 하지만, 이들 또한 트래픽별 경로 선택 방식의 지원이 어렵고, 기존 방식보다 높은 전력소비량을 보이는 AI 또는 ML 기술을 무선장치에서 분산 방식으로 동작시키기에 한계가 존재한다.

CRN에 소프트웨어 정의 네트워킹(SDN: Software Defined Networking) 방식을 채용할 경우, 중앙 집중형 또는 계층형으로 네트워크를 제어할 수 있다. 이에 따라, 트래픽 별 경로 선택 방식을 기반으로 CRN의 성능 뿐만 아니라 네트워크 제어 기술의 적용성 및 확장성을 향상시킬 수 있다. 컨트롤러에는 수집한 무선환경에 대한 정보를 기반으로 동작하는 AI 또는 ML 기술을 손쉽게 적용할 수 있다. 하지만, 현재는 SDN 방식을 채용한 CRN을 대상으로 주파수 자원 관리 기술이 주로 제안되고 있다¹⁸⁻²²⁾.

본 논문에서는 SDN 지원 CRN의 아키텍처를 제안하고, 제안 구조를 갖는 CRN에 적용 가능한 강화학습 기반 부하 분산 컨트롤러의 아키텍처를 제안한다. 또한, 대표적인 강화학습 알고리즘인 심층 결정적 정책 경사법 (DDPG: Deep Deterministic Policy Gradient) 기반의 SDN 지원 CRN 환경 모델링 결과를 보인다. 이후, 모델링 및 시뮬레이션 (M&S: Modeling and Simulation) 결과를 통해 트래픽 분산이 효과적으로 이루어질 수 있음을 보인다.

본 논문의 구성은 다음과 같다. 2장에서는 강화학습의 개요 및 DDPG에 대해 설명하며, CRN을 위한 AI 및 ML 기반 라우팅 연구와 CRN을 위한 SDN 기술 연구에 대해 설명한다. 3장에서는 SDN 기술을 적용하기 위한 CRN 아키텍처에 대해 설명하며, 제안 CRN에서 구동시킬 수 있는 DDPG 기반 트래픽 분산 라우팅 기술을 제안한다. 4장에서는 제안방법의 M&S 결과에 대해 설명하며, 5장에서 결론을 내린다.

II. 배경 및 관련연구

2.1 강화학습 및 DDPG 개요

강화학습은 변화하는 환경에서 시행착오를 통해 현재 상태에 대해 최적의 행동이 무엇인지 학습하는 기술의 종류로, Q-learning이 기본적인 강화학습 알고리즘으로 유명하다.

문제의 규모가 클수록 학습을 위한 메모리 요구량이 증가하는 Q-learning의 문제를 극복하기 위해, 인공신경망을 기반으로 상태 및 행동에 따른 보상을 저장하기 위한 Q-테이블을 함수로 근사하는 Deep Q Network (DQN)이 제안되었으나⁵⁾, 행동을 이산공간 값으로 가정하여 행동공간이 연속적인 환경에는 부적합하다.

심층 결정적 정책 경사법 (DDPG: Deep Deterministic Policy Gradient)은 상기 강화학습 알고리즘들의 문제를 해결하기 위해, Deep Policy Gradient (DPG)와 Deep Q Network (DQN)의 구조를 결합한 형태를 가진다⁶⁾. DPG의 Actor-Critic 구조는 주어진 상태에서 행동의 보상을 학습시켜 연속적인 행동공간에서 학습을 수행할 수 있도록 하며, DQN의 Replay Buffer 구조는 학습 결과가 편향되지 않도록 샘플들 사이의 상관관계를 제거한다. DDPG는 상기 설명한 장점을 가지고 있어, 복잡한 무선 네트워크 환경을 위한 최적화 기술들의 기반으로써 활용되고 있다⁷⁾.

2.2 CRN을 위한 AI 및 ML 기술 연구

인지무선 네트워크(CRN: Cognitive Radio

Network)를 위한 네트워킹 알고리즘의 경우¹⁸⁻¹⁰⁾, 부사용자(SU: Secondary User)의 트래픽 송신을 위한 자원 할당 및 경로 선택에 있어 기존 네트워킹 알고리즘의 고려사항과 더불어 주사용자(PU: Primary User) 및 SU의 채널 사용량을 포함한 다양한 요소를 고려하기 때문에, 구성 방식이 복잡하다.

CRN에서 PU의 주파수 자원 점유 및 다른 SU의 자원 사용이 SU 트래픽 전달 성능에 큰 영향을 미친다. 따라서, CRN의 성능 향상을 위해서 노드별 채널 사용량에 따른 트래픽 별 전달 경로 제어 방식이 요구된다. 그러나, CRN 대상 애드혹 라우팅 방식에서는¹⁰⁻¹²⁾ 각 노드가 분산 라우팅에 의해 구성된 테이블을 기반으로 트래픽을 전달한다. 따라서, 트래픽별 경로 선택 방식이 적용되기 어렵다. 일부 CRN 멀티 홉 라우팅 기술들은¹³⁻¹⁷⁾ 무선환경을 고려하기 위한 다양한 메트릭을 기반으로 성능을 향상시키기 위해 인공지능(AI: Artificial Intelligence) 및 기계학습(ML: Machine Learning) 기술을 기반으로 동작하기도 한다. 에너지 사용 효율을 고려한 Q-learning 기반 라우팅 연구는¹³⁾ 처리량을 향상시키면서 전력소비를 최소화하기 위한 분산 라우팅 알고리즘을 제안한다. 클러스터 기반 AI 라우팅 연구는¹⁴⁾ 전력소비를 고려한 라우팅 알고리즘과 함께 Cluster Header (CH) 선정 방법을 제안한다. 협력적 스펙트럼 센싱 기반 Q-라우팅 연구는¹⁵⁾ SU 트래픽 전송 경로의 단절을 최소화하기 위해, 에너지 인과성, 센싱 신뢰성, 간섭 임계값 및 개별 링크 처리량을 정의하여 학습에 활용한다. 스펙트럼 학습 기반 AI 라우팅 연구는¹⁶⁾ PU의 주파수 자원 사용 경향을 Support Vector Machine (SVM)을 기반으로 분석하며, 그 결과를 라우팅 경로 학습에 활용한다. CR 지원 사물인터넷(IoT: Internet of Things) 대상 AI 라우팅 연구는¹⁷⁾ 높은 처리량 및 낮은 단대단 지연시간을 달성하기 위한 라우팅 기술을 제안한다. 하지만, 이들 역시 트래픽별 경로 선택 방식이 적용되기 어려우며, AI 또는 ML 기술을 무선장치에서 분산 방식으로 동작시키기에는 여전히 전력소비 측면에서 한계가 존재한다.

2.3 CRN을 위한 AI 및 ML 기술 연구

CRN에 소프트웨어 정의 네트워킹(SDN: Software Defined Networking) 방식을 채용할 경우, 중앙 집중형 또는 계층형으로 망을 제어할 수 있다. 이 경우, 트래픽별 경로 선택 방식의 적용이 가능해지고, 망 제어 기술의 적용성 및 확장성이 향상된다. 특히, 네트워크 구성 노드들이 수집한 무선환경에 대한 정보를 컨트롤러가 활용하기 때문에, AI 또는 ML 기술을 보다 손쉽게 적용

가능하다.

SDN 지원 CRN 아키텍처 기반의 주파수 자원 관리 연구들은¹⁸⁻²¹⁾ SDN 컨트롤러가 스펙트럼 센싱 결과를 종합하여 SU에 주파수 자원을 할당하는 방식을 제안하며, 별도의 라우팅 기술을 제안하지 않는다. SDN 지원 미래 이동 네트워크를 위한 라우팅 기반 간섭 방지 기술 연구들은²²⁾ 컨트롤러를 기반으로 무선 신호 간섭이 적은 구간을 선택하는 라우팅 기술을 제안한다, SDN 지원 CRN 환경에서 병목구간을 최소화하기 위한 트래픽 분산 기술을 제안하지 않는다.

III. SDN 방식을 채용한 CRN을 위한 DDPG 기반 트래픽 분산 라우팅 기술

3.1 SDN 기술의 적용을 위한 CRN 아키텍처

SDN 기술을 적용하기 위한 계층적인 CRN 아키텍처를 그림 1과 같이 나타낼 수 있다. CRN은 크게 SU 네트워크 도메인, PU 네트워크 도메인, 그리고 컨트롤 도메인으로 구분된다.

SU 네트워크들은 PU들과 연결되며, SU 네트워크를 통틀어서 SU 네트워크 도메인이라 부른다. SU 노드들은 자신과 연결된 PU 노드를 통해 컨트롤러에 사전 등록한다. SU 노드들은 CR 매체접근제어(MAC: Medium Access Control) 기술을 이용하여 PU 노드들에 의해 주파수 자원이 점유되지 않았음을 확인하고, 이후 자신의 데이터를 전송 시도한다. 만약 다른 SU 네트워크에 있는 노드에 데이터를 전송하고자 할 경우, 소스 SU 노드가 자신과 연결된 PU 노드에 릴레이를 요청해야 한다.

PU 네트워크 도메인은 상호 연결된 PU 노드들로 이루어져 있다. PU들은 인가된 스펙트럼에 상시 접근

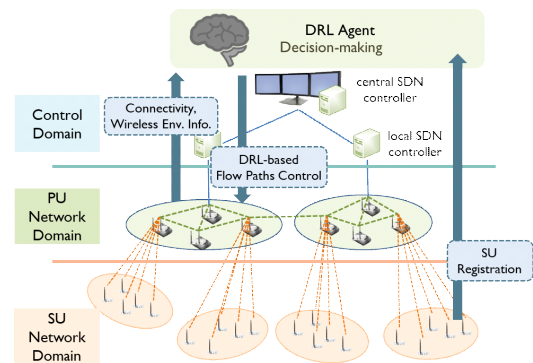


그림 1. SDN 기술의 적용을 위한 계층적 CRN 아키텍처
Fig. 1. Hierarchical CRN Architecture for applying the SDN technology

하여 상호 통신 할 수 있으며, SU의 딜레이 요청을 받을 경우, SDN 컨트롤러에 의해 계산된 경로를 통해 트래픽을 전달한다.

컨트롤 도메인은 SDN 컨트롤러가 존재하는 네트워크 도메인으로, 트래픽 경로를 계산하여 PU 네트워크 도메인으로 전달한다. 다음 절에서 설명할 DDPG 기반 라우팅 기술은 컨트롤 도메인의 SDN 컨트롤러 내부에서 Deep Reinforcement Learning (DRL) 에이전트에 의해 동작한다.

3.2 CRN에서 트래픽 분산을 위한 DDPG 기반 라우팅 기술

트래픽 분산 라우팅 알고리즘을 적용하지 않는 경우, 사용 가능한 대역폭이 고려되지 않는다. 따라서, SU가 송신하는 트래픽이 병목구간을 형성하여 해당 경로를 통해 전달되는 트래픽의 전달 성능이 낮아지고, 결국 SU 네트워크 도메인의 성능이 낮아질 수 있다.

본 논문에서 제안하는 SDN 기반 CRN 아키텍처에는 심층 결정적 정책 경사법 (DDPG: Deep Deterministic Policy Gradient) 기반 트래픽 분산 라우팅 기술을 적용한다. 즉, SDN 컨트롤러가 SU 노드의 송신 트래픽을 분산시켜서 PU가 주파수 자원을 많이 사용하는 경우에도 트래픽이 특정 노드 또는 링크에서 정체되지 않도록 제어한다.

그림 2는 SDN 컨트롤러에서 동작하는 DDPG 기반 라우팅 기술의 시뮬레이션을 위한 아키텍처를 나타낸다. CRN 환경에 대한 정보를 나타내는 C , T_{SU} , 그리고 U_{PU} 는 파라미터화되어 컨트롤러 내부에 사전 저장되

며, 시뮬레이션 시작시에 1번 전달된다. 네트워크에 n 개의 노드가 있다고 하자. 이때, C 는 노드 사이의 연결 관계를 나타내는 $n \times n$ 행렬, T_{SU} 는 SU에 의해 송수신되는 트래픽별 소스 및 목적지 행렬, 그리고 U_{PU} 는 PU의 링크 사용률 행렬을 의미한다. DRL 에이전트는 t 번째 ($t=1,2,3,\dots$) 링크 가중치 행렬의 초기값 W^{init}_t 를 계산하여 컨트롤러에 전달한다. 컨트롤러는 CRN 파라미터와 W^{init}_t 또는 $W^{updated}_t$ 를 이용하여 트래픽 전달 경로를 계산한 이후, 다음 $W^{updated}_t$ 를 계산하는 과정을 반복한다. 이때, $W^{updated}_t$ 는 트래픽 전달 경로에 해당하는 링크에 대해, W^{init}_t 또는 $W^{updated}_t$ 의 최소 가중치 값을 더해주는 방식으로 계산된다. 모든 트래픽에 대한 경로 계산이 완료되면, 경로 계산 결과에 따른 현재 상태 S_t 와 보상 값 R_t 를 계산해서 DRL 에이전트로 전달한다. 에이전트는 S_t 와 R_t 에 대한 action으로 W^{init}_{t+1} 을 계산해서 컨트롤러에 전달한다. 컨트롤러와 에이전트는 위 과정을 반복하며, 높은 보상 값을 갖는 트래픽 경로 F_{SU} 를 PU 네트워크 도메인으로 전달한다.

현재 상태를 나타내는 행렬 S_t 에는 링크의 사용률을 나타내는 행렬 U^{link}_t 가 저장된다. U^{link}_t 의 i 번째 행 j 번째 열에 해당하는 요소인 $U^{link}_t[i,j]$ 는 노드 i 와 노드 j 를 연결하는 링크에 대한 사용률을 나타낸다. 1홉 관계인 PU i 와 j 가 각 노드의 1홉 이웃으로 전송하는 트래픽에 의한 링크 사용률을 각각 $U_{PU}[i]$ 와 $U_{PU}[j]$ 라 하자. 그리고 t 번째 라우팅 결과에 따라 PU i 와 j 가 k 로 전달하는 SU 트래픽에 의한 링크 사용량을 각각 $U^{su}[i,k]$ 와 $U^{su}[j,k]$ 라 하자. 또한, 노드 i 의 이웃으로 구성된 집합을 N_i 라 하자. 이때, $U^{link}_t[i,j]$ 는 다음 식과 같이 계산된다. 이때, B 는 링크의 대역폭을 나타낸다.

$$U^{link}_t[i,j] = \frac{U_{PU}[i] + U_{PU}[j] + \sum_{k \in N_i} U^{su}[i,k] + \sum_{k \in N_j} U^{su}[j,k]}{B} \quad (1)$$

CRN 환경은 컨트롤 도메인에 의한 제어 정보를 기반으로 동작하는 PU 및 SU 네트워크 도메인을 나타낸다. PU 및 SU 네트워크 도메인은 네트워크 시뮬레이터를 기반으로 모델링하거나 실장비를 이용하여 테스트베드 형태로 구성할 수 있다. CRN에 적용되는 파라미터를 컨트롤 도메인에 전달하기 위해서는 수동 작업을 수행하거나, 컨트롤 도메인에 사전 전달하는 형태로 구성해야 한다.

DRL 에이전트에 의해 반복되며 얻어지는 보상 값을 어떤 목적으로 설계하는지에 따라, 트래픽 전송 경로가

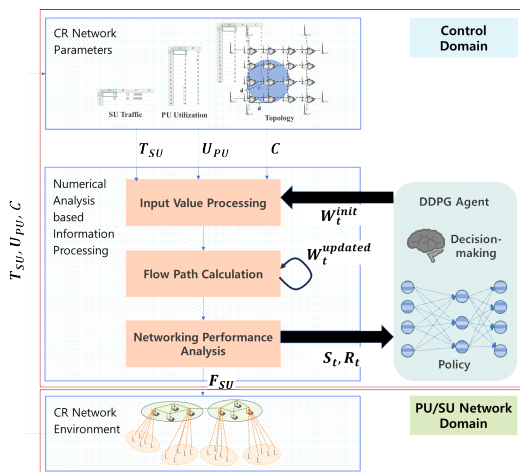


그림 2. DDPG 기반 트래픽 전달 경로 계산 기술의 시뮬레이션을 위한 아키텍처
Fig. 2. Architecture for simulation of DDPG-based traffic forwarding path calculation scheme

바뀐다. 본 절에서는 SU가 송수신하는 트래픽 전달 경로의 홉 수를 낮게 유지하면서 병목구간을 최소화하기 위한 보상 값 정의 결과에 대해 제시한다. 총 m 개의 트래픽이 발생한다고 할 때, 트래픽별 홉 수 정보를 포함하는 행렬을 \mathbf{H}_m^t 이라 하자. 또한, $E[\cdot]$, $\text{Var}(\cdot)$, 그리고 $\max(\cdot)$ 를 각각 행렬 또는 벡터의 모든 요소에 대한 평균, 분산, 그리고 최대값을 나타내는 함수라 하자. 보상 값 R_{t+1} 은 다음과 같이 계산된다. 이때, α , β , γ , 그리고 δ 는 사용자 설정 파라미터이다.

$$R_t = \frac{\alpha}{\text{Var}(\mathbf{U}_{i_n k}^t)} + \frac{\beta}{\max(\mathbf{U}_{i_n k}^t)} + \frac{\gamma}{\max(\mathbf{H}^t)} + \frac{\delta}{E[\mathbf{H}^t]} \quad (2)$$

IV. 시뮬레이션 및 분석

시뮬레이션 시나리오는 그림 3에 나타낸다. PU는 4x4 격자구조로 배치한다. PU의 번호는 좌상단부터 우측으로 1, 2, ...과 같이 지정되고 우하단 노드가 16이다. SU는 격자구조로 배치된 PU에 연결된다. SU의 번호는 좌상단부터 시계방향으로 17, 18, ...과 같이 지정된다.

노드들은 수직 및 수평으로 무선링크를 통해 연결된 노드들과 데이터 송수신이 가능하다. 이때, 무선링크를 통해 연결되지 않은 노드에는 영향을 미치지 않는다 가정한다. 무선링크의 최대 전송량은 2Mbps이다. 원으로 표시된 노드는 50%의 PU 링크 사용률을 가진다. 이는 해당 PU 노드들이 1홉 이웃과의 무선링크가 가진 대역폭의 50%만큼을 사용한다는 것을 의미한다. 트래픽의

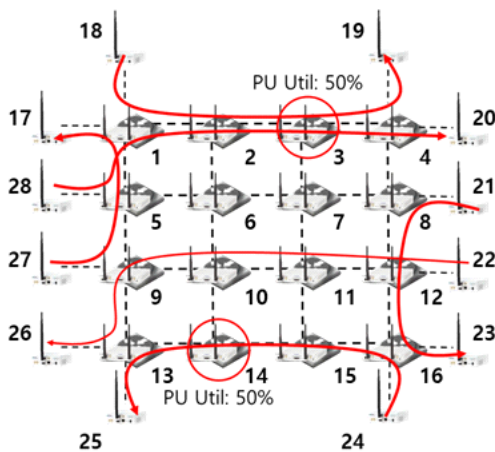


그림 3. DDPG 기반 트래픽 분산 라우팅의 시뮬레이션 시나리오
Fig. 3. Simulation scenario of DDPG-based traffic distribution routing

소스 및 목적 노드는 모두 SU이며, 화살표를 통해 확인된다. 트래픽 전송량은 300Kbps 또는 200Kbps이다.

SDN 지원 CRN의 컨트롤 도메인은 MATLAB R2020b를 이용하여 모델링했고, 대표적인 강화학습 알고리즘인 DDPG를 적용하기 위해 MATLAB의 애드온인 Reinforcement Learning Toolbox를 사용했다^[23]. AI/ML 기술의 실험 결과를 확인하기 위한 파라미터는 표 1과 같다. Episode의 횟수는 1000회이고, episode당 step의 횟수는 500회이다. 학습 속도 결정에 사용되는 learning rate는 0.005이다. 링크 사용률의 분산, 최대 링크 사용량, 트래픽 경로의 최대 홉 수, 그리고 트래픽 경로 홉 수 평균의 비율을 고려하여 강화학습 보상 값 계산 결과에 각 항이 비슷한 영향을 미치도록 α , β , γ 그리고 δ 를 각각 1000, 250, 1, 1로 설정하였다. 트래픽 분산 결과를 비교하기 위해, 최단경로 알고리즘 기반의 라우팅 및 동일 비용 다중경로 (ECMP: Equal Cost Multi Path) 라우팅^[24]과 DDPG 기반의 트래픽 분산 라우팅에 대해 시뮬레이션을 수행한다.

그림 4는 최단경로 기반 라우팅 방식을 사용했을 때, 링크별 사용률을 측정할 결과를 분포 형태로 나타낸다. 가로축은 두 노드 사이의 무선링크들을 나타내며, 각 링크는 양 끝에 위치한 두 노드의 번호로 표현된다. 이때, 링크 사용률의 색은 파란색이 40% 미만이고, 노란색이 40% 이상 80% 미만이며, 빨간색이 80% 이상을 나타낸다. 따라서, PU 링크 사용률이 50% 이상인 노드의 주변 링크에서 사용률이 80% 이상인 병목현상이 발생하는 모습을 확인할 수 있다. 반면, 그 외의 무선링크들은 다른 링크에 비해 사용률이 낮은 모습을 확인할 수 있다. 이러한 결과를 통해, 일부 트래픽의 경로를 바꿔서 병목현상을 해소할 수 있음을 보인다.

그림 5는 부하분산을 위해 사용되는 ECMP 라우팅 기술을 SDN 컨트롤러를 사용하지 않는 최단경로 라우팅과 함께 사용했을 때, 링크별 사용률을 측정할 결과를 분포 형태로 나타낸다. 결과를 통해 병목현상이 발생하는 링크의 사용률 최대치가 감소한 모습을 확인할 수

표 1. CRN 환경에서의 트래픽 분산을 위한 강화학습 파라미터

Table 1. Parameters for reinforcement learning in CRN environment

Parameter	Value			
The number of episodes	1000			
The number of steps	500			
Learning rate	0.005			
Parameters for calculation of reward (α , β , γ , δ)	$\alpha=1000$	$\beta=250$	$\gamma=1$	$\delta=1$

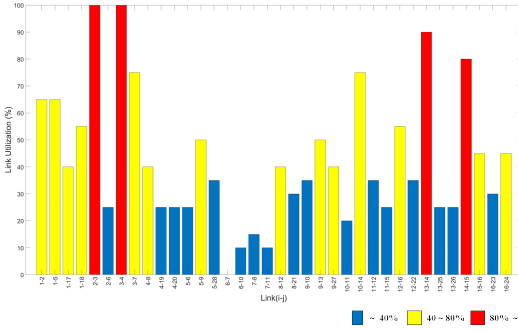


그림 4. 최단경로 라우팅에 의한 링크별 사용률 계산 결과
Fig. 4. Link utilization calculation results using shortest path routing

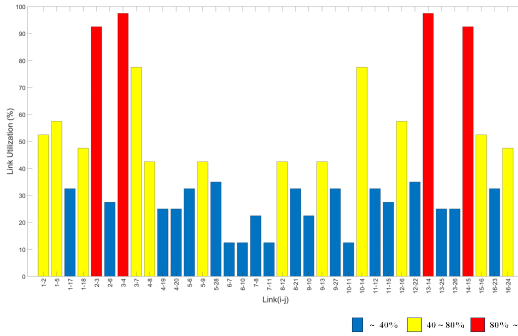


그림 5. 최단경로 라우팅 및 ECMP 라우팅에 의한 링크별 사용률 계산 결과
Fig. 5. Link utilization calculation results using shortest path and ECMP routing

있다. 하지만 병목구간의 수는 최단경로 라우팅만을 사용했을 때와 동일하고, 다른 병목구간의 링크 사용량이 증가한 모습을 확인할 수 있다. 이러한 현상이 발생하는 이유는 두 가지로, 첫째는 최소 비용을 갖는 경로에 대해서만 트래픽이 분산되기 때문이다. 둘째는 최소 비용을 갖는 모든 경로에 대해서 부하분산 기술이 동작하여 분산된 트래픽이 동일 링크를 경유할 수 있기 때문이다.

그림 6부터 그림 8까지는 DDPG 기반 트래픽 분산 라우팅 방식의 보상 값 R_t 가 5인 상황에서 수식 (2)의 분포 값에 따른 링크별 사용률 측정 결과를 분포 형태로 나타낸다. 3개의 결과를 통해 $\max(\mathbf{U}_{link}^t)$ 가 클수록 병목구간의 수가 적은 모습을 확인할 수 있다. 이러한 현상이 발생하는 이유는 동일한 보상 값에 대해 $\max(\mathbf{U}_{link}^t)$ 가 크면 $\text{Var}(\mathbf{U}_{link}^t)$ 는 작아야 하고, 링크별 사용률이 전체적으로 증가해야 $\max(\mathbf{U}_{link}^t)$ 가 크고 $\text{Var}(\mathbf{U}_{link}^t)$ 가 작은 결과를 나타내기 때문이다.

그림 9부터 그림 11까지는 DDPG 기반 트래픽 분산 라우팅 방식의 R_t 가 각각 6, 7, 그리고 8인 상황에서

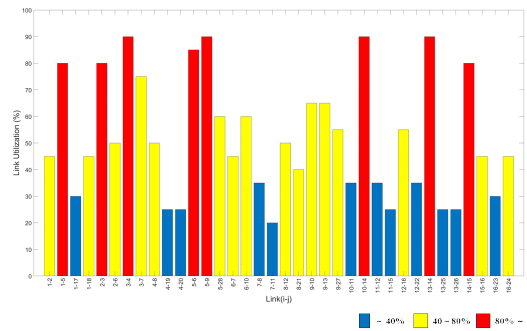


그림 6. DDPG 기반 트래픽 분산 라우팅에 의한 링크별 사용률 계산 결과 ($\text{Var}(\mathbf{U}_{link}^t)=498$, $\max(\mathbf{U}_{link}^t)=90$, $\max(\mathbf{H}^t)=9$, $E[\mathbf{H}^t]=6$, $R_t=5$)
Fig. 6. Link utilization calculation results using DDPG-based traffic distribution routing ($\text{Var}(\mathbf{U}_{link}^t)=498$, $\max(\mathbf{U}_{link}^t)=90$, $\max(\mathbf{H}^t)=9$, $E[\mathbf{H}^t]=6$, $R_t=5$)

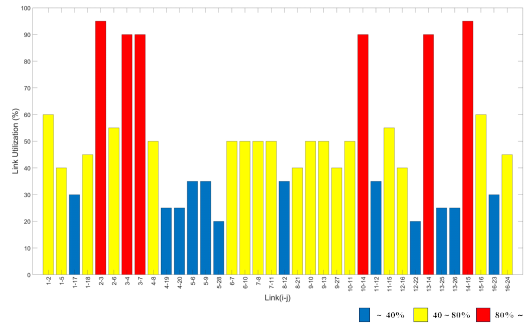


그림 7. DDPG 기반 트래픽 분산 라우팅에 의한 링크별 사용률 계산 결과 ($\text{Var}(\mathbf{U}_{link}^t)=489$, $\max(\mathbf{U}_{link}^t)=95$, $\max(\mathbf{H}^t)=6$, $E[\mathbf{H}^t]=6$, $R_t=5$)
Fig. 7. Link utilization calculation results using DDPG-based traffic distribution routing ($\text{Var}(\mathbf{U}_{link}^t)=489$, $\max(\mathbf{U}_{link}^t)=95$, $\max(\mathbf{H}^t)=6$, $E[\mathbf{H}^t]=6$, $R_t=5$)

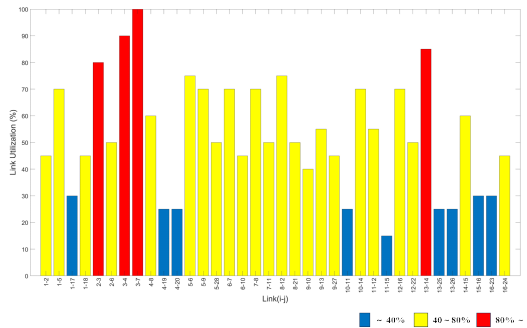


그림 8. DDPG 기반 트래픽 분산 라우팅에 의한 링크별 사용률 계산 결과 ($\text{Var}(\mathbf{U}_{link}^t)=446$, $\max(\mathbf{U}_{link}^t)=100$, $\max(\mathbf{H}^t)=9$, $E[\mathbf{H}^t]=6$, $R_t=5$)
Fig. 8. Link utilization calculation results using DDPG-based traffic distribution routing ($\text{Var}(\mathbf{U}_{link}^t)=446$, $\max(\mathbf{U}_{link}^t)=100$, $\max(\mathbf{H}^t)=9$, $E[\mathbf{H}^t]=6$, $R_t=5$)

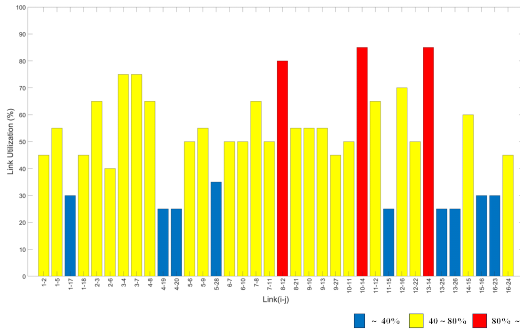


그림 9. DDPG 기반 트래픽 분산 라우팅에 의한 링크별 사용률 계산 결과 ($R_t=6$)
 Fig. 9. Link utilization calculation results using DDPG-based traffic distribution routing ($R_t=6$)

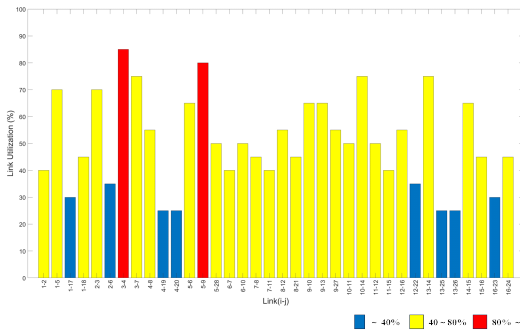


그림 10. DDPG 기반 트래픽 분산 라우팅에 의한 링크별 사용률 계산 결과 ($R_t=7$)
 Fig. 10. Link utilization calculation results using DDPG-based traffic distribution routing ($R_t=7$)

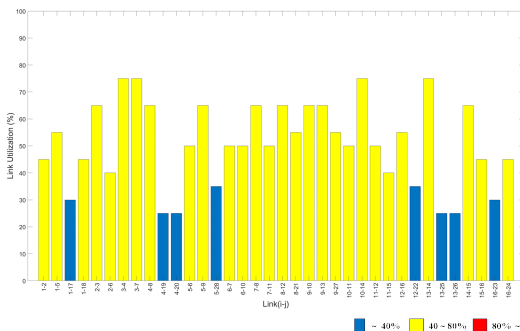


그림 11. DDPG 기반 트래픽 분산 라우팅에 의한 링크별 사용률 계산 결과 ($R_t=8$)
 Fig. 11. Link utilization calculation results using DDPG-based traffic distribution routing ($R_t=8$)

병목구간의 수가 최소일 때 링크별 사용률 측정 결과를 분포 형태로 나타낸다. 4개의 결과에서 보상 값이 클수록 병목구간의 수가 적은 모습을 확인할 수 있다.

DRL 에이전트는 보상을 크게 받는 행동을 취하도록

학습한다. 따라서, 학습이 진행될수록 병목구간의 수를 감소시키고, 우회경로에 해당하는 링크의 사용률이 증가하도록 라우팅을 수행함을 확인할 수 있다. 이때, SDN 컨트롤러가 트래픽별 제어를 수행하기 때문에 트래픽이 경유하는 링크를 선택할 수 있다. 따라서, 부하 분산이 효과적으로 이루어질 수 있다.

V. 결론 및 향후연구

본 논문에서는 SDN 컨트롤러에 의한 제어구조를 채용할 수 있는 CRN 아키텍처를 제안하였다. 또한, 제안 환경에 적용할 수 있는 DDPG 기반 SU 트래픽 분산 기술을 제안하였다. 제안 기술은 PU가 주파수 자원을 많이 사용하는 경우에도 SU 트래픽이 특정 노드 또는 링크에서 정체되지 않도록 동작한다. 컨트롤러에는 MATLAB 기반의 간단한 무선 환경 모델을 적용하였으며, 수학적인 방식으로 제안방법의 동작을 확인할 수 있도록 SDN 지원 CRN의 M&S를 수행했다. 시뮬레이션 결과는 제안하는 네트워크 제어 알고리즘이 효과적으로 트래픽을 분산시킬 수 있음을 보인다.

향후연구에서는 군에서 신뢰성을 인정받은 네트워크 시뮬레이터인 Riverbed Modeler를 기반으로 보다 복잡한 CRN 환경을 모델링하고자 한다. 그리고 링크 사용률 뿐만 아니라 다양한 환경 요소를 고려하는 AI/ML에 의한 네트워크 제어 기술을 적용한 뒤에, Riverbed Modeler 기반으로 모델링한 CRN 환경과 연동해서 네트워킹 성능을 측정 및 분석할 수 있는 SDN 기반 CRN의 시뮬레이션 환경을 구축하고자 한다.

References

[1] R. I. Ansari, H. Pervaiz, S. A. Hassan, C. Chrysostomou, M. A. Imran, S. Mumtaz, and R. Tafazolli, "A new dimension to spectrum management in IoT empowered 5G networks," *IEEE Netw.*, vol. 33, no. 4, pp. 186-193, Jul. 2019. (<https://doi.org/10.1109/MNET.2019.1800157>)

[2] Y. Zhou, L. Liu, L. Wang, N. Hui, X. Cui, J. Wu, C. Xing, et al., "Service-aware 6G: An intelligent and open network based on the convergence of communication, computing and caching," *Digital Commun. and Netw.*, vol. 6, no. 3, pp. 253-260, Aug. 2020. (<https://doi.org/10.1016/j.dcan.2020.05.003>)

- [3] J. K. Choi, M. Seo, S. Um, J. W. Choi, and C. Han, "Cooperative sensing parameter optimization based on the infimum of cooperative detection probability," *J. KICS*, vol. 46, no. 4, pp. 598-607, Jan. 2021. (<https://doi.org/10.7840/kics.2021.46.4.598>)
- [4] A. Nasser, H. A. H. Hassan, J. A. Chaaya, A. Mansour, and K. C. Yao, "Spectrum sensing for cognitive radio: Recent advances and future challenge," *Sensors*, vol. 21, no. 7, 2408, Mar. 2021. (<https://doi.org/10.3390/s21072408>)
- [5] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, "Playing atari with deep reinforcement learning," *arXiv preprint arXiv:1312.5602*, 2013. (<https://doi.org/10.48550/arXiv.1312.5602>)
- [6] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, et al., "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015. (<https://doi.org/10.48550/arXiv.1509.02971>)
- [7] N. C. Luong, D. T. Hoang, S. Gong, D. Niyato, P. Wang, Y. C. Liang, and D. I. Kim, "Applications of deep reinforcement learning in communications and networking: A survey," *IEEE Commun. Surv. & Tuts.*, vol. 21, no. 4, pp. 3133-3174, May 2019. (<https://doi.org/10.1109/COMST.2019.2916583>)
- [8] A. Kumar and K. Kumar, "Multiple access schemes for cognitive radio networks: A survey," *Physical Commun.*, vol. 38, Feb. 2020. (<https://doi.org/10.1016/j.phycom.2019.100953>)
- [9] R. A. Diab, N. Bastaki, and A. Abdrabou, "A survey on routing protocols for delay and energy-constrained cognitive radio networks," *IEEE Access*, vol. 8, pp. 198779-198800, Nov. 2020. (<https://doi.org/10.1109/ACCESS.2020.3035325>)
- [10] J. Wang and Y. Ge, "A radio frequency energy harvesting-based multihop clustering routing protocol for cognitive radio sensor networks," *IEEE Sensors J.*, vol. 22, no. 7, pp. 7142-7156, Mar. 2022. (<https://doi.org/10.1109/10.1109/JSEN.2022.3156088>)
- [11] K. A. Darabkh, O. M. Amro, R. T. Al-Zubi, and H. B. Salameh, "Yet efficient routing protocols for half-and full-duplex cognitive radio ad-hoc networks over IoT environment," *J. Netw. and Comput. Appl.*, vol. 173, 102836, Jan. 2021. (<https://doi.org/10.1016/j.jnca.2020.102836>)
- [12] M. Zheng, C. Wang, M. Song, W. Liang, and H. Yu, "SACR: A stability-aware cluster-based routing protocol for cognitive radio sensor networks," *IEEE Sensors J.*, vol. 21, no. 15, pp. 17350-17359, Aug. 2021. (<https://doi.org/10.1109/JSEN.2021.3076995>)
- [13] X. He, H. Jiang, Y. Song, C. He, and H. Xiao, "Routing selection with reinforcement learning for energy harvesting multi-hop CRN," *IEEE Access*, vol. 7, pp. 54435-54448, Apr. 2019. (<https://doi.org/10.1109/ACCESS.2019.2912996>)
- [14] T. Stephan, F. Al-Turjman, K. S. Joseph, B. Balusamy, and S. Srivastava, "Artificial intelligence inspired energy and spectrum aware cluster based routing protocol for cognitive radio sensor networks," *J. Parallel and Distrib. Comput.*, vol. 142, pp. 90-105, Aug. 2020. (<https://doi.org/10.1016/j.jpdc.2020.04.007>)
- [15] V. N. J. Robert, and K. Vidya, "Effective cooperative spectrum sensing using deep recurrent reinforced learning based Q-routing in multihop cognitive radio networks," *Int. J. Commun. Syst.*, vol. 34, no. 17, e4982, Sep. 2021. (<https://doi.org/10.1002/dac.4982>)
- [16] A. Paul and S. P. Maity, "Machine learning for spectrum information and routing in multihop green cognitive radio networks," *IEEE Trans. Green Commun. and Netw.*, vol. 6, no. 2, pp. 825-835, Nov. 2021. (<https://doi.org/10.1109/TGCN.2021.3127308>)
- [17] T. S. Malik, K. R. Malik, A. Afzal, M. Ibrar, L. Wang, H. Song, and N. Shah, "RL-IoT:

Reinforcement learning-based routing approach for cognitive radio-enabled IoT communications,” *IEEE Internet of Things J.*, vol. 10, no. 2, pp. 1836-1847, Sep. 2022. (<https://doi.org/10.1109/JIOT.2022.3210703>)

[18] M. Cicioğlu, S. Cicioğlu, and A. Çalhan, “SDN enabled cognitive radio network architecture,” *IET Commun.*, vol. 14, no. 18, pp. 3153-3160, May 2020. (<https://doi.org/10.1049/iet-com.2019.1346>)

[19] N. M. El Azaly, E. F. Badran, H. N. Kheirallah, and H. H. Farag, “Centralized dynamic channel reservation mechanism via SDN for CR networks spectrum allocation,” *IEEE Access*, vol. 8, pp. 192493-192505, Oct. 2020. (<https://doi.org/10.1109/ACCESS.2020.3032666>)

[20] K. Venugopalan and V. Sumathy, “Interference mitigation with spectrum sharing and aggregation for SDN based cognitive networks,” *J. Ambient Intell. and Humanized Computi.*, pp. 1-11, Mar. 2021. (<https://doi.org/10.1007/s12652-021-02996-2>)

[21] L. Abbas, U. Shoaib, and A. K. Bashir, “Priority based dynamic spectrum management using virtual utility functions in cognitive radio enabled internet of things,” *Comput. Commun.*, vol. 196, pp. 239-248, Dec. 2022. (<https://doi.org/10.1016/j.comcom.2022.10.002>)

[22] S. H. A. Kazmi, F. Qamar, R. Hassan, and K. Nisar, “Routing-based interference mitigation in SDN enabled beyond 5G communication networks: A comprehensive survey,” *IEEE Access*, pp. 4023-4041, Jan. 2023. (<https://doi.org/10.1109/ACCESS.2023.3235366>)

[23] Mathworks, “Reinforcement learning toolbox,” Retrived Apr. 24, 2024, from <https://kr.mathworks.com/products/reinforcement-learning.html>

[24] F. Rhamdani, N. A. Suwastika, and M. A. Nugroho, “Equal-cost multipath routing in data center network based on software defined network,” in *2018 6th ICoICT, IEEE*, pp. 222-226, May 2018. (<https://doi.org/10.1109/ICoICT.2018.8528730>)

이철웅 (Cheol-woong Lee)



2015년 2월 : 아주대학교 정보 컴퓨터공학과 졸업
2017년 2월 : 아주대학교 컴퓨터공학과 석사
2023년 8월 : 아주대학교 AI융합네트워크학과 박사
2023년 2월~현재 : LIG 빅스원 재직

<관심분야> 전송네트워크 아키텍처, 지능형 네트워킹, SDN/NFV, 네트워크 M&S

[ORCID:0000-0001-9287-7034]

이호진 (Hojin Lee)



2021년 2월 : 아주대학교 사이버보안학과 졸업
2023년 2월 : 아주대학교 AI융합네트워크학과 석사
2023년 2월~현재 : LIG 빅스원 재직

<관심분야> 통신공학, 네트워크 M&S, SDN

[ORCID:0000-0001-9337-1940]

이규민 (Gyu-min Lee)

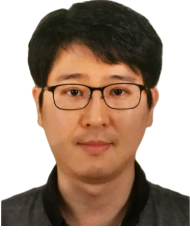


2014년 2월 : 아주대학교 정보 및 컴퓨터공학부 졸업
2016년 2월 : 아주대학교 컴퓨터공학과 석사
2022년 8월 : 아주대학교 컴퓨터공학과 박사
2023년 2월~현재 : LIG 빅스원 신입연구원

<관심분야> C5ISR, 전송네트워크 아키텍처, SDN/NFV

[ORCID:0000-0002-6384-795X]

이 병 규 (Byungkyu Lee)



2005년 8월 : 성균관대학교 기
계공학 학사

2005년 8월 : 성균관대학교 정
보통신공학 학사

2005년 7월~현재 : LIG넥스원
C4I연구소 수석연구원

<관심분야> 군통신시스템, 항
재밍통신기법