

# KREONET-S의 지능적 네트워크 자원 및 전력 관리를 위한 강화학습 알고리즘 적용 및 검증 모델

박성진, 김용환, 김승해, 김동균\*

한국과학기술정보연구원

{xkqpekdl, yh.kim086, shkim, mirr}@kisti.re.kr

## Reinforcement Learning Algorithm for Intelligent Network Resource and Power Management in KREONET-S

Seongjin Park, , Yong-hwan Kim, Seunghae Kim, Dongkyun Kim\*

Korea Institute of Science Technology and Information (KISTI)

### 요약

본 논문은 국가과학기술연구망(KREONET)의 SDN 기반 KREONET-S에서의 지능적 네트워크 자원 및 전력관리를 위한 강화학습 알고리즘 및 검증 모델을 제시한다. 강화학습 기반 지능적 전력관리 모델의 목적은 사용자들의 실시간 네트워크 자원 요청에 대한 요구조건을 만족시킴과 동시에 네트워크 자원 할당으로 인한 전력 소모를 최소화하는 것이다. 이를 위해, 해결하고자 하는 문제를 MDP(Markov Decision Process)로 정의하고 컨트롤러로부터 수집한 실시간 데이터를 활용한 KREONET-S 전력 플랫폼 구축과 시뮬레이션 구축 환경을 소개한다. 최종적으로 제안 방법의 시험 및 검증의 단계별 모델을 제시한다.

### I. 서론

본 논문에서는 국내의 과학기술연구망(KREONET)의 SDN 기반 네트워크 환경인 KREONET-S에서 인공지능 기반 전력 관리 모델을 소개한다. KREONET-S에서는 사용자에게 지정된 네트워크 자원을 할당할 수 있는 응용 서비스인 VDN(Virtual Dedicated Network)을 제공하고 있다. VDN은 KREONET-S에서 사용자가 요청하는 품질이 보장된 네트워크 자원을 동적으로 할당해 줄 수 있는 네트워크 슬라이싱 서비스이다. 이때, 사용자의 요청을 만족시키지 못하면 SLA 위반이 발생할 수 있다. 물리적으로 한정된 네트워크 자원을 가지고 있는 환경에서 사용자의 서비스 요청을 최대한 만족시키면서도 그로 인한 전력 소모량은 최소화 시킬 수 있어야 한다[1]. 따라서 본고에서 제안하는 인공지능 기반 전력 관리 솔루션의 목적은 KREONET-S의 전체 에너지 소모량을 최소화하는 동시에 사용자의 VDN 서비스 요청을 만족시키는 것이다. 이를 위해, KREONET-S의 토폴로지, H/W 구성, 전력소모 현황 데이터를 활용한 강화학습 기반 지능적 전력 관리 솔루션을 제안한다.

### II. 본론

본 논문에서는 KREONET-S의 지능적 전력관리를 위한 해결책으로 강화학습 기반의 알고리즘을 적용하고자 한다. 이를 위해 우선 해결하고자 하는 문제를 MDP(Markov Decision Process)로 정의한다. MDP는 상태 공간, 행동 공간, 상태 전이 확률 함수, 보상 함수의 총 4가지 튜플로 구성된다. 이 때, KREONET-S의 지능적 전력관리와 같이 대부분의 실제 환경에서는 상태 전이 확률이 사전에 주어지지 않는 Model-free 기반의 강화학습 알고리즘 적용이 필요하다. 상태공간은 해결하고자 하는 문제의 현재 환경 정보를 의미하는 것으로 본 논문에서는 KREONET-S의 실시간 운영 정보를 의미한다. 구체적으로는 KREONET-S의 네트워크 장비 전력 소모 현황, VDN 운영 현황(VDN별 사용자 그룹, 할당 대역폭, 트래픽 사용량 등), SDN 컨트롤러 수집 정보(플로우 정보)를 포함한다. 행동

공간은 agent가 실시간으로 결정할 수 있는 범위를 정의하는 것이다. 본 논문에서 agent가 결정해야 할 실시간 의사결정은 각 스위치의 port speed를 조절하는 것이므로 행동공간의 크기는 port configuration의 가능한 모든 경우의 수이다. 즉시 보상함수에는 전력 소모 감소량을 이득으로 VDN 서비스 SLA 위반이 있다면 패널티로 정의한다.

본 논문에서 제안하는 강화학습 기반 지능적 전력관리 알고리즘은 그림 1과 같다. 강화학습 알고리즘에서 Agent는 주어진 환경으로부터 현재 상태 정보를 얻으며 현재 정책에 따른 의사결정을 내린다. 현재 상태에서 정책에 따라 선택된 현재 행동에 대한 결과로 보상 값이 주어진다. 현재 상태 정보, 현재 행동, 보상 값은 현재 환경에 반영되며 다음 상태로 넘어 가게 된다. 이러한 일련의 과정들은 누적 보상 값이 최대화가 되기 위하여 반복적으로 수행하게 되며, 미리 정의된 종료 조건을 충족시킬 만큼 더 좋은 정책을 얻을 때까지 학습을 수행하게 된다.

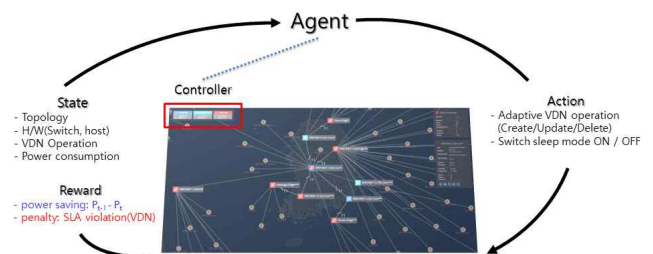


그림 1. KREONET-S의 강화학습 기반 지능적 전력관리 알고리즘

본 논문의 강화학습 기반의 제안 알고리즘의 실험 및 성능 검증을 위한 모델은 다음과 같다. 우선, KREONET-S의 토폴로지 및 SDN 스위치 및 VM 구성 현황을 반영한 가상의 시뮬레이션 환경을 구축하는 것이 필요하다. 이 때, 현재 VDN 서비스 현황을 반영하되 시뮬레이션 환경에서 임의로 생성/삭제/갱신이 가능하도록 구현한다. KREONET-S에서 실제

할당되어 운영되고 있는 VDN list는 그림 2와 같이 현재 14개이지만, 시뮬레이션 상에서는 보다 다양한 시나리오를 가정하기 위해서 변경이 가능하도록 설계해야 한다.

그림 2. KREONET-S의 VDN list 할당 현황

KREONET-S의 전력 현황은 실제 계측 값을 사용하는 대신에 SDN 스위치 별 할당된 포트의 구성을 반영하여 계산된 값을 활용한다. 수식 (1)은 SDN 스위치의 전체 에너지 소모량을 크게 4가지로 분류한 것으로 각 각 스위치 기본 전력 소모, 포트 설정, 컨트롤 패킷 처리, 오픈플로우 트래픽 처리에 대한 전력 소모이다. SDN 전력 소모에 가장 큰 비중을 차지하는 부분은 기본 전력과 포트 설정에 관한 부분이다[2]. 즉, 스위치의 Sleep mode를 ON/OFF 하는 것과 스위치의 포트별 할당된 전송 속도를 결정하는 것이 전력 소모 절감에 있어 가장 중요하다. 따라서 강화학습 agent가 실시간으로 결정해야 할 가장 중요한 의사결정은 스위치 별 포트 설정과 Sleep mode on/off에 대한 것이다.

$$P_{switch} = P_{base} + P_{config} + P_{control} + P_{OF} \quad (1)$$

이 때, KREONET-S의 소모 전력 값은 선행연구들과 KREONET-S에 구축된 스위치들의 공식 specification을 고려하였으며, 각 스위치들의 가능한 포트 설정을 반영하였다. KREONET-S의 가상환경 시뮬레이션 구축 및 인공지능 학습을 위한 데이터는 KREONET-S의 자체 구축한 데이터 수집 및 모니터링 플랫폼으로부터 추출한다.

KREONET-S의 지능적 전력관리를 위한 강화학습 알고리즘의 시험 및 검증 모델은 크게 두 단계로 나눌 수 있다. 첫 번째 단계에서는 비교적 간단한 서비스 시나리오 환경에서의 시험 및 검증을 수행한다. 상태 공간과 행동 공간의 크기가 이산적인 값으로 표현된다면 Q-learning과 같이 기초적인 강화학습 알고리즘의 적용이 가능하다. KREONET-S의 지능적 전력관리 문제에서는 상태공간으로서 실시간 네트워크 장비와 전력소모 현황, VDN 운영에 대한 세부 정보들을 포함한다. 이 때, 특정 파라미터들은 연속적인 실수 값의 범위를 가지는 경우도 있다. 이러한 상황에서 특정 상황별 행동에 따른 가치 값을 Q-table 형식으로 저장하는 Q-learning 방식은 분명한 한계점을 가지고 있다. 이에 대한 대안으로 신경망을 적용함으로써 상태공간을 추상화할 수 있는 DQN(Deep Q Network)과 같은 알고리즘을 적용하는 것이 필요하다.

이에 첫 번째 단계에서는 그림 3과 같이 MDP 문제를 정의한다. 상태 공간은 신경망을 KREONET-S 전력 소모 현황과 VDN 서비스 운영 현황을 반영한다. 행동 공간은 스위치 별 포트 구성 조절과 SDN switch의 sleep mode를 결정하는 것이다. 이 때, 행동 공간의 크기는 조절 가능한 포트 구성의 경우의 수와 스위치의 개수에 의해 결정되므로 이산적인 값이므로 DQN 알고리즘을 적용하기에 적합하다. 즉시 보상함수로는 이전 time slot 대비 절감된 전력 소모량으로 정의한다.

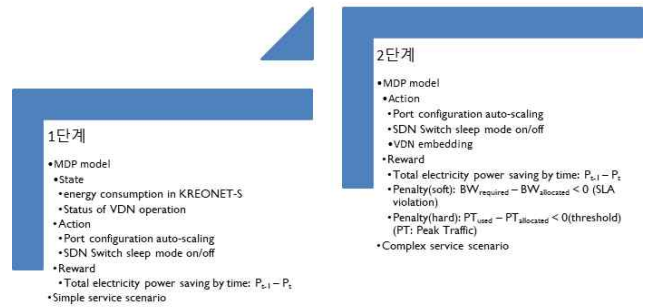


그림 3. 인공지능 기반 전력 관리 모델 단계별 시험 및 검증 방안

두 번째 단계는 보다 고도화된 시험 및 검증을 진행한다. 따라서 실제와 유사한 환경을 반영하기 위하여 복잡하고 어려운 문제를 해결하기에 적합한 알고리즘을 설계 및 구현한다. 먼저, VDN 서비스 생성/삭제/갱신에 대한 관리와 함께 할당된 VDN 자원에 적합한 트래픽을 생성시키는 것이 필요하다. 이 때, 트래픽의 시간별 발생 분포는 요청 VDN의 특성을 고려하도록 한다. 두 번째 단계에서의 MDP 모델은 그림 3과 같이 첫 번째 단계보다 복잡한 환경을 반영한다. 행동 공간의 경우, SDN 스위치의 포트 구성을 조절하는 것 이외에도 실시간 사용자의 요청에 따른 VDN embedding의 세부 과정을 조절한다. 이 경우에는 확률 값과 같이 실수의 값을 결정해야 하는 경우가 발생할 수 있다. 따라서 DDPG(Deep Deterministic Policy Gradient)나 PPO(Proximal Policy Optimization), TRPO(Trust Region Policy Optimization)와 같은 상황별 다양한 강화학습 알고리즘을 적용할 수 있다. 또한, 즉시 보상함수 정의할 때도 실제 사용자 요청을 위반했을 때 뿐만 아니라 실제 사용 트래픽 대비 할당된 트래픽이 낮은 경우를 고려하여 차등된 SLA(Service Level Agreement)를 부여하도록 한다.

### III. 결론

본 논문에서는 KREONET-S에서 수집하고 있는 운영 데이터를 기반으로 지능적 전력 관리에 대한 요구사항을 분석하고 환경에 적응적인 실시간 의사결정이 가능한 강화학습 기법을 통한 지능적 네트워크 자원 및 전력 관리 모델을 설계하고 그에 대한 시험 및 검증 방안을 제시하였다. 제안 모델의 시험 및 검증 이후에는 KREONET-S의 지능적 전력관리 연구의 고도화를 위해 5G 등 무선 네트워크 서비스 환경 반영 및 실제 전력 계측을 통한 모델 개선 등을 추진할 예정이다.

### ACKNOWLEDGMENT

본 논문은 2023년 한국과학기술정보연구원(KISTI) 주요사업 과제로 수행한 연구결과입니다.

### 참 고 문 헌

- [1] B. G. Assefa, Ö. Özkasap, "survey of energy efficiency in SDN: Software-based methods and optimization models", Journal of Network and Computer Applications, vol. 137, pp. 127 - 143, Apr. 2019.
- [2] F. Kaup, S. Melnikowitsch, D. Hausheer, "Measuring and Modeling the Power Consumption of OpenFlow Switches", 10th International Conference on Network and Service Management (CNSM), pp. 181-186, 2014.