

에지 클라우드 네트워크에서 계층적 심층강화학습을 활용한 부분 작업 오프로딩 최적화

석영준¹, 임현교¹, 올라 이산², 한연희^{1*}

¹한국기술교육대학교 컴퓨터공학과 미래융합공학전공

²한국기술교육대학교 첨단기술연구소

dsb04163@koreatech.ac.kr, glenn89@koreatech.ac.kr,

ihsan@koreatech.ac.kr, yghan@koreatech.ac.kr,

Hierarchical Deep Reinforcement Learning-Based Partial Task Offloading Optimization in Edge Cloud Networks

Yeong-Jun Seok¹, Hyun-Kyo Lim¹, Ihsan Ullah², Youn-Hee Han^{1*}

¹Future Convergence Engineering, Dept. of Computer Science and Engineering,
KOREATECH

²Advanced Technology Research Center, KOREATECH

요약

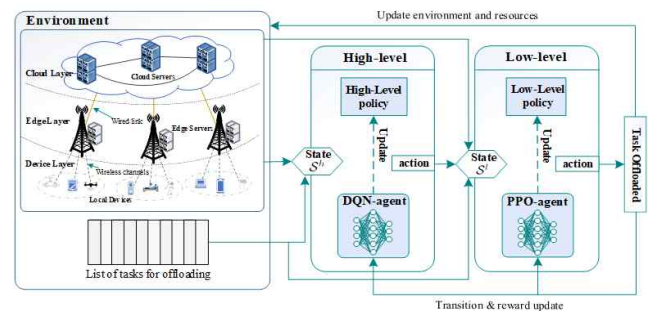
IoT 및 5G 네트워크에서 부분 작업 오프로딩(Partial Task Off-loading)은 자원이 제한된 로컬 장치에서 요구되는 에너지 및 지연시간을 만족시키기 위한 중요한 솔루션 중 하나이다. 하지만 부분 작업 오프로딩에서 오프로딩 비율 선택에 조합 최적화 문제가 발생한다. 본 논문에서는 이러한 문제 해결을 위해 에지 클라우드 컴퓨팅 체계에서 계층적 심층 강화학습(Hierarchical Deep Reinforcement Learning)을 적용한 부분 작업 오프로딩을 제안한다. 또한 계층적 심층 강화학습 기반 작업 분배 최적화 문제에 대한 Markov Decision Process (MDP) 및 목적함수를 정의한다. 계층적 심층 강화학습에서의 각 계층은 서로 다른 알고리즘을 활용하는 것을 장려하며 작업 분배 문제에 대한 시뮬레이션 결과는 제안하는 방법이 다른 방법 보다 지연시간 및 자원소모 측면에서 우수하다는 것을 보여준다.

I. 서론

IoT 및 5G 무선 네트워크는 증강/가상 현실과 같은 새로운 애플리케이션의 사용을 가능하게 했다. 하지만, 자원이 제한된 로컬 장치가 지연에 민감하고 계산 집약적인 애플리케이션을 지원하기는 어렵다[1][2]. 따라서 에지 클라우드 컴퓨팅 체계에 작업을 오프로딩하여 로컬 장치에서 이러한 애플리케이션을 지원하도록 할 수 있다[3].

계층적 심층 강화학습(Hierarchical Deep Reinforcement Learning)은 복잡한 문제를 상위 레벨과 하위 레벨로 분할하여 정복하는 접근법을 제안함으로써 강화학습을 확장한다[4][5]. 이 접근법에서, 문제의 복잡성은 여러 개의 작은 문제로 분할되며 이러한 문제들은 일반적으로 해결하기가 더 쉽고, 그 해결책들은 다른 문제들을 해결하기 위해 재사용될 수 있다.

부분 작업(Partial Task Off-loading) 오프로딩 비율에 따라 지연시간 및 에너지 소모가 다양해진다. 따라서 최적의 오프로딩 비율을 선택해야 하는 조합 최적화 문제가 발생한다. 본 논문에서는 위 문제를 해결하기 위해 에지 클라우드 컴퓨팅 체계에서 계층적 심층강화학습 기반 부분 작업 오프로딩을 제안하고, 다양한 오프로딩 방법과 비교하여 에너지 소모 최소화 및 지연시간 만족 측면에서 우수하다는 것을 증명한다.



[그림 1] 계층적 강화학습 방식의 에지 클라우드 체계

II. 시스템 모델 및 문제 정의

본 장에서는 에지 클라우드 컴퓨팅 체계의 시스템 모델 및 문제 정의를 설명한다. 그림 1은 에지 클라우드 체계에서 작업 오프로딩을 위해 제안된 계층적 심층 강화학습 방식을 보여준다. 시스템 모델은 작업 모델, 에너지 모델, 지연시간 모델로 구성된다.

1. 작업 모델(Task Model)

제안하는 모델에서 Base Station은 M 개 존재하며 각 Base Station에 N 개의 장치가 있다. 각 장치에는 처리할 작업이 하나 존재하며 작업은 데이터의 사이즈 σ (MB) 및 지연시간 λ (s)으로 구성되어 있다. 부분 작업 오프로딩을 통한 분할 비율은 다음과 같이 표시된다. $V = \{x, y, z\}$. 순서대로 로컬, 에지, 클라우드 별 작업의 오프로딩

* 한연희: 교신저자

비율을 나타낸다. 오프로딩 비율의 총합은 1로 제한된다.

2. 에너지 모델(Energy Model)

작업 당 소요되는 에너지 E 는 전송 에너지 E^{tr} 와 처리 에너지 E^{ex} 의 합이며 정규화 되어 0에서 1사이의 값을 가진다. 처리 에너지 E^{ex} 는 다음과 같이 계산된다. $E^{ex} = P^{ex} \sigma V$. P^{ex} (J/bit)는 계층의 컴퓨팅 에너지를 나타내며 V 는 계층의 작업 비율을 나타낸다. 전송 에너지 E^{tr} 는 다음과 같이 계산된다. $E^{tr} = P^{tr}(\sigma V + \alpha)$. P^{tr} (J/bit)는 전송 컴퓨팅 에너지를 나타내며 α 는 처리 후 반환되는 데이터의 크기이다.

3. 지연시간 모델(Delay Model)

지연 시간 T 는 전송 지연 시간 T^{tr} 과 처리 지연 시간 T^{ex} 의 합이다. 작업 당 전송 지연시간 T^{tr} 은 다음과 같이 계산된다. $T^{tr} = \sigma V / \mu$. μ (Gbps)는 오프로딩 계층의 전송률이다. 작업 당 계층에서 처리 지연시간 T^{ex} 은 다음과 같이 계산된다. $T^{ex} = \sigma V / F$. F (CPU cycle/s)는 계층의 컴퓨팅 사이클을 나타낸다. 지연시간은 각 계층의 지연시간 중 가장 큰 값을 사용한다.

4. 문제 정의

작업을 분할하여 오프로딩 한다면 분할 비율 $V = \{x, y, z\}$ 에 따라 작업 처리에 소모되는 에너지 및 지연시간이 다양해진다. 모든 Base Station M 의 모든 작업 N 을 효율적으로 오프로딩하여 소모되는 에너지를 최소화하고 지연시간을 만족하기 위해서 본 논문에서 해결하고자 하는 문제를 다음과 같이 정의한다.

$$\min \sum_{m=1}^M \sum_{n=1}^N (E_n^m) \text{ s.t. } T_n^m \leq \lambda_n^m$$

III. 계층적 심층 강화 학습

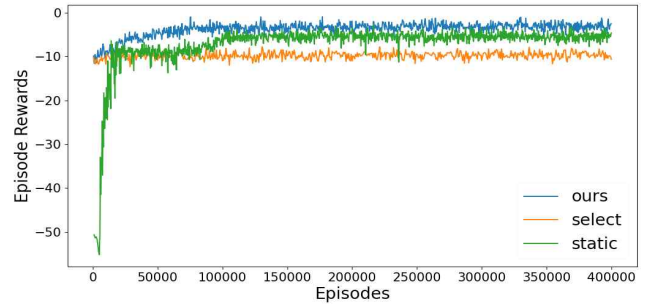
본 장에서는 시스템 모델에서 작업 오프로딩을 위한 계층적 심층 강화학습의 상위 레벨 MDP(Markov Decision Process)와 하위 레벨 MDP를 설명한다.

1. 상위 레벨 MDP(High-Level MDP)

상위 레벨 에이전트는 상태로 네트워크 정보를 입력받아 행동으로 작업의 분할 오프로딩 위치를 출력한다. 행동은 [로컬, 에지, 클라우드] 같은 형태를 가진다. 예시로 로컬, 에지, 클라우드에 작업을 분할해서 할당한다면 행동의 형태는 [1, 1, 1] 같으며, 로컬 및 에지에만 작업을 분할해서 할당한다면 행동의 형태는 [1, 1, 0]과 같다. 이러한 상위 레벨의 행동 형태는 하위 레벨의 상태로 전송된다. 보상은 하위 레벨 에이전트의 행동 후 계산된다.

2. 하위 레벨 MDP(Low-Level MDP)

하위 레벨 에이전트는 상태로 네트워크 정보 및 상위 레벨의 행동을 입력받아 행동으로 오프로딩 비율을 출력한다. 행동의 형태는 상위 레벨과 같으며 예시는 [0.4, 0.4, 0.2] 같다. 보상은 에너지 모델을 통해 계산된 작업의 컴퓨팅 에너지와 지연시간 모델을 통해 계산된 작업의 지연시간을 통해 계산된다. 리워드 R 은 $T < \lambda$ 라면 $1 - E$, $T > \lambda$ 라면 $-E$ 이다. 계산된 보상은 상위 레벨에도 전송된다.



[그림 2] 다양한 알고리즘의 에피소드 보상 비교 결과

IV. 실험 및 결과

비교 실험에서 계층적 심층 강화학습 알고리즘은 상위 레벨은 DQN, 하위 레벨은 PPO를 사용했다. static은 심층 강화학습 PPO만을 사용했고 select는 static과 알고리즘은 같지만 작업을 분할하여 오프로딩 하지 않는다. 결과는 3회 학습의 평균을 사용했으며 [그림 2]와 같다. 환경에서 $M=3$ 이며 $N=32$ 이다. X축은 에피소드를 나타내며, Y축은 학습 후 테스트에서 받은 보상을 나타낸다. 제시한 기법이 $T < \lambda$ 를 만족하며 E 는 최소화 하는 것을 알 수 있다.

V. 결론

본 연구에서는 에지 클라우드 네트워크에서 작업의 지연시간은 만족하며 에너지 소모는 최소화하는 계층적 심층 강화학습 방법을 제시하고 효과를 검증했다. 본 연구는 현재 작업의 오프로딩 순서는 선택하지 않는다. 향후 연구에서 오프로딩 순서 또한 강화학습을 통해 결정하는 것으로 더욱 효율적으로 오프로딩 할 수 있도록 기대한다.

ACKNOWLEDGMENT

이 논문은 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (No. 2020R1I1A3065610 & No. 2018R1A6A1A03025526)

참고문헌

- [1] A. Singh, S. C. Satapathy, A. Roy, and A. Gutub, "AI-Based Mobile Edge Computing for IoT: Applications, Challenges, and Future Scope," Arab. J. Sci. Eng., pp. 1 - 31, 2022.
- [2] A. S. Mohammed, K. Venkatachalam, S. Hubálovský, P. Trojovský, and P. Prabu, "Smart Edge Computing for 5 g/6 g Satellite IOT for Reducing Inter Transmission Delay," Mob. Netw. Appl., pp. 1 - 10, 2022.
- [3] B. Dai, J. Niu, T. Ren, and M. Atiquzzaman, "Towards Mobility-Aware Computation Offloading and Resource Allocation in End-Edge-Cloud Orchestrated Computing," IEEE Internet Things J., 2022.
- [4] T. D. Kulkarni, K. Narasimhan, A. Saeedi, and J. Tenenbaum, "Hierarchical Deep Reinforcement Learning: Integrating Temporal Abstraction and Intrinsic Motivation," Advances in neural information processing systems, vol. 29, 2016.
- [5] M. M. Botvinick, "Hierarchical Reinforcement Learning and Decision Making," Current opinion in neurobiology, vol. 22, no. 6, pp. 956 - 962, 2012.