

저탄소 AI 데이터센터 작업 분배를 위한 계층 강화학습 프레임워크

이현수, 최대현*, 김중현

고려대학교, *중앙대학교

hyunsoo@korea.ac.kr, *dhchoi@cau.ac.kr, joongheon@korea.ac.kr

Hierarchical Reinforcement Learning Framework for Low-Carbon AI Data Center Workload Distribution

Hyunsoo Lee, Dae-Hyun Choi*, and Joongheon Kim

Korea Univ., *Chung-Ang Univ.

요약

AI 애플리케이션의 급격한 확장에 따른 에너지 소비와 탄소 배출 급증으로 인해 AI 데이터센터(AIDC)의 친환경적인 에너지 관리가 중요해지고 있다. 이에, 본 논문은 배전망의 저탄소 운영을 보장하면서 불확실성 하에서도 AIDC를 효율적으로 운영하기 위한 계층적 탄소 인식 다중 에이전트 강화학습(CA-MARL) 프레임워크를 제안하였다. 이 프레임워크는 멀티 에이전트 트랜스포머 방식을 사용하여 학습된 글로벌 워크로드 관리자 에이전트와 다수의 로컬 AIDC 에이전트로 구성된다. 워크로드 관리자 에이전트는 노드별 탄소 강도(NCI)를 기반으로 AI 작업을 공간적으로 할당하며, 각 AIDC 에이전트는 작업의 시간적 이동, GPU 할당 및 냉각 시스템 제어를 수행한다.

I. 서론

최근 거대 언어 모델(Large Language Model, LLM)과 생성형 인공지능(Artificial Intelligence, AI) 기술이 급격하게 확산됨에 따라, 이를 지탱하는 AI 데이터센터(AI Data Center, AIDC)의 전력 소비량이 폭발적으로 증가하고 있다. 2024년 기준 AIDC는 전 세계 전력 사용량의 약 1.5%를 차지하고 있으며, 2030년에는 이 수치가 두 배로 늘어날 것으로 전망된다 [1]. 이러한 에너지 소비 급증은 막대한 탄소 배출을 유발할 뿐만 아니라 전력 배전망의 운영 불안정성을 초래할 위험이 있다.

기존의 인터넷 데이터센터(Internet Data Center, IDC)의 경우 전력 최적화를 위해 배전망과의 상호 작용을 통해 비용을 최소화하거나 강건성을 향상시키는 연구가 수행되었다. 예를 들어, IDC의 워크로드를 공간적으로 스케줄링하여 배전망 운영 비용을 줄이거나 [2], 노드별 탄소 강도(Nodal Carbon Intensity, NCI) 정보를 이용하여 워크로드를 시간적으로 이동시켜 탄소 배출을 저감하려는 연구도 수행되었다 [3]. 그러나 이러한 기존 연구들은 학습(Training) 및 추론(Inference) 워크로드가 혼재된 AIDC의 복잡한 운영 특성을 반영하지 못하였다. AIDC는 지역에 민감한 추론 작업과 지역 허용이 가능한 학습 작업의 서로 다른 특성의 워크로드를 처리한다는 점에서 차별화된다 [4]. 추론 작업은 개별 작업의 크기가 학습 작업에 비해 작지만, 사용자의 요청에 따른 즉각적인 응답이 필요하여 지역적으로 분산된 데이터센터 간의 공간적 배분(Spatial Allocation)이 중요하다. 반면에 학습 작업은 개별 실행 시간이 수 시간에서 수 주에 걸쳐 수행되기도 하는 큰 작업이지만, 마감 기한 내에서 실행 시점을 조절할 수 있는 시간적 이동(Temporal Shifting)이 가능하다 [5]. 본 연구는 이러한 AI 워크로드의 특성을 고려하여, 계층적 강화학습 방식을 통해 전력 최적화와 저탄소 운영을 동시에 달성할 수 있는 새로운 프레임워크를 제안한다.

II. 본론

II-1. 시스템 모델 및 환경

본 프레임워크의 환경 모델은 배전망의 물리적 특성을 반영하기 위해 탄소 배출 흐름(Carbon Emission Flow, CEF) 모델을 통합하였다. 이는 단순히 발전원의 위치만을 고려하는 기존 방식과 달리, 전력 조류 방정식에

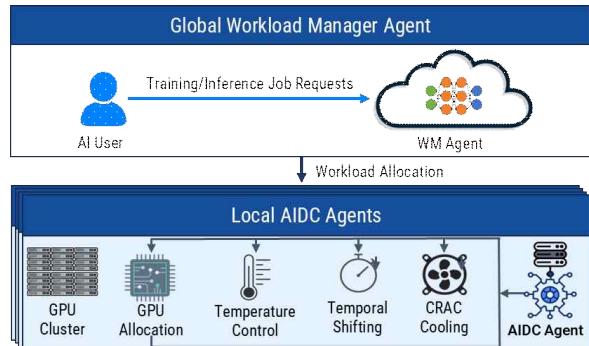


그림 1 데이터센터 작업 분배를 위한 계층 강화학습 프레임워크.

따라 탄소 배출이 발전원에서 소비 노드로 흘러가는 가상의 경로를 추적 한다. 이를 통해 산출된 NCI는 각 AIDC가 소비하는 전력에 포함된 탄소 강도를 수치화하여 에이전트에게 전달된다.

본 논문에서 상정하는 시스템은 크게 사용자로부터 훈련 워크로드와 추론 워크로드 요청을 받는 작업 관리자 에이전트와, 이를 처리하는 AIDC 모델로 구성된다. 작업 관리자 측면에서는 워크로드를 크게 학습과 추론으로 분류한다. 학습 작업은 대규모 연산이 필요하지만 장기간의 지연을 허용하므로, 전력 가격이나 탄소 집약도가 높은 시간대를 피하여 작업을 미루는 시간적 유연성을 가진다. 반면, 추론 작업은 사용자 요청에 즉각 반응해야 하므로 대기열 관리가 중요하며, 주로 데이터센터 간의 부하 분산을 통해 최적화를 도모한다. 또한, AIDC의 전력 소비는 서버의 IT 전력 뿐만 아니라 냉각 시스템(Computer Room Air Conditioning, CRAC)의 전력 소모까지 포함하여 모델링되었으며, 서버의 부하 상태에 따라 냉각 공급 온도를 제어하는 열역학적 모델도 함께 고려하였다.

II-2. 학습 프레임워크

이러한 복잡한 환경에서 최적의 의사결정을 내리기 위해, 본 연구에서는 그림 1에서 나타낸 것과 같은 ‘계층적 다중 에이전트 강화학습(Hierarchical Multi-Agent Reinforcement Learning)’ 프레임워크를 제안한다. 이 프레임워크는 방대한 행동 공간으로 인한 학습의 어려움을 해결하고 효율적인 협력을 유도하기 위해, 전체 시스템을 ‘글로벌 작업 관리

자 에이전트'와 '로컬 AIDC 에이전트'라는 두 개의 계층으로 나누어 설계하였다.

상위 계층인 작업 관리자 에이전트는 AIDC 통합 관리자 역할을 수행한다. 작업 관리자 에이전트는 사용자의 작업 요청과 각 AIDC의 운영 상태를 관측한 뒤, 도착하는 대규모 학습 및 추론 작업을 지리적으로 분산된 여러 AIDC에 최적 비율로 할당한다. 즉, 탄소 집약도가 낮거나 여유 자원이 있는 데이터센터로 작업을 보내는 '공간적 할당'을 전달하며, 이때 '최대 나머지 규칙(Largest Remainder Rule)'을 적용하여 연속적인 확률값을 실제 할당 가능한 정수 단위의 작업 개수로 변환한다. 하위 계층인 로컬 AIDC 에이전트는 개별 데이터센터 운영자 역할을 맡는다. 각 AIDC 에이전트는 작업 관리자 에이전트로부터 할당받은 작업량을 바탕으로 구체적인 실행 계획을 수립한다. 구체적으로는 학습 작업의 실행 시점을 조절하여 탄소 배출이 적은 시간대로 미루는 '시간적 이동(Task 1)', 주어진 작업을 처리하기 위해 GPU 자원을 효율적으로 분배하는 '자원 할당(Task 2, 3)', 그리고 서버 온도에 맞춰 냉각 시스템에 공급되는 공기의 온도를 조절하는 '냉각 제어(Task 4)'를 수행한다.

작업 관리자와 AIDC 에이전트는 협력을 통해 하나의 통합된 보상 함수를 최대화하도록 학습한다. 이 보상 함수는 총 작업 처리량(Throughput)에 대한 보상과 전력 구매 비용, 총 탄소 배출량, 그리고 마감 기한을 넘겨 수행에 실패한 작업 수에 대한 폐널티의 가중 합으로 구성된다. 이를 통해 에이전트는 경제성과 탄소 배출량, 그리고 서비스 품질(QoS) 사이의 균형을 스스로 학습하게 된다.

이 두 계층의 에이전트는 멀티 에이전트 트랜스포머(Multi-Agent Transformer, MAT) 구조를 통해 유기적으로 연결된다. MAT는 '다중 에이전트 이점 분해(Multi-Agent Advantage Decomposition)'에 기반하여 에이전트 간의 협력을 모델링하는데, 구체적으로 인코더-디코더 구조와 순차적인 의사결정 방식(Auto-regressive)을 사용하였다 [6]. 즉, 작업 관리자 에이전트가 먼저 공간적 배치를 결정하면, 이 정보가 인코딩되어 각 AIDC 에이전트에게 전달되고, AIDC 에이전트는 이를 조건부로 하여 자신의 로컬 스케줄링을 결정한다. 이러한 구조는 에이전트 간의 불필요한 충돌을 방지하고, 기하급수적으로 늘어날 수 있는 조합 탐색 공간을 효과적으로 줄여준다.

II-3. 성능 평가

본 연구에서 제안한 계층적 CA-MARL 프레임워크의 유효성을 검증하기 위해, 3개의 AIDC가 포함된 IEEE 33-노드 배전망 환경에서 시뮬레이션을 수행하였다. 특히, 제안된 기법이 NCI의 변동에 따라 워크로드를 얼마나 효과적으로 공간 배분(Spatial Allocation)하는지 분석하기 위해, 서로 다른 NCI 값을 갖는 3개의 AIDC에 대한 성능 지표를 비교하였다. 표 1은 하루 동안의 NCI 변동 평균값과 그에 따른 작업 할당, 총 전력 소비, 전력 효율 지수(Power Usage Effectiveness, PUE)를 나타낸 것이다. 작업 관리자(WM) 에이전트는 탄소 배출을 최소화하기 위해 NCI 값과 반비례하여 작업을 할당하는 경향을 보였다. 표 1에서 확인할 수 있듯이, NCI가 0.420으로 가장 낮은 AIDC 1에는 전체의 약 43%에 해당하는 76,596개의 작업이 집중적으로 할당되었다. 반면, NCI가 0.841로 가장 높은 AIDC 3에는 상대적으로 적은 37,586개의 작업만이 할당되었다. 이는 에이전트가 NCI가 높은 AIDC의 부하를 줄임으로써 시스템 전체의 탄소 발자국을 저감하려는 의도된 행동을 학습했음을 시사한다.

또한, 이러한 작업 할당은 AIDC의 IT 전력과 냉각 시스템 전력의 비율인 PUE에도 영향을 미쳤다. 많은 작업이 할당되어 IT 전력 소모가 커던 AIDC 1은 냉각 시스템이 높은 가동률로 효율적으로 운영되면서 1.05라는

구분	NCI (kgCO ₂ /kWh)	할당된 작업 수	총 소비전력 (MW)	전력 효율 지수
AIDC 1	0.420	76,596	11.86	1.05
AIDC 2	0.569	62,956	10.65	1.06
AIDC 3	0.841	37,586	8.44	1.08

표 1 NCI에 따른 AIDC별 할당된 작업 수와 총 전력, 전력 효율 지수.

가장 우수한 PUE 수치를 기록하였다. 이와 대조적으로 작업 할당량이 적은 AIDC 3은 1.08의 상대적으로 높은 PUE 수치를 보였다. 이는 탄소 저감을 위한 부하 분산이 AIDC의 전력 효율성과 밀접한 Trade-off를 가지며, 제안된 프레임워크가 복잡한 상관관계 속에서도 NCI를 고려한 최적의 운영 지점을 찾아낼 수 있음을 보인다.

III. 결 론

본 연구에서 제안된 계층적 MARL 프레임워크는 글로벌 작업 관리자 에이전트와 로컬 AIDC 에이전트의 역할을 명확히 분리함으로써, 불확실한 워크로드 환경에서도 안정적이고 효율적인 운영을 가능하게 하였다.

결과적으로 이 프레임워크는 AI 워크로드의 시공간적 유연성을 극대화하여 데이터센터의 경제성을 확보함과 동시에 시스템 전체의 탄소 배출을 저감하면서도 원활하게 AI 워크로드를 처리하는 데 기여할 수 있다. 특히, 계층적 접근 방식은 대규모 시스템에서의 복잡한 상호작용 문제를 해결하고, 데이터센터 운영과 전력망 관리의 공존을 모색하는 새로운 방향성을 제시하고 있다. 향후 연구로는, 이기종 전력망과의 결합을 통해 서로 다른 전력 공급원이 있는 상황에서의 전력 최적화 및 워크로드 분배 연구가 가능할 것으로 전망된다.

ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 지원으로 정보통신기획평가원-대학HCT 연구센터(ITRC)의 지원을 받아 수행된 연구임 (IITP-2026-RS-2024-00436887). 본 논문의 교신저자는 최대현, 김중현입니다.

참 고 문 헌 [1] IEA, "Tracking Clean Energy Progress 2023", IEA, 2023, (<http://www.iea.org/reports/tracking-clean-energy-progress-2023>).

[2] M. Chen, C. Gao, M. Shahidehpour and Z. Li, "Incentive-Compatible Demand Response for Spatially Coupled Internet Data Centers in Electricity Markets," *IEEE Trans. Smart Grid*, vol. 12, no. 4, pp. 3056–3069, Jul. 2021.

[3] P. Wiesner et al., "Let's wait awhile: How temporal workload shifting can reduce carbon emissions in the cloud," in *Proc. International Middleware Conference (Middleware)*, Québec City, Canada, Dec. 2021, p. 260 - 272.

[4] X. Chen, X. Wang, A. Colacelli, M. Lee, and L. Xie, "Electricity demand and grid impacts of AI data centers: Challenges and prospects," *arXiv preprint arXiv:2509.07218v2*, Sep. 2025.

[5] W. Zhang, Z. Zhang, S. Zeadally, H. -C. Chao and V. C. M. Leung, "Energy-efficient Workload Allocation and Computation Resource Configuration in Distributed Cloud/Edge Computing Systems With Stochastic Workloads," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 6, pp. 1118–1132, Jun. 2020.

[6] M. Wen, J. G. Kuba, R. Lin, W. Zhang, Y. Wen, J. Wang, Y. Yang, "Multi-agent reinforcement learning is a sequence modeling problem," in *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, New Orleans, LA, USA, Dec. 2022, pp. 16509 - 16521.