

자연어 기반 디지털 트윈 환경 제어를 위한 멀티에이전트 구조 설계 및 실증

김혜영, 정영환, 최원기, 이진영, 이상신*

한국전자기술연구원

{boramaey, cjstntjd, cwk1412, jylee}@keti.re.kr, *sslee@keti.re.kr

Design and Demonstration of a Multi-Agent Architecture for Natural Language-Based Digital Twin Environment Control

Kim Hye Young, Jeong Young Hwan, Choi Won Gi, Lee Jin Young, Lee Sang Shin*

Korea Electronics Technology Institute

요약

최근 디지털 트윈 환경을 활용하여 복합 시스템의 상태를 분석하고 이를 기반으로 실제 환경을 제어하려는 요구가 증가하고 있다. 이러한 환경에서는 자연어 질의를 통해 센서 데이터를 해석하고 제어 도구 호출을 적절하게 수행할 수 있는 지능형 제어 모델이 요구된다. 기존의 키워드 매핑 및 단일 대규모 언어 모델(LLM) 기반 접근법은 명시적인 요청에는 대응할 수 있으나 질의가 모호하거나 다단계 추론이 필요한 상황에서는 제어의 일관성과 정밀도가 저하되는 한계를 보인다. 이를 해결하기 위해 본 논문에서는 역할 분담 기반의 멀티에이전트 모델을 제안한다. 관리, 평가, 계획, 제어 기능으로 에이전트를 분리하여 구성함으로써 복잡적이고 불완전한 질의에 대해서도 단계적인 판단과 제어 수행이 가능하도록 설계되었다. 실제 돈사 디지털 트윈을 활용하여, 제안된 멀티에이전트 기반 모델이 기존 방식과 비교하여 자연어 질의에 대해 보다 안정적이고 일관된 제어 결정을 수행함을 보여준다.

I. Introduction

최근 디지털 트윈 기술의 발전과 함께 가상 환경에서 시스템의 상태를 정밀하게 모사하고 이를 기반으로 실제 환경을 제어하고자 하는 요구가 다양한 산업 분야에서 증가하고 있다.[1] 그러나 디지털 트윈 기반 제어 시스템은 다수의 센서와 제어 요소가 복잡하게 얽혀 있고 도메인 지식에 대한 이해를 전제로 하므로 주로 전문가를 위한 도구로만 활용되어 왔고, 비전문가가 이를 직관적으로 활용하기에는 진입장벽이 높다는 한계가 있었다.[2] 이러한 문제를 완화하기 위해 자연어 인터페이스 기반의 LLM을 통한 디지털 트윈 방식이 제안되었으나, 기존의 키워드 매핑 방식 또는 단일 LLM 방식은 모호한 자연어 질의나 다단계 추론이 필요한 상황에서 제어의 일관성과 정밀도가 저하되는 한계를 보였다.

본 논문에서는 이러한 문제를 해결하기 위해, 전문가의 디지털 트윈 내 작업 흐름을 모사한 멀티에이전트 구조를 제안한다. 제안 모델은 관리, 평가, 계획, 제어에 해당하는 의사결정을 수행하는 개별 에이전트로 구성되며, 각 에이전트는 상호 협력적으로 동작하면서 복합적인 자연어 질의를 단계적으로 해석하고 제어 전략을 도출한다. 이를 통해 사용자의 질의에 포함된 정보가 불완전하거나 모호한 경우에도 환경 상태를 점진적으로 분석하고 일관된 제어 결정을 수행할 수 있도록 설계되었다.

실제 돈사 디지털 트윈 환경을 실증 테스트베드로 활용한 비교 실험 결과, 제안한 멀티에이전트 기반 모델은 단순 질의와 복합 질의 모두에서 기존 모델 대비 우수한 제어 성능을 보였다. 또한 자연어 질의에 포함된 정보의 명확성 수준과 무관하게 환경 상태와 제어 맥락을 일관되게 반영하고 제어 목적에 부합하는 MCP 도구를 선별하여 호출하는 과정에서도 높은 성공률을 보이는 것을 확인하였다.

II. Proposed Method

멀티에이전트 시스템은 상호작용하는 여러 개의 에이전트로 구성된 시스템으로 하나의 복잡한 문제에 대해 각 에이전트가 부분적인 의사결정을 수행하고, 이들의 협력을 통해 전체 문제를 해결하는 인공지능 구조이다. 이러한 구조는 단일 모델 기반 접근 방식에 비해 복잡한 판단 과정과 단계적 추론이 요구되는 환경에서 더욱 안정적인 의사결정을 가능하게 한다.[3] 그림 1은 제안한 멀티에이전트 구조와 디지털 트윈 및 실제 환경 간의 연계 구조를 나타낸다. 제안 모델은 관리(Supervisor), 평가(Critic), 계획(Planner), 제어(Controller)에 해당하는 역할을 담당하는 개별 에이전트로 구성되며, 각 에이전트가 상호 협력하여 문제를 반복적으로 추론하는 구조를 가진다.

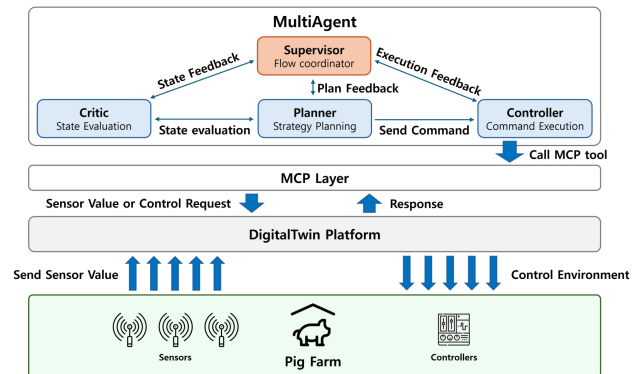


그림 1. 멀티에이전트-MCP-디지털 트윈 플랫폼 구조도

구체적으로, Supervisor 에이전트는 전체 질의 처리 흐름을 관리하며, 각 에이전트 간의 상호작용을 조정한다. Critic 에이전트는 센서 데이터 및 환경 상태를 기반으로 현재 상황을 평가하고, Planner 에이전트는 평가

결과를 토대로 제어 전략을 수립한다. Controller 에이전트는 Planner로부터 전달된 제어 계획을 실제 실행 가능한 명령으로 변환하여 MCP(Model Control Protocol) Layer를 통해 디지털 트윈 플랫폼에 전달한다. 각 에이전트는 독립적인 언어 모델을 기반으로 구성되었으며, 역할의 복잡도에 따라 최적화된 크기의 언어 모델을 배치하였다. 전체 맥락 파악이 중요한 Supervisor에는 상대적으로 큰 모델을, 실행 효율성이 중요한 하위 에이전트에는 경량 모델을 배치하여 시스템 전체의 응답 속도와 추론 정확도 사이의 균형을 맞추었다. 이를 통해 단일 모델이 모든 판단을 수행하는 구조에서 발생할 수 있는 추론 부담을 분산시키고, 질의 해석 및 제어 의사결정의 안정성을 확보하고자 하였다. 표 1은 멀티에이전트 구성에 사용된 모델과 각 역할 간의 대응 관계를 정리한 것이다.

모델	크기	역할
Deepseek-r1	8b	Supervisor
Gemma 3	4b	Planner
Gemma 3	1b	Critic
Gemma 3	1b	Controller

표 1. 멀티에이전트 역할별 모델 구성

본 연구는 실제 돈사의 센서 데이터가 실시간으로 수집되는 디지털 트윈 플랫폼을 실증 테스트 베드로 활용하였다. 그림 2는 채팅 인터페이스를 통한 디지털 트윈의 실시간 제어 화면으로, 사용자의 추상적 자연어 질의에 대해 데이터 조회, 환경 상태 분석, 제어 판단, 명령 실행으로 이어지는 추론 과정이 논리 흐름에 맞게 순차적으로 수행되며 그 결과가 디지털 트윈에 적용되어 실제 환경을 제어하는 과정을 보여준다.



그림 2. 돈사 디지털 트윈 실제 적용 화면

III. Simulation Result

이 절에서는 멀티에이전트 시스템의 성능을 평가한다. 자연어 질의에 대한 제어 성공 여부는 제어 명령의 정상적인 실행과 MCP 도구 호출의 적합성 여부를 기준으로 정의하였다. 단순 질의와 복합 질의를 구분하여 키워드 매핑 기반, 단일 LLM 기반, 멀티에이전트 기반 모델의 제어 수행 결과를 비교하고 그 결과를 분석하였다.

표 2는 질의 유형별 제어 성능을 비교한 결과를 나타낸다. 키워드 매핑 기반 모델은 사전에 정의된 키워드와 정확히 일치하는 경우에 한해 제한적인 성능을 보였으며, 표현 방식이 달라지거나 질의가 복잡한 경우에는 대부분의 질의에서 제어 수행에 실패하였다. 단일 LLM 기반 모델은 키워드 매핑 방식에 비해 질의 대응 범위는 확장되었으나, 복합 질의 상황에서는 환경 상태 해석과 제어 판단의 일관성이 저하되는 경향을 보였다. 반면 멀티에이전트 기반 모델은 단순 질의와 복합 질의 모두에서 상대적으로 높은 성공률을 보였으며, 다양한 질의 유형에 대해 일관된 응답을 생성하는 것을 확인하였다.

모델	단순 질의	복합 질의
키워드 매핑	40%	5%
단일 LLM	50%	20%
멀티에이전트	95%	85%

표 2. 질의 유형별 제어 성능 비교

표 3은 각 모델의 MCP 도구 호출 과정에서의 세부 수행 결과를 비교한 것이다. 키워드 매핑 기반 모델은 사전에 정의된 규칙에 포함되지 않는 경우 적절한 도구를 선택하지 못해 대부분의 단계에서 실패하는 양상을 보였다. 단일 LLM 기반 모델의 경우 일부 제어 명령은 성공적으로 생성하였으나, 도구 이름 오류, 인자 형식 부적합, 동일 도구의 반복 호출 등으로 인해 실제 MCP 도구 호출 과정에서 실패 사례가 다수 관찰되었다. 이에 비해 멀티에이전트 기반 모델은 센서 목록 조회, 센서값 확인, 제어 목록 조회, 제어 명령 실행에 이르는 각 단계에서 목적에 부합하는 MCP 도구를 적절하게 선별하여 호출하는 것을 확인하였다. 이는 질의에 대한 제어 정확도를 확보하는 데 그치지 않고, 불필요한 도구 호출이나 중복 호출을 최소화함으로써 연산 자원 및 응답 시간 측면에서도 보다 효율적인 제어 수행이 가능함을 보여준다.

모델	센서 목록 조회	센서 값 조회	제어 수행
키워드 매핑	75%	33.3%	11.1%
단일 LLM	100%	66.7%	44.4%
멀티에이전트	100%	88.9%	92.6%

표 3. MCP 도구 호출 세부 성능 비교

IV. Conclusion

본 논문에서는 디지털 트윈 환경에서 자연어 기반 제어를 보다 안정적이고 정밀하게 수행하기 위한 역할 분담 기반의 멀티에이전트 모델을 제안하고, 실제 돈사 디지털 트윈 환경에 적용하여 그 효과를 검증하였다. 제안 모델은 기존 모델이 분석하지 못했던 모호하고 복잡한 사용자 질의에 대해서도 각 에이전트의 상호작용을 통해 판단 오류를 줄이고 최적의 제어 결과를 도출하는 것을 확인하였다. 이는 제안 모델이 비정형 환경에서도 훨씬 더 안정적이고 정밀한 의사결정을 수행할 수 있음을 시사한다. 향후 연구에서는 상황과 질의 특성에 따라 프로토타입을 동적으로 선택 및 구성하는 방식으로 발전시켜 좀 더 다양한 환경과 도메인에 유연하게 적용 가능하도록 구현할 예정이다.

ACKNOWLEDGMENT

이 논문은 2026년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No. RS-2022-II220545, 지능형 디지털 트윈 연합 객체 구성 및 데이터 프로세싱 기술 개발, 기여율 100%)

참 고 문 헌

- [1] SAEEDI, Sarow. Digital Twins and Their Applications in Modeling Different Levels of Manufacturing Systems: A Review. arXiv preprint arXiv:2511.06119, 2025.
- [2] OPOKU, De-Graft Joe, et al. Barriers to the adoption of digital twin in the construction industry: A literature review. In: Informatics. MDPI, 2023. p. 14.
- [3] LI, Xinyi, et al. A survey on LLM-based multi-agent systems: workflow, infrastructure, and challenges. Vicinagearth, 2024, 1.1: 9.