

트랜스포머 기반 다중 에이전트 동적 전장 돌발 상황 대응 기법

Transformer-Based Response Technique for Dynamic Battlefield Incidents in Multi-Agent Systems

김 성 민¹ · 김 희 연² · 이 재 호² · 홍 민 지² · 최 한 림[†]

Seong-Min Kim¹, Hee-Yeon Kim², Jae-Ho Lee², Min-Ji Hong², Han-Lim Choi[†]

Abstract: This study proposes a transformer-based approach for real-time response to dynamic battlefield incidents in heterogeneous multi-agent environments consisting of UAVs and UGVs. Initially, an offline global optimization was conducted to generate three million optimal response datasets for various incident scenarios, capturing optimal combinations of agents and corresponding missions based on agent, mission status, and incident type. Subsequently, transformer-based model was trained to approximate these global optimal solutions, enabling responses within approximately one second. Experimental results indicated that the transformer-based model significantly reduced computational time compared to global optimization while maintaining high response fidelity (75-83%). Additionally, the transformer-based model exhibited superior generalization performance over FCNN as the number of agents and missions increased.

Keywords: Multi-Robot System, Transformer, Unexpected Event, Collaborative Task

1. 서 론

최근 전장 환경은 Unmanned Aerial Vehicle (UAV), Unmanned Ground Vehicle (UGV)와 같은 고도화된 무인체제들이 인간을 보조하거나 대체하여 다양한 임무를 수행하는 방향으로 빠르게 진화하고 있다^[1]. 특히 인간의 개입 없이 무인 지상차량 (UGV)과 무인 항공기(UAV)가 협동하여 임무를 수행하는 사례가 늘어나면서 주요 임무를 자율적으로 수행할 수 있는 다중 에이전트 시스템이 주목받고 있다^[2]. 이러한 시스템은 의사결정 방식에 따라 크게 중앙집중형(Centralized)과 분산형(Decentralized)으로 나뉜다^[3].

중앙집중형은 지휘소나 호스트컴퓨터와 같은 중앙 노드가 접근 가능한 정보를 바탕으로 의사결정을 수행하는 방식이다. 중앙 노드가 의사결정 알고리즘을 통해 명령을 생성하면 각 에이전트는 해당 명령을 수행한다. 이 방식은 접근 가능한 전체 정보를 바탕으로 명령을 생성하므로 전역 최적화를 기대할 수 있다. 하지만 통신 지연이나 장애가 발생하면 시스템 전체가 마비될 수 있으며, 에이전트의 수가 늘어나거나 임무의 수가 늘어나는 등 전장 상황의 복잡도가 증가하는 경우 중앙 노드의 부하가 급증한다는 단점이 있다^[4]. 또한 중앙 노드가 공격받는 경우 전체 임무 과정이 중단될 수 있다는 취약점이 존재한다^[5].

분산형은 각 UGV와 UAV가 지역 센서 정보와 근처 동료 에이전트의 상태를 주고받으며 자율적으로 의사결정을 수행하는 방식이다. 통신 품질이 불안정한 전장에서도 근접 에이전트끼리 통신을 주고받으며 임무를 수행할 수 있으며, 일부 에이전트가 공격 또는 장애로 인해 사용이 불가능하더라도 나머지 에이전트를 통해 임무를 수행할 수 있다는 장점이 있다^[6]. 실제 전장 상황에는 여러 장애 요소가 발생할 수 있으므로 이러한 방식은 임무 수행에서 중앙집중형보다 높은 안정성을 가진다. 하지만 전체 정보가 아닌 일부 정보를 기반으로 판단하기 때문에

Received : Aug. 9. 2025; Revised : Jan. 28. 2026; Accepted : Mar. 9. 2026

※ This work was supported by the Agency For Defense Development grant funded by the Korean Government in 2025.

1. Principal Researcher, Aerospace Engineering, Korea Advanced Institute of Science and Technology, Daejeon, Korea (bcn05002@kaist.ac.kr)

2. Manager, Aerospace Engineering, Korea Advanced Institute of Science and Technology, Daejeon, Korea (dkssud715, 16jhlee, alswl@kaist.ac.kr)

† Associate Professor, Corresponding author: Aerospace Engineering, Korea Advanced Institute of Science and Technology, Daejeon, Korea (hanlimc@kaist.ac.kr)

전역 최적해를 보장하기 어려워 지역 최적해에 머무를 수 있다는 단점이 있다.

본 논문은 실제 전장 환경에서 발생할 수 있는 다양한 돌발 상황에서의 중앙집중형 의사결정에 대해 다룬다. 특히 실제 전장 환경에서 발생할 수 있는 적 발견, 위험물 발견 등의 돌발 상황에 효율적으로 대응하는 에이전트와 임무를 선택하는 알고리즘에 대해 다룬다. 돌발 상황 대응 알고리즘은 다음과 같은 요구사항을 가진다. 첫째, 기존 임무 계획을 최대한 유지하며 돌발 상황에 적절히 대응하여야 한다. 기존 임무 계획을 크게 변경하여 돌발 상황에 대응하는 알고리즘은 전체 임무 수행 면에서 전역 최적해를 찾기에 적절하다. 하지만 이는 돌발 상황 대응이 아닌 임무 재계획으로 다루어져야 한다. 둘째, 돌발 상황에 빠르게 응답할 수 있도록 실시간성이 보장되어야 한다. 돌발 상황 대응은 대부분 수 초의 빠른 시간 내에 이루어져야 한다. 오랜 시간이 걸리는 최적해 보다는 상황에 적절히 대응 가능한 실시간 해가 요구된다.

첫 번째 요구 사항을 만족하기 위해 기존 임무 계획을 유지하되, 돌발 상황 대응을 위해 새롭게 추가되는 임무를 기존 임무 계획 사이에 추가하는 방식을 사용하였다. 즉 돌발 상황 대응을 위해 일부 에이전트가 기존 임무를 미완료 상태로 두거나, 기존 임무로의 이동을 멈추고 돌발 대응 임무에 참여한다. 이후 돌발 대응 임무를 완료하면 각 에이전트는 기존 수행 임무에 복귀하여 남은 임무를 수행한다. 이를 통해 기존 임무의 틀을 유지하며 어떤 에이전트가 돌발 대응에 참여할 지와 어떤 임무를 수행할지를 결정하는 문제로 단순화하였다.

두 번째 요구의 충족을 위해 본 연구에서는 전역 최적 알고리즘을 통해 돌발 대응 해를 대규모로 생성한 후, 이를 인공 신경망(Neural Network)의 학습 레이블로 활용하는 신경망 근사 기법을 사용한다. 전역 최적화 기반 돌발 대응 방식은 문제 조건 하에서 가능한 모든 경우의 수를 고려하므로 계산 시간이 오래 걸리며, 에이전트의 수가 늘어남에 따라 계산 시간이 기하급수적으로 증가한다는 단점을 지닌다. 하지만 인공 신경망 근사를 통해 돌발 대응 해를 구하면 계산 시간을 현저히 줄일 수 있다.

전체 해를 탐색하는 전역 최적 알고리즘으로 돌발 상황 2,800,000 건에 대한 초기 임무 계획 및 최적 대응 에이전트, 임무 조합을 오프라인에서 계산하여 저장한다. 이후 이 데이터를 바탕으로 에이전트 유형(UAV/UGV) 및 정보(위치, 수행 중인 임무 등), 기존 계획된 임무 종류(감시, 정찰, 위험물 탐지, 교전) 및 진행 상태(임무 진행률, 참여중인 에이전트), 돌발 유형(적 또는 위험물 발견)의 정보를 입력으로 받아 사전 계산된 최적 대응을 근사하는 신경망을 학습한다. 학습 구조로 완전 연결 신경망, 임베딩 기반 트랜스포머 두 구조를 사용하고 비교한다. 각 모델의 추론 응답 시간과 근사 정확도, 실제 전장 비용을 계산하여 종합 평가한다.

본 논문의 남은 부분은 다음과 같이 진행된다. 2절에서 임무 계획을 최적화하는 연구와, 돌발 임무가 추가되어 임무 계획 변경을 수행하는 연구에 대해서 탐색한다. 이후 3절에서 본 연구에서 사용한 알고리즘을 검증하기 위한 검증 환경에 대해 다룬다. 4절에서는 효율적인 신경망 근사를 위해 사용한 데이터 구조 및 네트워크 구조에 대해 다루고, 5절에서는 그에 따른 결과를 다룬다. 최종적으로 6절에서는 결과를 분석하고 향후 연구 방향을 제시한다.

2. 선행 연구

다중 에이전트 동적 임무 할당 기법들은 크게 네 갈래로 나뉜다. 첫째, 경매 기반 접근은 에이전트들이 각자 임무에 입찰하고, 높은 입찰가를 제시한 에이전트가 해당 임무를 수행하도록 하는 분산 경매 메커니즘이다. 합의 기반 변형 알고리즘(Consensus-Based Bundle Algorithm)^[7]과 같은 기법은 임무 변들을 구성한 뒤 합의 단계를 통해 충돌 없는 할당을 도출하며, CBBA-PR (Partial Replanning)^[8]과 SWCBBA-PR (Soft-Window CBBA with Partial Reallocation)^[9]는 신규 임무가 발생했을 때 절반 정도의 이전 할당만 배제하고 재입찰하는 방식으로 응답 시간을 대폭 단축하였다. 그러나 이러한 경매 방식은 전역 최적해를 보장하기 어렵고, 가치 평가 설계 함수에 따라 해의 품질의 변동 폭이 크다는 한계가 있다.

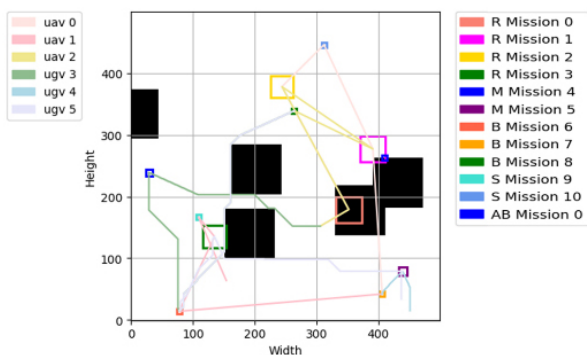
둘째, 계약망 프로토콜(Contract-Net Protocol)은 요청-응찰-선정-확인 네 단계의 협상을 통해 임무를 분배한다. 전통적인 CNP^[10]를 기반으로 한 BSE-CNP (Bidder Selection Enhanced CNP)^[11] 등은 입찰자 선정 메커니즘과 가상 의사결정자(Virtual Decision Maker)를 도입해 동적 돌발 상황에서도 유연한 재할당을 실현한다. 협상 기반 방식은 에이전트 개별의 의사결정을 존중한다는 강점을 지닌다. 하지만 다단계 협상이 누적되면 통신 비용과 응답 지연이 증가하고 복잡한 제약 조건(Time Window, 에너지 제약 등)을 동시에 처리하기 어렵다는 단점이 있다.

셋째, 수리최적화(Optimization) 방식의 접근은 MILP (Mixed-Integer Linear Programming)^[12], SMT (Satisfiability Modulo Theories), PSO (Particle Swarm Optimization)^[13], GA (Genetic Algorithm)^[14]와 같은 기법으로 글로벌 최적해를 산출한 뒤 돌발 임무가 발생하면 기존 해를 증분으로 갱신한다. SMT 기반 증분 최적화 연구는 이전 풀이과정의 학습을 재사용하여 응답 속도를 높였으나, 에이전트 수나 임무 수가 많아지면 여전히 계산 시간이 길어 실시간성 확보에 제약이 따른다. 계층형 동적 할당 기법(Stochastic Conflict-Based Allocation (SCoBA))^[15]은 개별 에이전트의 행동 정책 탐색과 상위층의 충돌 해소(conflict resolution)를 분리하여 해품질과 속도 간 적절한 균형을 보였으나, 여전히 대규모 환경에서는 부담이 크다.

넷째, 강화학습(Reinforcement Learning) 기반 기법은 다중 에이전트 심층 정책 강화학습(Multi-Agent RL (MARL))을 통해 에이전트들이 시행착오를 거쳐 돌발 상황 대응 전략을 학습한다. MADDPG (Multi-Agent Deep Deterministic Policy Gradient)^[16], PPO (Proximal Policy Optimization)^[17] 등을 변형한 방법론들은 모델-프리 특성으로 불확실성과 예측 불완전성에 강인하며, 공중 및 지상 군집 시뮬레이션에서 높은 임무 완수율을 입증했다. 하지만 충분한 학습 데이터를 수집하는 데 필요한 시간과 계산 자원이 방대하고, 새로운 돌발 유형이 등장했을 때 재훈련 없이 즉각 대응하기 어렵다는 문제가 있다.

3. 검증 환경

본 절에서는 신경망 근사 돌발 대응 알고리즘의 성능을 평가하기 위해 사용한 검증 환경^[18]에 대해 다룬다. [Fig. 1]은 본 연구에서 사용한 검증 환경 중 하나의 예시 시나리오를 나타내며, 이종 로봇들의 이동 경로와 임무 할당, 장애물 구성을 포함한다. 본 연구는 주어진 초기 임무 계획 하에서 다양한 돌발 상황에 직면한 이종 로봇(UAV, UGV)들의 실시간적 대응을 목표로 가진다. 이를 위해 여러 에이전트들의 협업을 모델링 할 수 있으며 현실성을 갖춘 시뮬레이션 기반의 평가 환경을 사용하였다. 검증 환경은 500×500으로 이루어진 2.5D 환경으로, UAV와 UGV의 다양한 물리적 특성을 반영하며, 각 로봇의 수와 세부 성능 파라미터, 임무의 수와 요구 조건을 조정함으로써 시나리오 별로 다른 상황을 구성할 수 있다. 또한 정찰, 감시, 위험물 탐지, 교전 등 복합 임무를 포함하고 있으며, 각 임무에 대한 중단·재개 및 시간 제약 조건이 명확히 정의되어 있다. 이외에도 장애물을 고려하기 위해, 장애물의 개수, 위치 및 크기를 무작



[Fig. 1] Example of simulation environments. Agent trajectories in a scenario with 6 agents and 11 missions. The colored lines represent the paths of UAVs and UGVs, while colored boxes indicate assigned missions (R: reconnaissance, M: mine detection, B: battle, S: surveillance, A: accident). Black rectangles denote obstacles in the environment

위로 설정하여 지도에 배치할 수 있다. 각 임무의 완수율과 관련된 보상 함수 역시 설정되어 있어 알고리즘이 시공간적 맥락을 인지하고 동적으로 전략을 수립하는 능력을 정량적으로 평가할 수 있다.

본 검증 환경의 UAV, UGV 유형은 이동 속도, 센서 범위, 내구성 등의 물리적 특성에서 뚜렷한 차이를 보이며, 적합한 임무의 종류 또한 상이하다. 예를 들어 UAV는 광범위한 시야 확보를 바탕으로 감시 및 정찰에 적합하고, UGV는 지상 기반의 위험물 탐지 및 제거 등 특수 임무 수행에 최적화되어 있다. 또한 동일한 유형 내에서도 속도와 등 세부 성능은 일정 범위 내에서 균등 분포로 무작위 샘플링되어, 현실적인 성능 분포를 반영하도록 구성되었다. 구체적으로, UAV는 이동속도가 10~15 grid/s 범위에서, UGV는 4~9 grid/s 범위를 가진다. 또한 UAV의 경우 장애물을 통과할 수 있어 돌발 대응에 대한 참여와 복귀를 신속하게 수행 가능하다. 이러한 이종성은 임무 간 이동 과정에서 특히 두드러지며, 로봇 유형에 따라 경로 계획 방식과 수행 전략이 달라지게 된다. 따라서, 이는 각 로봇 유형의 수행 특성과 임무 적합성을 반영하며, 환경 내 다양한 시나리오에서 로봇 간 차별화된 행동 패턴을 유도한다.

본 환경은 이러한 로봇 특성을 반영하며 지형 정보를 고려한 연속 공간 기반의 경로 계획 최적화를 지원한다. 각 로봇은 주어진 임무를 수행하기 위해 지형 장애물을 회피하며 목표 지점으로 이동해야 한다. UAV는 고도 제약이 없는 상황을 가정하여 일정한 속도로 비행할 수 있지만, UGV는 지형에 따라 이동 속도가 변화한다. 시뮬레이션은 이러한 제약을 정량적으로 반영함으로써 로봇 실제 이동 시간을 보다 정확히 추정하고, 다양한 임무 시나리오 하에서 돌발 대응 모델 성능 평가가 가능하도록 한다.

또한 본 환경은 연속적인 시공간 모델을 기반으로 한다. 공간적으로 로봇이 연속적인 위치에서 자유롭게 이동할 수 있어 장애물 회피, 목표 접근, 돌발 상황 대응 등의 경로를 현실적으로 모사할 수 있다. 시간적으로는 고정된 타임 스텝이 아닌 연속 시간에 기반한 계산 방식을 적용하여 임무 수행의 동적 변화를 정밀하게 추적할 수 있다. 이를 통해 감시, 정찰, 교전, 위험물 탐지와 같은 임무 수행 과정에서 로봇의 센서 범위, 가시성 조건, 협업 수행을 위한 경로 계획, 임무 거리 등을 연속적으로 계산할 수 있으며, 임무 성공률, 총 수행 시간, 자원 소모량 등의 주요 성능 지표를 정밀하게 측정하고 비교 평가할 수 있는 기반을 제공한다.

이러한 점은 로봇의 수, 임무의 수 외에도 여러 파라미터들을 조정하여 다양한 환경을 구상할 수 있다는 장점을 지닌다. 또한 서로 다른 특성을 지니는 이종 로봇을 고려함으로써 알고리즘이 특정 임무에 적합한 로봇을 적절히 선택할 수 있는지 역시 평가할 수 있다.

정찰 임무는 특정 지역 내의 적 존재 여부를 탐지하고 위치 정보를 수집하는 것을 목표로 하며, 본 환경에서는 최대 두 대의 로봇이 동시에 임무에 투입될 수 있도록 설계되었다. 두 개체가 협업할 경우 정찰 범위의 중복을 최소화하도록 경로가 분할되며, 센서 범위와 이동 성능을 고려한 효율적인 영역 정찰이 달성되도록 한다. 또한 정찰 임무는 수행 도중 돌발 상황이 발생하더라도 이후 유연한 재개가 가능하다는 특징을 갖는다. 임무가 일시 중단되면 기존에 탐색이 완료된 영역을 제외한 미탐색 구역을 기준으로 새로운 경로가 자동으로 계산되며, 해당 시점에서 로봇의 위치 변화나 협업 개체의 투입 여부에 따라 정찰 전략이 동적으로 조정된다. 이를 통해 돌발 대응 이후에도 불필요한 자원 낭비 없이 효율적인 임무 수행이 가능하도록 모델링되었다.

감시 임무는 임무 수행 전 사전 수집된 정보에 따른 주요 위치에서 정보를 지속적으로 수집하는 것을 목표로 하며, 지정된 시간 구간 내에 임무가 완료되어야 하는 시간 제약 조건이 부여된다. 각 감시 임무는 고유한 수행 시간 구간을 갖고 있으며, 해당 시간 구간 내에 임무를 수행해야만 고위험 대상에 대한 효과적인 대응이 가능하다. UAV는 기동성이 뛰어나 시간 구간 내에 임무를 완수하기에 비교적 유리하다. 감시 지점에는 최대한대의 로봇만 배치되며, 이를 통해 제한된 자원을 효율적으로 분배할 수 있다. 감시 임무는 정보의 시간적 연속성이 핵심이다. 따라서 수행 도중 임무가 중단되거나 다른 작업으로 전환되면 해당 지점에 대한 감시가 실패한 것으로 간주되며 재개는 허용되지 않는다. 이는 실시간 정보 확보가 중요한 감시 임무의 특수성을 반영한 제약으로, 돌발 대응에 있어 하나의 고려 요소로 작용한다.

위험물 탐지 임무는 특정 구역에 대하여 잠재적 위험 요소가 존재할 가능성이 높을 것으로 판단 시 수행된다. 본 검증 환경에서는 위험물 탐지 장비를 갖춘 로봇만이 해당 임무를 수행할 수 있도록 제한함으로써 실제 운용 환경의 제약 조건을 반영하였다. 해당 지역은 탐지가 전 영역에 대하여 완료될 때까지 고위험 구역으로 간주되며, 다른 로봇의 접근이 제한되어 경로 회피를 야기하여 전체 임무 효율에 영향을 미친다. 따라서 위험물 탐지 임무의 신속한 수행은 다른 임무 수행 경로의 최적화로 이어지며, 결과적으로 전체적인 효율을 높이는 핵심 요인으로 작용한다. 본 임무 역시 정찰 임무와 동일하게 중단 후 재개가 가능하며, 탐색이 완료되지 않은 구역을 기준으로 수행 경로가 동적으로 계산된다.

교전 임무는 적의 제압과 동시에 아군 로봇의 피해를 최소화하는 것을 주 목표로 하며, 이는 연속적으로 수행되는 후속 임무의 성공을 보장하기 위한 핵심 조건이다. 본 환경에서는 교전 임무의 비가역성을 현실적으로 반영하여 일단 전투가 개시되면 도중에 중단하거나 타 임무로 전환하는 것이 불가능하다고

가정한다. 전투 중 로봇의 상태 변화(예: 파손, 생존 여부)와 적 개체의 제거 여부는 임무 진행 상황을 실시간으로 평가하는 기준으로 활용되며, 이를 통해 교전의 결과와 그 영향이 후속 임무에 반영된다. 로봇 간의 전투 수행 능력은 이동 속도, 무장 성능, 내구성 등 물리적 특성에 따라 상이하게 나타나므로, 각 임무에 적합한 개체를 선별하고 배치하는 전략이 교전 결과에 큰 영향을 미친다. 특히 매복이나 기습 공격과 같은 돌발 상황에 적절히 대응하지 못하는 경우 해당 지역의 안전 확보가 지연되거나 아군 전력이 손실되어 이후 정찰, 감시, 위험물 탐지 등의 연속 임무 수행이 불가능해지는 연쇄적 실패로 이어질 수 있다. 따라서 교전 임무는 돌발 상황에 대한 신속한 판단 및 정확한 대응 전략을 요구한다.

본 연구는 돌발 상황 발생 시의 실시간 대응 알고리즘을 중심으로 하며, 이를 위해 각 시나리오마다 고유한 초기 임무 계획이 주어진 상태를 전제로 한다. 초기 계획은 각 임무의 소요 시간, 임무 달성 보상, 패널티(자원 손실)를 고려한 가중 합 기반의 목적 함수를 설정하고, 이에 따라 순차적 탐색 알고리즘(Sequential Greedy Algorithm)^[18]을 통해 수립하였다. 목적 함수에 대한 자세한 수식은 5절에서 다룬다. 이 초기 계획은 다양한 요소가 복합적으로 반영된 결과로 실제 운용 가능성을 고려한 현실적인 임무 배분 구조를 제공한다. 그러나 본 연구의 초점은 계획 수립 자체가 아닌 해당 계획 하에서 돌발 상황 발생 시 어떻게 협업을 고려하며 효과적으로 대응할 수 있는가에 있다. 초기 임무 계획은 타 로봇이 다른 개체의 임무 수행을 대체하는 것이 불가능하다고 가정하며, 특정 로봇이 기능을 상실할 경우 후속 임무의 연쇄적 실패가 발생하는 것으로 모델링된다. 이는 실제 전장 상황에서의 비가역적 제약을 반영한 것으로 돌발 상황에 대한 개별 로봇의 대응 전략뿐 아니라, 다른 로봇의 임무 수행에도 영향을 미치는 협업적 요소를 평가하기 위한 기반이 된다. 결과적으로 본 환경은 다양하고 구체적인 시나리오에서 정해진 계획 하에 돌발 상황이 전체 임무 흐름에 미치는 영향을 최소화하는 알고리즘의 대응 능력과 협업 효율성을 측정하는 데 목적이 있다.

본 환경에서는 실제 작전 상황에서 발생할 수 있는 다양한 돌발 상황을 모사하였다. 대표적인 돌발 상황으로는 이동 중 위험 가능성이 높은 지역 진입, 기습 공격 발생 또는 일방적인 적 탐지, 전투 중 적군의 후속 병력의 도착, 위험물 발견 등이 있다. 이러한 상황에 직면했을 때, 로봇은 즉각적인 대응을 수행하거나 회피 또는 후퇴와 같은 미대응 전략을 선택할 수 있으며, 각 선택은 전체 임무 흐름에 상이한 영향을 미친다.

돌발 상황에 대한 대응 여부는 단순히 개별 로봇의 생존과 임무 수행 성공률뿐 아니라, 기존에 진행 중이던 임무의 중단 및 재개 가능성, 협업 구조의 재구성, 자원 제약 조건의 유지 여부에 밀접하게 연관된다. 예를 들어, 특정 로봇이 돌발 상황에

대응하기 위해 기존 임무를 중단한 경우 정찰이나 위험물 탐지와 같은 일부 임무는 재개가 가능하지만, 감시 임무는 실패로 간주된다. 또한 교전 중인 개체는 돌발 상황에 대한 지원이 불가능하며, 위험물 제거 등 특수 장비가 필요한 상황에서는 해당 기능을 갖춘 로봇만이 대응할 수 있어 돌발 대응이 불가능할 수 있다. 이러한 다양한 돌발 상황을 환경 내에 반영함으로써 알고리즘이 예상치 못한 상황 변화에 대해 얼마나 유연하고 강건하게 대응할 수 있는지를 정량적으로 평가한다. 이는 단순한 경로 최적화나 임무 할당 정확도를 넘어 실제 전장 환경에서의 실시간 대응력과 시스템 신뢰성 확보에 중요한 역할을 한다.

4. 데이터셋 및 모델 구조

본 절에서는 돌발 상황 대응을 위한 데이터셋과 네트워크 구조를 설명한다. 데이터셋으로 이전에 설명한 환경에서 전역 최적 알고리즘을 통해 생성한 약 2,800,000개의 돌발 대응 데이터를 사용한다.

데이터셋 생성 과정에서 사용한 전역 최적화 알고리즘은 전체 로봇 집합 R 중 돌발 대응이 가능한 로봇들로 구성된 부분 집합 $R_f \subseteq R$ 를 먼저 정의한 후, R_f 로부터 생성 가능한 부분 집합들을 모두 고려하여 각 조합의 돌발 대응 비용을 계산하고 최소 비용을 갖는 조합을 선택하는 방식으로 수행된다. 집합 R_f 는 돌발 상황 발생 시점에서 각 로봇의 현재 임무 상태 및 수행 능력을 고려하여 구성된다. 예를 들어 교전 임무를 수행 중인 로봇은 돌발 대응에서 제외되며, 장애물이 존재하는 지역의 탐색 임무에는 UGV가 투입될 수 없고, 위험물 탐색 및 위험물 제거 임무에는 UAV가 투입될 수 없다. 이러한 조건을 통해 돌발 대응이 불가능한 로봇들은 사전에 제거한다. 전체 로봇 수를 $|R|=n$, 돌발 대응이 가능한 로봇 수를 $|R_f|=m$ 이라 할 때 고려되는 로봇 조합의 수를 최대 2^m 이며, 데이터 생성 환경에서는 상기 조건에 의해 $m \leq n$ 이 되어 탐색 공간이 감소한다.

각 데이터는 환경 내 다양한 무작위성 요소를 반영하여 생성한다. 이에 따라 임무 위치, 교전 임무의 적 규모 및 공격력, 감시 임무의 TimeWindow 시작 시간, 아군 로봇 초기 배치 등이 시나리오별로 다르도록 구성한다. 무작위성 요소들은 균등 분포를 따르도록 모델링한다. 이러한 환경 설계를 통해 매 에피소드마다 미세하게 다른 시나리오를 생성하여, 알고리즘이 다양한 상황 변화에 대해 유연하게 대응할 수 있는 능력을 평가한다. 대응 방식의 일관성을 위하여 목적 함수에서 사용하는 가중치 및 임무 별 보상 값은 고정한다.

데이터셋은 에이전트와 임무 수가 각각 6개와 11개, 9개와 17개, 15개와 24개인 세 가지 상황에서 각각 생성한다. 각 상황별 데이터셋의 크기는 순서대로 936,060개, 1,143,682개,

764,304개이다. 데이터셋은 입력 데이터로 에이전트, 임무, 돌발 상황 정보를 가지고, 출력 데이터로는 돌발 상황에 대응할 에이전트 조합과 대응 임무 정보를 포함한다.

에이전트와 임무의 수 별로 모델의 입출력의 크기가 다르기 때문에 각 상황 별로 파라미터 수를 달리한다. 파라미터 수는 베이지안 최적화를 통해 해당 상황에 가장 적절한 파라미터의 수로 설정한다. 완전 연결 신경망의 경우 은닉층 2개의 크기를 최적화하고, 임베딩 기반 트랜스포머 구조의 경우 인코더와 각 임베딩 벡터의 크기를 최적화한다.

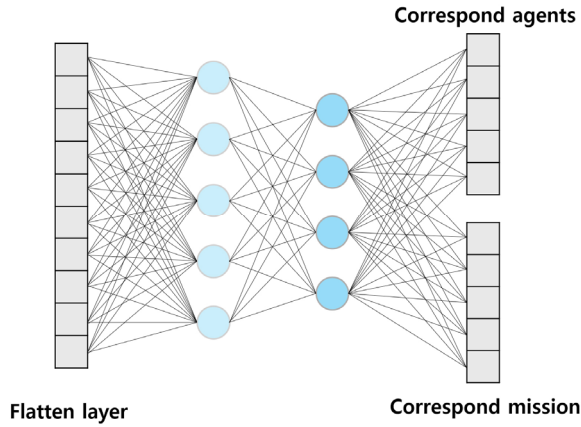
에이전트 정보는 위치, 돌발 대응 시점의 체력, 에이전트 종류, 그리고 에이전트가 돌발 상황 발생 시 대응 가능 여부를 포함한다. 이때 에이전트의 대응 가능 여부는 에이전트가 교전 등의 이유로 인해 돌발 상황에 참여가 불가능함을 나타낸다.

임무 정보는 각 임무에 할당된 에이전트들의 인덱스, 교전 임무의 경우 돌발 발생 시점에서의 교전 여부를 포함한다. 정찰이나 지뢰 제거 임무의 경우 임무 영역의 크기, 모든 임무에 대해서는 예상 종료 시각, 위치, 임무의 종류 정보가 제공된다. 돌발 상황 정보는 돌발 발생 시각, 위치 및 돌발 상황 유형 정보를 포함한다.

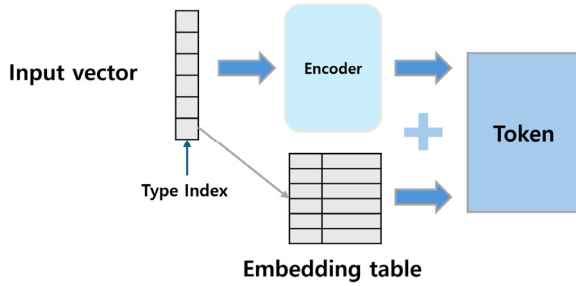
완전 연결 신경망은 입력 정보를 스케일링 한 후 다차원의 입력 데이터를 1차원 벡터로 변환하여 사용한다. 출력층은 원-핫 인코딩 형태로 대응 임무의 종류를, 멀티-핫 인코딩 형태로 대응 에이전트의 조합을 출력하도록 설계하였다. 신경망은 전역 최적 알고리즘 결과와 동일한 대응 결과를 출력하도록 학습된다. 사용된 완전 연결 신경망은 [Fig. 2]와 같이 은닉층을 갖는다. 그러나 이러한 방식은 입력 데이터를 1차원 벡터로 변환하는 과정에서 데이터가 본래 가지고 있던 구조적, 의미적 정보를 손실한다는 한계가 있다.

이러한 한계를 극복하기 위해, 임베딩 단계와 트랜스포머를 결합한 모델을 도입한다. 본 모델은 에이전트 종류, 임무 종류, 돌발 임무 종류와 같은 범주형 데이터를 학습 가능한 임베딩 테이블을 통해 고유한 벡터로 변환 후 트랜스포머의 입력 토큰으로 사용한다. 이를 통해 단순히 정수 인덱스나 원-핫 인코딩보다 더 유용한 정보를 모델에 전달할 수 있게 된다.

트랜스포머의 셀프 어텐션 메커니즘은 입력 데이터 내 요소들의 관계와 고유한 의미를 효과적으로 보존하고 학습할 수 있다. 임베딩 기반 트랜스포머 구조는 [Fig. 3]과 같이 입력 데이터를 토큰 형태로 변환 후 처리한다. 에이전트와 임무, 그리고 돌발 대응 임무의 종류 정보를 제외한 나머지 입력 데이터는 인코더를 통과하여 모델이 필요한 추가적인 정보를 추출할 수 있도록 하였다. 이후 인코더를 통과한 정보와 임베딩 벡터가 결합되어 최종 토큰을 형성하며, 이를 통해 데이터 크기나 구조가 달라도 일관된 형태의 토큰으로 트랜스포머 모델에서 처리 가능하다.



[Fig. 2] Fully connected neural network structure: takes flattened input features, processes them through three fully connected layers, and outputs the combination of responding agents along with the corresponding mission



[Fig. 3] Transformer tokenizer: the input vector is divided into a type index and the remaining components; the type index is mapped to an embedding via a lookup table, while the remaining components are encoded through an encoder. The embedding and the encoded representation are then concatenated to form the final token

5. 실험 결과

[Table 1]은 각 알고리즘이 서로 다른 시나리오에 대해 돌발 대응을 수행한 결과를 나타낸다. 에이전트와 임무의 수, 그리고 에이전트의 종류를 고정하고 나머지 요소는 모두 무작위로

선택되도록 하였다. 즉 임무의 종류와 각 시나리오의 지도는 무작위로 생성하였다. 각 에이전트와 임무 개수별로 200개의 시나리오를 생성하였으며, 개별 시나리오에 대해 각 알고리즘이 돌발 대응한 결과를 평균으로 나타내었다. 평가를 위해 CPU : i7-14700K, GPU : RTX 4090를 장착한 컴퓨터를 이용하였다. 각 신경망 모델은 Python의 torch 라이브러리를 이용해 학습하였으며, Window 11 환경에서 평가하였다.

$$J = w_1 \sum u_i - w_2 T - w_3 P \quad (1)$$

목적 함수(J)는 현재 임무 할당을 환경에서 평가하는 값으로써 높을수록 좋은 돌발 대응을 의미한다. 즉 선택된 대응 에이전트 및 임무가 적절한지를 평가한다. 목적 함수는 식 (1)과 같이 세 개의 항의 가중합으로 구성된다. 각 항은 완료한 임무 i 개를 각각 완료함으로써 얻을 수 있는 점수인 s_i 의 합과 최종 임무 완료 시각인 T , 그리고 에이전트가 받은 피해를 반영하는 P 로 구성된다. 가중치 w_1, w_2, w_3 은 실험적 방법을 통해 전역 최적 알고리즘의 결과가 이상적으로 나올 수 있는 수치를 선택하였다. 전역 최적 알고리즘, 완전 연결 신경망 그리고 트랜스포머 구조를 이용하여 평가를 수행하였고 그 결과를 비교하였다.

목적 함수만으로는 임무나 에이전트의 수가 달라지는 경우에 대해 비교하기 어렵다. 따라서 전역 최적 알고리즘의 목적 함수 값에 비해 각 알고리즘의 목적 함수 값이 어느 정도의 성능을 내는 지를 계산하여 Fidelity로 나타내었다. 즉 전역 최적 알고리즘의 목적 함수 값을 기준으로 몇 %의 성능을 가지는지를 표현하였다. 개별 Fidelity의 평균을 통해 해당 에이전트 및 임무의 개수에서 특정 알고리즘이 어느 정도의 성능을 내는지 확인할 수 있다. Global optimal은 전역 최적 알고리즘을 의미하고, Fully Connected Neural Net은 완전 연결 신경망 구조를 의미한다. Embedding based Transformer는 임베딩 기반 트랜스포머 구조를 의미한다. 각 알고리즘의 성능을 에이전트 및 임무의 개수가 다른 경우에 대해서 각각 비교하고 성능 지표를 산출하였다. 전역 최적 알고리즘의 경우 에이전트 및 미션의 개수가 늘어남에 따라 계산 시간이 기하급수적으로 증가한다. 최종적

[Table 1] Experimental results

Agents_Missions (UAVs, UGVs)	Global Optimal			Fully Connected Neural Net			Embedding based Transformer		
	Time(s)	J	Fidelity (%)	Time(s)	J	Fidelity (%)	Time(s)	J	Fidelity (%)
6_10 (3, 3)	8.290 ±10.268	-134.85 ±49.09	-	0.708 ±0.300	-162.13 ±63.66	77.274 ±31.38	0.801 ±0.311	-155.48 ±60.31	82.14 ±26.18
9_17 (5, 4)	28.146 ±48.913	-102.49 ±42.58	-	0.908 ±0.268	-117.29 ±46.39	80.290 ±31.62	0.973 ±0.347	-114.33 ±44.67	83.81 ±24.76
15_24 (8, 7)	135.566 ±269.27	-21.83 ±16.29	-	1.067 ±0.485	-27.55 ±17.85	66.756 ±43.06	1.230 ±0.615	-25.86 ±17.47	75.52 ±34.16

으로 15개의 에이전트, 24개의 임무가 있는 경우 평균적으로 135s의 시간이 소요된다. 이와 같이 긴 계산 시간은 돌발 상황 대응이라는 목표에 부합하지 않는다.

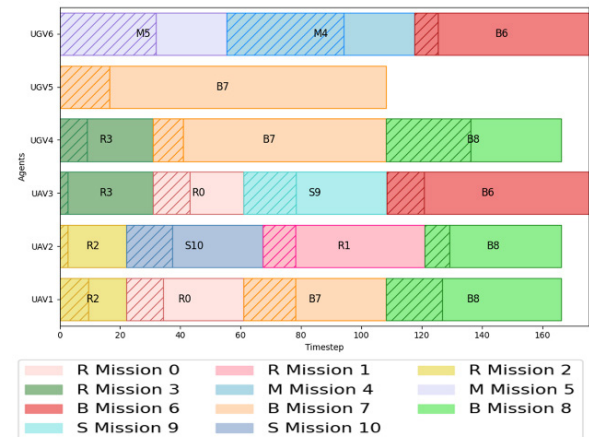
완전 연결 신경망 구조 및 임베딩 기반 트랜스포머 구조를 통한 돌발 대응은 약 1초 내외의 계산 시간을 보인다. 완전 연결 신경망 구조는 임베딩 기반 트랜스포머 구조보다 약간 적은 시간을 소요하나 그리 큰 차이를 보이지 않는다. 에이전트 및 임무의 수가 증가함에 따라 요구되는 파라미터의 개수가 증가하기 때문에 계산 시간이 다소 증가하는 것을 확인할 수 있다. 하지만 항상 전역 최적 알고리즘의 계산 시간에 비해 적은 계산 시간이 소요되며, 이러한 경향은 에이전트 및 임무의 개수가 늘어날수록 더 드러난다. 전역 최적 알고리즘의 경우 에이전트 및 임무의 수가 증가할수록 계산 시간의 표준편차가 상당히 증가하여 불안정하지만, 나머지 두 구조는 표준편차가 작아 안정적이다.

Fidelity의 경우 9개의 에이전트 및 17개의 임무에서 가장 높은 값을 보이며, 15개의 에이전트 및 24개의 임무에서 가장 낮은 값을 보인다. 일정 수 이하의 에이전트 및 임무에서는 높은 성능을 보이지만, 수가 늘어나면 성능이 일부 감소한다. 이는 대응 가능한 에이전트의 수가 증가함에 따라 문제의 난이도가 증가하기 때문으로 보인다. 임베딩 기반 트랜스포머 구조는 항상 완전 연결 신경망 구조에 비해 좋은 성능을 보인다. 문제의 복잡도가 높아지는 경우에 완전 연결 신경망과의 격차가 더 벌어지는 것 역시 확인할 수 있다. 또한 표준편차 역시 제안한 모델 구조가 더 작은 값을 보여 안정적인 성능을 기대할 수 있다.

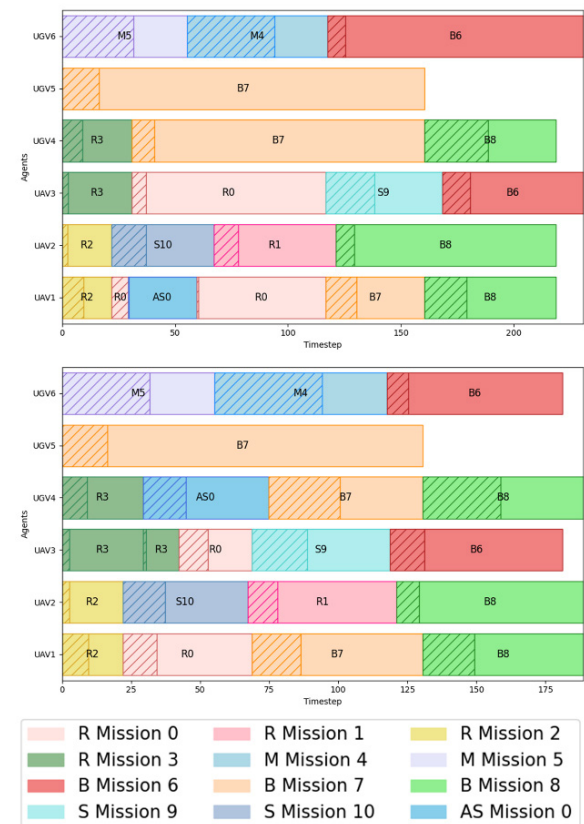
임베딩 기반 트랜스포머가 완전 연결 신경망보다 높은 성능을 보임을 확인하기 위해, 통계적 방법을 이용하여 검증하였다. 현재 결과는 동일한 시나리오에 대해 각 알고리즘이 대응한 결과이므로 윌콕슨 부호 순위 검정(Wilcoxon signed-rank test)을 수행하였다. 수행 결과, 상대적으로 규모가 작은 6_11 시나리오에서는 두 알고리즘 간의 유의미한 차이가 나타나지 않았으나($p = 0.163$), 문제의 복잡도가 증가함에 따라 제안 방법의 우수성이 뚜렷하게 나타났다. 특히 중형 규모인 9_17 시나리오($p = 0.038$)와 대형 규모인 15_24($p = 0.030$) 모두에서 통계적 유의수준($p < 0.05$)을 만족하며 임베딩 기반 트랜스포머 구조가 완전 연결 신경망을 유의미하게 앞서는 것으로 나타났다. 이는 탐색 공간이 확장될수록 임베딩 기반 트랜스포머 구조의 어텐션 매커니즘이 다중 에이전트 간의 상관관계를 보다 효과적으로 학습하며, 문제 규모에 따른 우수한 확장성을 가짐을 입증한다.

본 연구의 다양한 복합 임무 시나리오 중 제안된 모델과 비교 모델 간의 대응 전략 차이를 가장 명확하게 확인할 수 있는 사례를 발췌하여 [Fig. 4]-[Fig. 7]에 제시하였다. 모든 실험 시나리오는 감시, 정찰, 교전, 위험물 탐지 등 여러 임무가 동시에

진행되는 복합적 전장 환경을 반영하며, 돌발 상황의 발생 위치와 시점을 무작위로 변화시켜 구성되었다. 이 가운데 각 알고리즘의 특성이 가장 두드러지게 드러나는 대표 사례 두 가지를 선택하여, 알고리즘의 실시간 대응 품질 및 그로 인한 전체 임무 흐름의 변화 양상을 비교, 분석하였다.

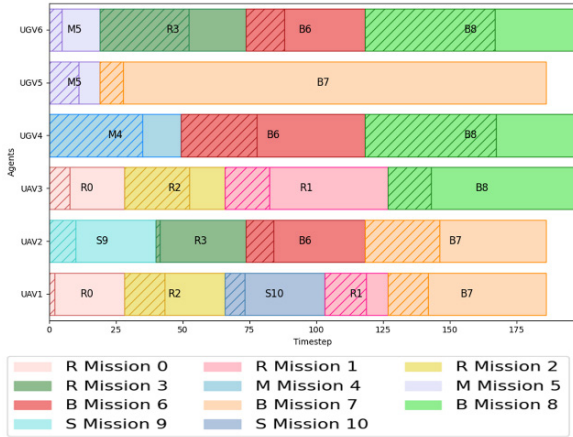


[Fig. 4] Initial allocation (Case 1): 1st example scenario of initial allocation with 6 agents and 11 missions

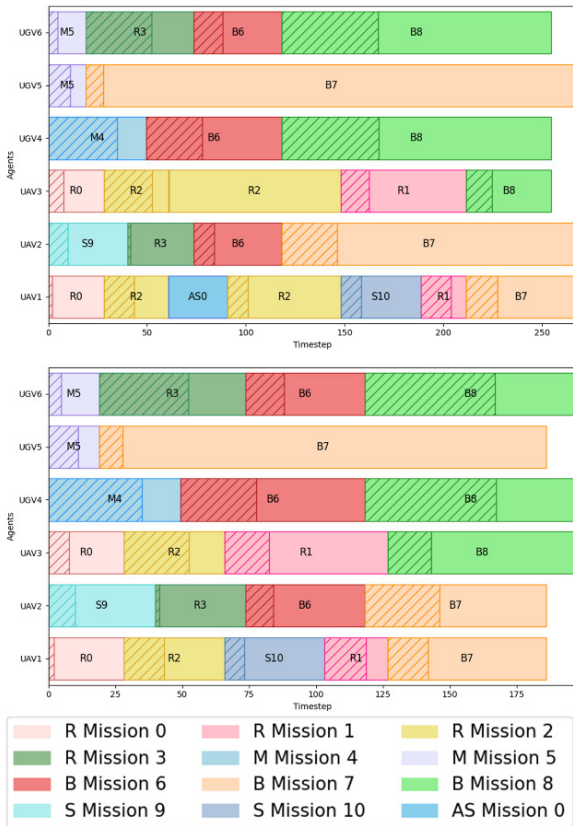


[Fig. 5] Accident corresponding (Case 1) - comparison of response results between neural network (top), transformer (bottom): AS Mission 0 represents the accident mission

[Fig. 4]-[Fig. 7]은 돌발 상황 발생 전 초기 전장 임무할당테이블로 가로축은 시간, 세로축은 로봇을 의미한다. 정찰(R), 위험물탐지(M), 교전(B), 감시(S) 과업들이 각 로봇에 할당된 것을 확인할 수 있고, 각 과업 블록 안의 빗금 친 구간은 임무 지역까지의 이동시간을 의미한다.



[Fig. 6] Initial allocation (Case 2): 2nd example scenario of initial allocation with 6 agents and 11 missions



[Fig. 7] Accident corresponding (Case 2) - comparison of response results between neural network (top), transformer (bottom): AS Mission 0 represents the accident mission

[Fig. 4]-[Fig. 5]는 첫 번째 예시 시나리오를 나타낸다. 케이스 1에서는 초기 임무계획을 따라 각 에이전트가 임무를 수행하던 중 29.39초에 돌발이 발생하였다. UAV1이 정찰(R) 0을 수행하기 위해 이동 중 적을 발견하였다. 이때 돌발대응 알고리즘은 환경의 복합적 상황을 고려하여 감시(AS), 교전(AB), 또는 미 대응을 선택하는 판단을 내려야 한다.

[Fig. 5]는 해당 시나리오에 대해 두 알고리즘의 돌발 대응 결과 차이를 보여준다. 완전 연결 신경망 구조의 경우, UAV1이 돌발상황에 대하여 적을 일정시간 감시하다 정찰(R) 0으로 복귀하였다. 이 과정에서 늘어난 이동 거리로 인해 후속 과업들인 교전(B) 7, 8이 연쇄적으로 지연되었다. 또한 정찰(R) 0을 협업하여 수행하던 UAV3이 영향을 받아 감시(S) 9의 시작이 늦어졌고, 결과적으로 전체 임무 수행 효율이 저하되었다.

임베딩 기반 트랜스포머 구조의 경우, 마찬가지로 감시를 선택했지만 대응 주체로 UGV4를 할당하였다. UAV1이 대응하는 경우보다 UGV4가 수행하는 것이 지연을 최소화하였는데, UGV4는 교전(B) 7, 8에만 영향을 주며, 특히 교전(B) 7은 UAV1 도착 이후 시작되기 때문에 전체 임무에 대해 영향이 작았다. 이를 통해 임베딩 기반 트랜스포머 구조가 더 복합적인 협업 구조와 후속 영향까지 고려한 의사결정을 수행함을 확인할 수 있다.

[Fig. 6]-[Fig. 7]은 두 번째 예시 시나리오를 나타낸다. 케이스 2에서는 돌발 상황의 영향도가 낮아 의도적으로 대응하지 않는 것이 오히려 효과적일 수 있는 경우를 보여준다. [Fig. 6]는 두 번째 예시 시나리오의 초기 임무 계획을 나타낸다.

[Fig. 7]은 60.71초 시점, UAV 0이 정찰(R) 2를 수행하던 중 실시간으로 탐지된 적군에 대해 각 알고리즘이 어떻게 대응한지를 보여준다. 완전 연결 신경망 구조는 임무계획에 반영되지 않은 적에 대한 불확실성을 줄이기 위해 감시를 선택하였지만, 이로 인해 UAV1에 할당된 후속 과업들이 모두 지연되었다. 결과적으로 전체 임무 종료 시각이 70초가량 늦어지게 되었다. 반면, 임베딩 기반 트랜스포머 구조는 해당 적군이 이후 작전에 미칠 영향이 미미 할 것으로 판단하여 미대응을 선택하였다. 이는 UAV1의 이동 경로가 불필요하게 늘어나는 것을 방지하며 전체 임무 소요 시간을 단축시켰다. 결과적으로 이는 후속 임무의 연쇄적인 지연을 방지하고 자원의 낭비를 줄이는 등 전장 상황 전체의 종합적인 효율성의 향상으로 이어졌다.

6. 결론 및 향후 연구

본 연구에서는 이기종 다중 에이전트 시스템 운용 중 발생하는 돌발 상황에 대응하기 위한 새로운 의사결정 모델을 제안하고 그 성능을 검증하였다. 기존의 전역 최적 알고리즘은 최적의

대응 결과를 보장하지만, 문제의 복잡도가 증가함에 따라 계산 시간이 기하급수적으로 늘어나 실시간성이 요구되는 돌발 상황에 적용하기 어렵다는 명백한 한계를 가진다. 이러한 문제를 해결하기 위해, 본 논문에서는 데이터 내의 구조적, 의미적 관계를 효과적으로 학습할 수 있는 임베딩 기반 트랜스포머 구조를 제안한다. 실험 결과, 제안한 모델은 전역 최적 알고리즘 대비 현저히 짧은 약 1초 내외의 계산 시간으로 실시간 대응 가능성을 확보하였다.

대응의 질적 측면에서도 제안 모델의 우수성이 뚜렷하게 나타났다. 임베딩 기반 트랜스포머 구조 모델은 모든 실험 환경에서 완전 연결 신경망보다 높은 Fidelity를 기록했으며, 특히 에이전트와 임무의 수가 늘어나는 복잡한 시나리오일수록 그 성능 격차는 더욱 크게 벌어졌다. 이는 제안 모델이 문제의 복잡성에 더 강건하며 뛰어난 일반화 성능을 가짐을 시사한다. 또한, 특정 시나리오 분석을 통해 임베딩 기반 트랜스포머 구조가 당면한 돌발 상황만 처리하는 것이 아니라, 대응으로 인한 후속 임무들의 연쇄적 지연까지 고려하여 전체 임무 효율을 극대화하는 방향으로 대응함을 확인하였다.

결론적으로 임베딩 기반 트랜스포머 구조는 신속성과 정확성을 모두 만족시키는 강력한 돌발 대응 알고리즘이다. 본 연구에서는 복잡한 시공간적 관계 파악이 핵심인 이기종 다중 에이전트 전장 환경의 의사결정 문제에 임베딩 기반 트랜스포머 구조가 매우 효과적인 접근법이 될 수 있음을 입증했다.

하지만 에이전트 및 초기 임무의 수가 증가함에 따라 문제의 복잡도가 증가하여 적절한 에이전트를 선택하지 못하는 문제가 발생하였다. 또한 네트워크의 구조는 에이전트와 임무의 수가 달라질 때 마다 새로운 네트워크를 학습시켜야 된다는 단점을 가진다.

향후 연구는 이러한 문제를 해결하는 방향으로 진행되어야 한다. 먼저, 에이전트의 수가 증가하여도 적절한 에이전트들이 적절한 수만큼 선정되도록 네트워크 구조를 변경하여야 한다. 현재는 개별 에이전트가 돌발 대응에 참여하기 적절한지 점수를 매긴 후 일정 점수를 넘으면 돌발 대응에 참여하도록 하는 방식을 사용한다. 이 대신 단순히 각 에이전트에 점수를 매기는 것이 아니라 서로의 협업을 고려하여 전체 조합에 대해 점수를 매기는 방식의 네트워크 구조 등이 고려될 수 있다.

에이전트의 수와 임무의 개수가 달라지는 것에 대응 가능한 네트워크 구조 역시 고려할 수 있다. 현재는 각 상황별로 독립적인 네트워크 구조를 사용하는데, 모든 상황에 적용 가능한 네트워크 구조를 이용한다면 하나의 모델에 대해 더 다양한 데이터로 학습이 가능하며 다양한 상황에 대응 가능하게 된다. 이를 통해 모델의 일반화 성능 향상을 기대할 수 있다.

References

- [1] I. Munasinghe, A. Perera, and R. C. Deo, "A Comprehensive Review of UAV-UGV Collaboration: Advancements and Challenges," *Journal of Sensor and Actuator Networks*, vol. 13, no. 6, pp. 81, DOI: 10.3390/jsan13060081.
- [2] Y. Rizk, M. Awad, and E. W. Tunstel, "Cooperative Heterogeneous Multi-Robot Systems: A Survey," *ACM Computing Surveys (CSUR)*, vol. 52, no. 2, pp. 1-31, DOI: 10.1145/3303848.
- [3] J. van der Horst and J. Noble, "Distributed and Centralized Task Allocation: When and Where to Use Them," *Fourth IEEE International Conference on Self-Adaptive and Self-Organizing Systems Workshop*, Budapest, Hungary, pp. 1-8, 2010, DOI: 10.1109/SASOW.2010.8.
- [4] B. P. Gerkey and M. J. Mataric, "A Formal Analysis and Taxonomy of Task Allocation in Multi-Robot Systems," *The International Journal of Robotics Research*, vol. 23, no. 9, pp. 939-954, DOI: 10.1177/0278364904045564.
- [5] L. Ratnabala, A. Fedoseev, R. Peter, and D. Tsetserukou, "MAGNET: Multi-Agent Graph Neural Network-Based Efficient Task Allocation for Autonomous Vehicles with Deep Reinforcement Learning," *2025 IEEE Intelligent Vehicles Symposium (IV)*, Cluj-Napoca, Romania, pp. 970-975, 2025, DOI: 10.1109/IV64158.2025.11097641.
- [6] H. A. Kadir and M. R. Arshad, "Decentralized cooperative communication framework for heterogeneous multi-agent system," *Journal of Communications*, vol. 9, no. 2, pp. 163-170, DOI: 10.12720/jcm.9.2.163-170.
- [7] H. L. Choi, L. Brunet, and J. P. How, "Consensus-based decentralized auctions for robust task allocation," *IEEE Transactions on Robotics*, vol. 25, no. 4, pp. 912-926, DOI: 10.1109/TRO.2009.2022423.
- [8] N. Buckman, H. L. Choi, and J. P. How, "Partial Replanning for Decentralized Dynamic Task Allocation," *AIAA Scitech Forum*, San Diego, USA, pp. 0915, 2019, DOI: 10.2514/6.2019-0915.
- [9] J. Li, B. Liu, C. Liu, and C. Lin, "Dynamic Task Allocation for Heterogeneous Multi-autonomous Underwater Vehicle Collaboration Under Mine Countermeasures Missions," *Journal of Marine Science and Engineering*, vol. 13, no. 3, pp. 465, DOI: 10.3390/jmse13030465.
- [10] R. G. Smith, "The Contract Net Protocol: High-Level Communication and Control in a Distributed Problem Solver," *IEEE Transactions on Computers*, vol. C-29, no. 12, pp. 1104-1113, DOI: 10.1109/TC.1980.1675516.
- [11] Z. Zhou, G. Shen, W. Niu, B. He, and Y. Shen, "A Task Assignment Strategy for Multi-AUV Collaborative Hunting Problem," *OCEANS 2022, Hampton Roads*, Hampton Road, VA, USA, pp. 1-6, 2022, DOI: 10.1109/OCEANS47191.2022.9977373.
- [12] N. Atay and B. Bayazit, "Mixed-Integer Linear Programming Solution to Multi-Robot Task Allocation Problem," *All Computer Science and Engineering Research*, Washington, USA, Rep. WUCSE-2006-54, 2006, DOI: 10.7936/K7R49P04.

- [13] C. Wei, Z. Ji, and B. Cai, "Particle Swarm Optimization for Cooperative Multi-Robot Task Allocation: A Multi-Objective Approach," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2530-2537, DOI: 10.1109/LRA.2020.2972894.
- [14] J. G. Martin, J. R. D. Frejo, R. A. García, and E. F. Camacho, "Multi-robot task allocation problem with multiple nonlinear criteria using branch and bound and genetic algorithms," *Intelligent Service Robotics*, vol. 14, no. 5, pp. 707-727, DOI: 10.1007/s11370-021-00393-4.
- [15] S. Choudhury, J. K. Gupta, M. J. Kochenderfer, D. Sadigh, and J. Bohg, "Dynamic multi-robot task allocation under uncertainty and temporal constraints," *Autonomous Robots*, vol. 46, no. 1, pp. 231-247, DOI: 10.1007/s10514-021-10022-9.
- [16] R. Wang, Y. Xia, Y. Wei, Z. Pan, and J. Li, "MADDPG-Based Distributed Cooperative Search Strategy for Heterogeneous Agents System," *7th Chinese Conference on Swarm Intelligence and Cooperative Control*, Nanjing, China, pp. 292-305, 2023, DOI: 10.1007/978-981-97-3332-3_26.
- [17] S. Wang, N. Niu, Z. Wang, Y. Lv, and J. Zhang, "Heterogeneous Multi-Agent Task Planning Method in Complex Marine Environment," *IEEE Access*, vol. 13, pp. 84202-84216, DOI: 10.1109/ACCESS.2025.3567503.
- [18] J.-H. Lee, M.-J. Hong, H.-Y. Kim, S.-M. Kim, and H.-L. Choi, "Continuous Verification Environment and Dynamic Incident Response Techniques for Collaborative Robots," *Journal of Korea Robotics Society*, vol. 20, no. 3. The Korea Robotics Society, pp. 520-528, DOI: 10.7746/jkros.2025.20.3.520.

**김 성 민**

2024 한국항공대학교 항공우주 및 기계공학부 (학사)

2024~현재 한국과학기술원 항공우주공학과 (석사)

관심분야: 다중 로봇 시스템, 최적화, 강화학습

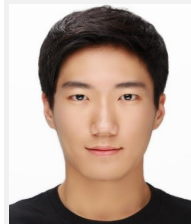
**김 희 연**

2023 한국과학기술원 기계공학과(학사)

2025 한국과학기술원 김재철AI대학원(석사)

2025~현재 한국과학기술원 항공우주공학과 (박사)

관심분야: 다중로봇 시스템, 강화학습

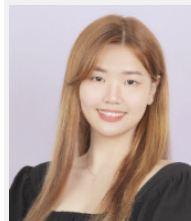
**이 재 호**

2021 한국과학기술원 항공우주공학과(학사)

2023 한국과학기술원 항공우주공학과(석사)

2023~현재 한국과학기술원 항공우주공학과 (박사)

관심분야: 다중로봇 시스템, 휴리스틱 서치

**홍 민 지**

2022 한국항공대학교 항공우주 및 기계공학부 (학사)

2024 한국과학기술원 항공우주공학과(석사)

2024~현재 한국과학기술원 항공우주공학과 (박사)

관심분야: 다중로봇 시스템, 경로 계획, 임무 할당

**최 한 림**

2000 한국과학기술원 항공우주공학과(학사)

2002 한국과학기술원 항공우주공학과(석사)

2009 메사추세츠 공과대학교 항공우주공학 (공학박사)

2008~현재 한국과학기술원 교수

관심분야: 다중 로봇 시스템, 지능형 항공우주 시스템, 불확실성 정량화 및 학습