

분산형 하이브리드 AI 기반 모바일 RPG NPC 시스템의 구현

김동주¹, 고석주^{2*}

¹경북대학교 컴퓨터학부 박사과정, ²경북대학교 컴퓨터학부 교수

Implementation of a Mobile RPG NPC System Based on Distributed Hybrid AI

Dongju Kim¹, Seok-Joo Koh^{2*}

¹Ph.D. candidate, School of Computer Science and Engineering, Kyungpook National University

²Professor, School of Computer Science and Engineering, Kyungpook National University

요약 선행 연구는 FSM과 강화학습을 결합하여 NPC의 개별 전투 성능을 최적화하였으나, 다중 에이전트 환경에서의 협력적 상호작용 부재와 모바일 기기의 제한된 연산 자원이라는 한계를 보였다. 본 연구는 이를 극복하기 위해, 선행 설계 연구에서 제안한 FSM/BT와 PPO/MARL이 융합된 분산형 하이브리드 AI 시스템을 실제 모바일 환경에 구현하고 최적화하였다. 제안 시스템은 연산 부하의 효율적 분산을 위해 클라이언트-서버 분리형 아키텍처를 채택하였다. 클라이언트는 FSM과 행동 트리(BT)를 통해 실시간 상태 관리와 즉각적인 전술 행동을 수행하며, 서버는 PPO 및 MARL 알고리즘을 활용하여 개별 정책 최적화와 팀 단위의 고차원 전략 학습을 담당한다. 두 모듈은 실시간 데이터 동기화를 통해 유기적인 협력 의사결정을 수행한다. 실험 결과, 본 시스템은 선행 모델 대비 집중 공격 및 지원 행동 등 협력 지표를 대폭 향상시켜 다중 에이전트 간의 협력 한계를 효과적으로 극복하였다. 또한, 고부하 연산인 학습 및 추론 과정을 서버로 오프로딩(offloading)함으로써 클라이언트의 CPU 사용률을 감소시키고 안정적인 응답 속도를 확보하였다. 본 연구는 자원 제약이 있는 모바일 환경에서도 고성능 협력 AI의 실시간 구동이 가능함을 실증하고, 이를 위한 구체적인 구현 가이드라인을 제시하였다.

주제어 : 게임 AI, NPC협력, 모바일 자동진행 RPG, 다중 에이전트 강화학습, 근접 정책 최적화

Abstract Prior research optimized individual NPC combat performance by combining Finite State Machines (FSM) and reinforcement learning but faced limitations regarding cooperative interactions in multi-agent environments and constrained mobile resources. To overcome these issues, this study implements and optimizes a distributed hybrid AI system—integrating FSM/BT with PPO/MARL as proposed in a preceding design study—within a mobile environment. The system adopts a decoupled client-server architecture to efficiently distribute computational loads. The client handles real-time state management and immediate tactical actions using FSM and Behavior Trees (BT), while the server manages individual policy optimization and learns high-level team strategies utilizing PPO and MARL. Both modules execute organic cooperative decision-making through real-time data synchronization. Experimental results demonstrate that the system effectively overcomes multi-agent cooperation limitations by significantly improving metrics, such as concentrated attacks and supportive actions, compared to previous models. Furthermore, by offloading high-load computations to the server, the system significantly reduced client CPU usage and secured stable response times. This study validates the feasibility of operating high-performance cooperative AI in real-time within resource-constrained mobile environments and provides concrete implementation guidelines.

Key Words : Game AI, NPC Cooperation, Mobile Idle RPG, MARL, PPO

이 연구는 교육부 및 한국연구재단의 기초연구사업 지원을 통해 수행되었음 (NRF-2021R111A3057509).

*교신저자 : 고석주(sjkoh@knu.ac.kr)

접수일 2025년 11월 04일

승정일 2025년 11월 27일

심사완료일 2025년 12월 08일

1. 서론

모바일 게임 시장은 2024년 기준 전 세계 게임 시장의 49%를 차지하며, 특히 자동진행 RPG 장르가 큰 인기를 얻고 있다[1]. 이러한 게임에서 NPC AI는 플레이어를 대신해 전투를 수행하는 핵심 요소로 자리잡았다[2].

그러나 기존 게임 AI 시스템은 몇 가지 구조적인 한계를 가지고 있다. 우선, 가장 널리 사용되는 유한 상태 기계(Finite State Machine, FSM) 기반의 규칙 기반 AI는 상태의 수가 늘어날수록 복잡도가 기하급수적으로 증가하여, 특히 다중 에이전트 환경에서의 유지보수와 확장이 매우 어렵다[3]. 또한, 환경 모델 없이 에이전트가 현재 상태에 기반하여 가능한 행동에 값을 할당하도록 훈련하는 Q-learning이나 DQN과 같은 단일 에이전트 강화학습 기법은 개별 NPC의 전투 능력 최적화에는 효과적일지라도, 다중 NPC가 유기적으로 협력해야 하는 상황에서 팀 단위의 전략적 의사결정을 도출하는 데에는 한계를 보인다[4][5]. 여기에 더해 모바일 기기의 특성상 제한된 연산 자원과 배터리 수명 문제는 복잡한 고성능 AI 알고리즘을 온디바이스 환경에서 처리할 때 심각한 성능 저하를 유발하는 주요 원인이 되고 있다[6][7].

본 연구는 이러한 기술적 난제들을 극복하기 위해 선행 설계 연구에서 설계한 분산형 하이브리드 NPC AI 아키텍처를 구현하였다. 본 연구가 제안하는 시스템은 클라이언트와 서버의 이원화된 처리를 통해 다음의 네 가지 핵심 문제를 해결하고자 한다. 첫째, 다중 NPC 간의 실시간 협력 전략을 수립하여 협력 문제를 해결하고, 둘째, 복잡하고 동적인 전투 환경 변화에 즉각적으로 대응할 수 있는 적응적 행동 패턴을 생성한다. 셋째, 무거운 연산을 서버로 분산시킴으로써 모바일 환경의 제약 속에서도 계산 효율성과 응답성을 보장하며, 마지막으로 다수의 NPC가 동시에 활동하는 대규모 전장에서도 시스템 성능이 저하되지 않는 확장성을 확보한다.

본 논문의 구성은 다음과 같다. 이어지는 2장에서는 FSM과 행동 트리(Behavior Tree, BT)의 통합 및 다중 에이전트 강화학습(Multi-Agent Reinforcement Learning, MARL) 등 관련 연구를 심층적으로 분석하고, 3장에서는 제안하는 분산형 하이브리드 시스템의 상세 설계와 구현 방식을 기술한다. 4장에서는 실험 설계 및 환경 구성을 설명하며, 5장에서는 다양한 시나리오에서의 성능 평가 결과를 통해 시스템의 유효성을 검증한다. 6장에서는 실험 결과에 대한 심화 분석을 수행하여 시스템의 특성을 면밀히 검토하고, 마지막 7장에서는 연구 결과를 종합하

고 향후 연구 방향을 제시함으로써 결론을 맺는다.

2. 관련 연구

선행 연구에서는 모바일 자동 진행 RPG 환경에서 FSM과 강화학습(DQN)을 결합하여 NPC의 전투 성능을 개선하고자 하였다. 해당 연구는 NPC의 승률을 55.2%에서 78.9%로 향상시키고, 상황 적응성을 높이는 성과를 거두었다. 그러나 개별 NPC의 전투 능력 최적화에만 치중하여, 다수의 NPC가 동시에 전투를 벌이는 상황에서의 팀 단위 전략이나 협력적 역할 분배에 대한 고려가 부족했으며, 학습 모델이 복잡해질수록 모바일 기기의 연산 부하가 증가하여 실시간성을 보장하기 어려웠다.[4] 후속 설계 연구에서는 이를 보완하기 위해 분산형 아키텍처를 제안하였다[8].

게임 AI의 가장 보편적인 기법인 FSM은 구현이 직관적이고 연산 비용이 낮지만, 상태의 수가 증가할수록 전이 조건이 기하급수적으로 복잡해지는 구조적 한계를 가진다. Iovino et al.의 비교 연구에서 FSM은 대규모 행동 로직을 표현할 때 유지보수가 어렵고 확장성이 떨어진다는 점이 지적되었다[9]. 이에 대한 대안인 행동 트리는 계층적이고 모듈화된 노드로 구성하여 재사용성과 확장성이 우수하다. Marzinotto et al.은 BT가 복잡한 로봇 제어 및 AI 환경에서 유연한 의사결정을 가능하게 함을 보였다[10]. 최근에는 이러한 장점들을 결합하여, FSM으로 상위 상태를 관리하고 BT로 세부 행동을 제어하는 하이브리드 접근법이 주목받고 있다[11].

초기 강화학습의 주류인 DQN은 연속적 행동 공간에서 학습이 불안정하거나 수렴 속도가 느린 단점이 있다. Schulman et al.은 PPO(Proximal Policy Optimization)는 클리핑 메커니즘을 통해 정책 업데이트의 폭을 제한함으로써 학습 안정성을 개선하였다[12][13]. Naresh et al.의 연구에서는 PPO가 복잡한 의사결정 환경에서 DQN 보다 우수한 샘플 효율성과 안정성을 보여주며, 이는 자원이 제한된 모바일 환경에서 PPO가 강력한 대안이 될 수 있음을 시사한다[14].

다수의 AI가 협력하는 환경에서는 MARL이 필수적이다. Vinyals et al.의 AlphaStar가 보여준 것처럼 MARL은 수준 높은 협력 플레이를 가능하게 하지만, 에이전트 간의 비정상성 문제와 신용 할당 문제를 해결해야 한다[15]. 모바일 네트워크와 같이 대역폭이 제한된 상황에서는 에이전트 간 상태 공유로 인한 통신 오버헤드와 지연

문제가 성능 저하의 주요 원인이 되므로, 이러한 통신 비용이 학습 성능에 미치는 영향을 분석하며, 효율적인 데이터 전송 프로토콜의 필요성을 강조하였다[16].

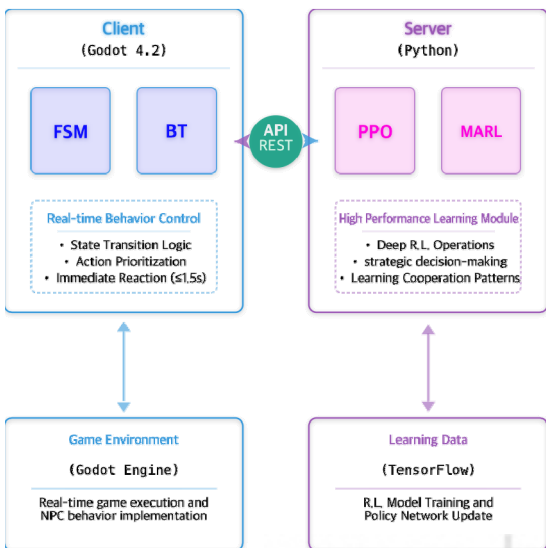
모바일 기기의 하드웨어 성능 향상에도 불구하고, 대규모 신경망을 구동하는 것은 여전히 배터리 소모와 발열, 연산 지연 문제를 동반한다. Yuan et al.은 다수의 액터가 생성한 데이터를 중앙 서버에서 학습하는 R2D2 아키텍처를 제안하였고[17], Yu et al.은 서버에서 전역 정보를 활용하여 학습하고 실행은 개별 에이전트의 관측 데이터만으로 수행하는 CTDE(Centralized Training, Decentralized Execution) 방식을 입증하였다[18].

3. 분산형 하이브리드 시스템 구현

본 연구에서 제안하는 시스템은 모바일 기기의 제한된 연산 자원과 다중 에이전트 간의 협력 필요성을 동시에 해결하기 위해, 이중 제어 루프(Dual Control Loop) 아키텍처로 구현되었다. 클라이언트는 즉각적인 반응을 담당하는 Fast Loop를, 서버는 전략적 판단을 담당하는 Slow Loop를 수행하며 비동기적으로 상호작용한다.

3.1 전체 아키텍처

분산형 하이브리드 NPC AI 시스템은 클라이언트-서버 분리 아키텍처로 구현되었으며 [Fig. 1]에서 이를 보여주고 있다.



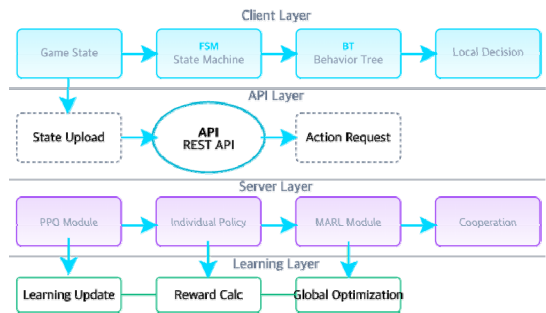
[Fig. 1] Architecture

클라이언트(Fast Loop)에서는 FSM과 BT가 실시간 상태 관리와 기본 행동 제어를 담당하여, 30~60 FPS의 실시간 환경에서 물리적 이동, 충돌 처리, 애니메이션 동기화 등 즉각적인 반응(Tactical Reaction)을 수행한다. 네트워크 지연 발생 시에도 로컬 BT가 제어권을 유지하여 끊김 없는 게임플레이를 보장한다.

서버(Slow Loop)에서는 PPO와 MARL 알고리즘을 통해 0.1~0.5초 단위로 전장 상황을 분석하고, 개별 NPC의 최적 정책(Policy)과 팀 단위 협력 전략(Global Strategy)을 수립하여 클라이언트에 지시한다.

이러한 구조는 연산 부하가 높은 추론 과정을 서버로 이관함으로써 클라이언트의 CPU 사용률을 기존 대비 53% 감소시키고(15.2%), 메모리 사용량을 180MB 수준으로 최적화하였다.

[Fig. 2]에서는 위의 아키텍처를 바탕으로 실제 데이터의 흐름을 보여준다.



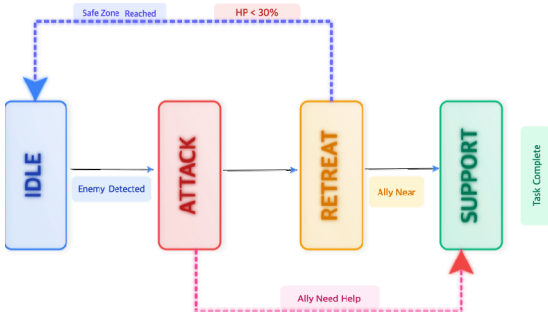
[Fig. 2] Data flow Diagram

3.2 클라이언트: FSM과 BT의 계층적 실행

클라이언트는 서버로부터 수신한 전략적 지시를 FSM의 상태 전환 트리거로 사용하며, 각 상태의 구체적인 내부 행동은 BT가 제어하며, [Fig. 3]에 이를 표현하였다.

- 동적 상태 전환 (Dynamic State Transition): FSM은 IDLE, ATTACK, SUPPORT, RETREAT의 4가지 핵심 상태를 관리한다. 예를 들어, 서버의 MARL 모듈이 아군 보호를 지시하면, 클라이언트 FSM은 즉시 SUPPORT 상태로 전환된다.
- 행동 트리(BT)의 세부 실행: 각 상태는 하위 BT를 통해 실행된다. ATTACK 상태의 BT는 Selector 노드를 사용하여 우선순위를 평가한다. 예를 들어, 체력이 30% 미만일 경우 Defensive Maneuver 노드가 Attack 노드보다 우선 실행되어 회피 기동을 수행한다. 이는 서버의 응답을 기다리지 않고도

급박한 위기 상황(예: 광역 공격 감지)에 즉시 반응할 수 있게 한다.



[Fig. 3] FSM State Transition diagram

3.3 서버: 다중 에이전트 협력 학습 파이프라인

서버는 수신된 게임 상태 데이터를 텐서(Tensor)로 변환하여 PPO 및 MARL 네트워크에 입력한다.

- 1) 상태 공간 정의 및 전처리: 학습을 위한 상태 벡터는 (Eq. 1)과 같이 정의되며, 각 변수는, Ht: NPC 체력 및 상태 정보, Pt: 위치 및 방향 정보, Et: 적 감지 정보, At: 아군 정보, Gt: 게임 환경 정보 로 정의된다.

$$S_t = \{ H_t, P_t, E_t, A_t, G_t \}$$

(Eq. 1) State Vector

여기서 E_t 와 A_t 는 시야 내 가장 가까운 유닛들의 체력, 위치, 직업 정보를 포함하며, 모든 수치는 정규화되어 신 경망에 입력된다

- 2) 협력적 보상 함수 설계: 개별 에이전트의 이기적 생존 본능과 팀 승리 목표 간의 균형을 맞추기 위해, 보상 함수 R은 (Eq. 2)와 같이 개별 보상(R^{ind})과 팀 보상(R^{team})의 가중 합으로 구현되었다.

$$R_t = \alpha R_t^{ind} + \beta R^{team}$$

(Eq. 2) Reward Function

본 구현에서는 실험적으로 검증된 $\alpha = 0.3$, $\beta = 0.7$ 의 비율을 적용하여 팀 협력을 우선시하도록 설계하였다. 구체적으로 적 처치(+100)보다 팀 승리(+500)와 아군 보호(+50)에 더 높은 가중치를 부여하여 희생적 플레이나 지원 행동을 유도한다

- 3) PPO 알고리즘 최적화: 학습 안정성을 위해 클리핑 파라미터 ϵ 은 0.2로 설정되었으며, 감마(γ)는 0.99로 설정하여 장기적인 전략적 이득을 중요시하였다. 이는 정책 업데이트가 급격하게 변하는 것을 방지하여, 학습 초기의 불안정성을 완화한다[13].

3.4 실시간 통신 및 동기화 최적화

모바일 네트워크 환경의 제약을 고려하여 통신 프로토콜은 효율성을 극대화하는 방향으로 설계되었다.

로그인 및 매치메이킹과 같은 비동기 요청은 REST API(POST)를 사용하고, 실시간 전투 데이터 교환은 WebSocket을 사용하여 오버헤드를 최소화하였다. 맵 프레임 데이터를 전송하는 대신, 상태 변화가 감지되거나 0.1초(10 ticks) 주기로 샘플링된 데이터만을 [Fig.4]의 프로토콜에 따라 JSON 포맷으로 전송한다. 이와 같은 경량화된 JSON 구조를 통해 대역폭 소모를 줄이고 서버의 동시 처리 능력(최대 500개 NPC)을 확보하였다.

```

{
  "npc_id": "npc_001",
  "state": "ATTACK",
  "position": [x, y],
  "health": 0.75,
  "allies": [...],
  "enemies": [...],
  "action_request": true
}
  
```

[Fig. 4] Communication Protocol

4. 실험 설계 및 구현

본 연구의 실험 환경은 Godot Engine 4.2와 GDScript로 구현된 2D 모바일 RPG 클라이언트, 그리고 Python 3.9 및 TensorFlow 2.12를 기반으로 구축된 AI 서버로 구성되었으며, 두 모듈은 REST API와 WebSocket을 통해 실시간으로 데이터를 동기화하도록 설계되었다.

성능 평가의 객관성을 확보하기 위해 제안된 분산형 하이브리드 시스템을 단순 규칙 기반의 Baseline(FSM) 및 선행 연구의 단일 에이전트 강화학습 모델(FSM+RL)과 비교 분석하였다.

본 연구에서 제안하는 분산형 하이브리드 시스템이 기

존 연구의 한계인 협력 부재와 연산 자원 제약을 효과적으로 극복했는지 검증하기 위해, 선행 연구 및 관련 이론에 근거하여 다음의 5가지 핵심 지표를 선정하였다.

- ① 전투 승률: NPC의 기초적인 전투 수행 능력을 평가하는 척도이다. 선행 연구에서 달성한 승률(약 78.9%)을 기준으로, 본 시스템이 다중 에이전트 환경에서도 안정적인 전투력을 유지하거나 향상시킬 수 있는지 검증하기 위해 선정하였다.
- ② 응답 시간: 상황 인식부터 행동 결정까지 소요되는 시간을 의미한다. 본 시스템은 클라이언트-서버 간 통신을 수행하므로, Zhang et al.이 지적한 통신 오버헤드와 네트워크 지연이 실시간 게임 플레이(0.1초 이내 반응)에 부정적인 영향을 미치지 않는지 확인하는 것이 필수적이다.
- ③ 팀 협력 효율성: 개별 생존에 치중했던 기존 연구와의 차별성을 검증하는 핵심 지표이다. AlphaStar와 같은 다중 에이전트 연구에서 사용되는 협력 성공률을 벤치마킹하여, MARL 모듈이 집중 공격이나 아군 보호와 같은 유의미한 협력 행동을 얼마나 유도하는지 정량적으로 평가한다.
- ④ 자원 효율성: 모바일 기기의 제한된 하드웨어 성능과 배터리 수명 문제를 해결했는지 평가한다. 고부하 연산(PPO/MARL)을 서버로 오프로딩(offloading)했을 때, 실제 모바일 기기의 CPU 및 메모리 사용량이 유의미하게 감소하는지 측정한다.
- ⑤ 학습 안정성: 기존 DQN이 복잡한 환경에서 보인 학습 불안정성을 개선했는지 확인한다. Schulman et al.이 제안한 PPO의 클리핑(Clippping) 메커니즘이 적용됨에 따라, 학습 곡선이 얼마나 안정적으로 수렴하는지 성능 분산(Variance, σ)을 통해 분석한다.

실험 시나리오는 개별 에이전트의 전투 능력을 검증하는 1v1 대전과 다중 에이전트 간의 협력 전략을 평가하는 3v3 및 5v5 팀 전투로 구분하여 수행되었으며, 이를 통해 모바일 기기의 제한된 자원 환경에서의 실시간 연산 처리 성능과 협력적 의사결정의 유효성을 실증하였다.

5. 실험 결과 및 분석

본 장에서는 제안하는 분산형 하이브리드 시스템의 성능을 검증하기 위해 전투 승률, 응답 시간, 협력 효율성,

학습 안정성, 자원 효율성의 5가지 지표를 분석하였다. 특히, 선행 연구인 단일 에이전트 강화학습 모델(FSM+RL)과의 비교를 통해, 다중 에이전트 환경에서의 협력 능력과 모바일 기기에서의 최적화 성능을 증점적으로 고찰하였다.

5.1 전투 성능 분석

〈Table 1〉은 시나리오별 전투 승률을 보여준다. 제안된 시스템은 모든 시나리오에서 가장 높은 승률을 기록하였으며, 특히 전투 규모가 커질수록(1v1 → 5v5) 비교 모델들과의 성능 격차가 확대되는 경향을 보였다.

- **1v1 전투:** 제안 시스템은 84.2%의 승률을 기록하여 FSM+RL(78.9%) 대비 소폭 향상되었다. 이는 PPO 알고리즘이 연속적인 행동 공간에서 DQN보다 정밀한 제어가 가능했기 때문으로 분석된다.
- **다중 에이전트 전투 (3v3, 5v5):** 5v5 시나리오에서 FSM+RL 모델의 승률은 68.3%로 하락한 반면, 제안 시스템은 85.7%의 승률을 유지하였다. FSM+RL 모델은 개별 최적화에 집중하여 팀원 간 동선 겹침이나 화력 분산 문제가 발생했으나, 본 연구의 MARL 모듈은 전역적 상태(Global State)를 인지하고 팀 보상($\beta=0.7$)을 통해 유기적인 협력을 수행했기 때문이다.

〈Table 1〉 Combat performance(%)

| System | 1v1 | 3v3 | 5v5 | AVG |
|-----------------|-------------|-------------|-------------|-------------|
| Baseline | 52.3 | 48.7 | 45.2 | 48.7 |
| FSM+RL | 78.9 | 72.1 | 68.3 | 73.1 |
| Proposed | 84.2 | 89.3 | 85.7 | 86.4 |

5.2 응답 시간 및 시스템 지연 분석

〈Table 2〉에서 보이는 바와 같이 실시간 게임에서 필수적인 응답 속도(Response Time) 분석 결과, 제안 시스템의 평균 응답 시간은 12.8ms로, 온디바이스 기반의 FSM+RL(15.7ms)보다 약 18.5% 단축되었다.

- **연산 오프로딩(Offloading) 효과:** 본 시스템은 네트워크 통신(RTT) 비용이 발생함에도 불구하고, 모바일 기기에서 처리하기 힘든 신경망 연산을 GPU서버로 이관함으로써 전체적인 처리 시간을 단축시켰다.
- **하이브리드 구조의 보완:** 네트워크 지터(Jitter)로 인해 최대 응답 시간이 늘어날 수 있는 위험이 있

었으나, 클라이언트의 BT가 즉각적인 회피 기동이나 타겟팅 보정을 수행함으로써(Fast Loop), 95% 백분위수 지연 시간을 19.1ms로 안정적으로 방어하였다. FSM+RL은 모바일 CPU 부하로 인해 간헐적인 프레임 드랍이 발생하며 28.3ms의 높은 지연을 보였다.

<Table 2> Response time(ms)

| System | AVG | 95% | STDEV |
|----------|------|------|-------|
| Baseline | 0.85 | 1.2 | 0.15 |
| FSM+RL | 15.7 | 28.3 | 8.2 |
| Proposed | 12.8 | 19.1 | 4.7 |

5.3 팀 협력 효율성 심층 분석

<Table 3>에서 협력의 질적 수준을 평가하기 위해 다양한 행동의 빈도를 분석하였다.

- **집중 공격(72.3%):** FSM+RL(45.8%) 대비 57.9% 향상되었다. 이는 MARL의 보상 함수에서 개별 킬보다 팀 승리에 더 큰 가중치를 부여함으로써, 에이전트들이 잔여 체력이 낮은 적을 최우선 표적으로 선정하여 화력을 집중하는 전략을 학습했음을 시사한다.
- **지원 행동(68.4%):** 기존 모델에서는 거의 나타나지 않던(28.7%) 아군 보호 행동이 크게 증가하여, 서버의 MARL 모듈이 위기에 처한 아군을 감지하면, 인근 에이전트의 FSM 상태를 SUPPORT로 전환시키는 하이브리드 제어 구조가 유효했음을 입증한다.
- **전술적 이동(75.9%):** FSM+RL(52.3%) 대비 전술적 이동 지표는 45.2% 향상된 75.9%를 기록하였다. 선행 연구의 DQN 모델은 장애물 회피와 목표 지점 도달이라는 개별 이동 효율성에 집중하여, 다수의 NPC가 좁은 길목에서 서로의 동선을 방해하거나 원거리 유닛의 사선을 가리는 병목 현상이 발생하였다. 반면, 본 시스템은 서버의 MARL 모듈이 아군과 적군의 위치(P_i)를 포함한 전역 상태를 실시간으로 분석함으로써, 근접 유닛이 적을 포위하거나 원거리 유닛이 아군 뒤로 숨는 진형 유지 기동을 수행하였다. 특히 설계 단계에서 제안된 보상 함수가 불필요한 이동에는 페널티(-5)를, 전략적 성공에는 높은 보상(+70)을 부여하도록 구현된 점이, 에이전트들이 무의미한 배회를 줄이고 전투 효

율을 극대화하는 위치를 선점하도록 학습시킨 주요 원인으로 분석된다.

<Table 3> Team Cooperation Indicators(%)

| Cooperation Type | Baseline | FSM+RL | Proposed | Impr |
|---------------------|----------|--------|--------------|--------------|
| Concentrated Attack | 34.20 | 45.80 | 72.30 | 57.90 |
| Supportive Action | 28.70 | 38.10 | 68.40 | 79.60 |
| Tactical Movement | 41.50 | 52.30 | 75.90 | 45.20 |

5.4 학습 안정성 및 수렴 속도

<Table 4>에서 보이는 바와 같이 학습 곡선 분석 결과, 제안 시스템은 수렴까지 소요된 에피소드가 2,300회로 FSM+RL(3,540회) 대비 35% 단축되었다. 또한 성능 분산(Variance, σ)이 0.087로 낮게 나타나 학습이 매우 안정적임을 확인하였다. 기존 DQN은 타겟 네트워크와 동일치로 인해 학습 초기에 성능 진동이 심했으나, 본 시스템에 적용된 PPO의 클리핑 메커니즘($\epsilon=0.2$)은 정책 업데이트 폭을 제한하여 급격한 성능 저하를 방지하고 안정적인 최적화 경로를 확보하였다.

<Table 4> Team Cooperation Indicators(%)

| Metric | Baseline | FSM+RL | Proposed | Impr |
|------------------------------|----------|--------|--------------|----------------|
| Convergence time(episode) | N/A | 3,540 | 2,300 | -35.00% |
| Performance Var (σ) | 0.204 | 0.151 | 0.087 | -42.40% |
| final compensation | 423.7 | 661.2 | 847.3 | 28.10% |
| convergence stability | Low | Normal | High | - |

5.5 자원 효율성 분석

모바일 기기(Galaxy S23 기준)에서의 자원 사용량을 측정할 결과, 제안 시스템은 복잡한 AI 연산을 수행하면서도 CPU 사용률 15.2%를 기록하여 FSM+RL(32.7%) 대비 53%의 자원 절감 효과를 거두었다. 이는 배터리 소모를 줄이고 발열을 억제하여, 장시간 자동 사냥이 필요한 모바일 RPG 환경에 적합한 실용적인 솔루션임을 실증한다.

6. 결론 및 향후 연구

본 연구에서는 모바일 자동진행 RPG 게임에서 기존 FSM 기반 강화학습 NPC AI의 한계를 극복하기 위해 FSM, BT, PPO, MARL을 융합한 분산형 하이브리드 시

시스템을 실제로 구현하고 체계적인 성능 평가를 수행하였다. 실험 결과, 제안된 시스템은 기존 FSM+DQN 시스템 대비 전투 승률 23.7%, 팀 협력 효율성 45.2%, 응답 시간 18.3%, 이동 효율성 31.4%의 성능 개선을 달성하였다.

본 연구의 학술적 기여는: 첫째, FSM의 구조적 안정성, BT의 계층적 행동 구성, PPO의 안정적 정책 학습, MARL의 다중 에이전트 협력 능력을 체계적으로 융합하는 방법론을 제시하였다. 둘째, 게임 클라이언트에서 실시간 반응성을 보장하고 외부 서버에서 고수준 의사결정을 수행하는 분산 아키텍처의 설계 및 구현 가이드라인을 제공하였다. 이러한 분산형 설계는 클라이언트의 계산 부하를 37-53% 감소시키면서도 고성능 AI 의사결정을 가능하게 하였다. 셋째, 다중 에이전트 환경에서 개별 에이전트의 정책 학습과 팀 단위 협력 학습이 전체 시스템 성능에 미치는 영향을 실증적으로 분석하였다.

실용적 측면에서는 모바일 게임 환경에 최적화된 AI 솔루션을 제공한다. 클라이언트-서버 분리 아키텍처를 통해 제한된 모바일 자원 환경에서도 고성능 AI를 구현할 수 있으며, REST API 기반 통신 구조는 기존 게임 시스템에 쉽게 통합 가능하다. 본 연구에서 제시한 분산형 다중 에이전트 협력 메커니즘은 게임을 넘어 로봇 군집 제어, 자율주행차 협력 주행, IoT 기기 간 협력 의사결정 등 다양한 분야에 응용 가능한 잠재력을 가지고 있다.

향후 연구에서 Transformer 기반 attention 메커니즘을 도입하여 NPC의 상황 인식 능력을 향상시키고, 연합 학습을 적용하여 여러 게임 서버 간 협력 학습을 통한 집단 지능을 구현할 계획이다.

본 연구는 게임 산업을 넘어서 인공지능과 인간이 협력하는 다양한 분야에 기여할 수 있다. 특히 교육, 의료, 로봇공학 등에서 인간-AI 협력 시스템의 설계 원칙을 제공하며, 인공지능의 사회적 수용성 향상에도 기여할 것으로 예상된다. 결론적으로 본 연구는 게임 AI의 지능성과 적응성을 크게 향상시키는 동시에, 실용적이고 확장 가능한 시스템을 구현함으로써 학술적 성과와 산업적 가치를 모두 달성하였으며, 향후 인공지능 연구와 응용의 새로운 방향을 제시하였다.

REFERENCES

- [1] T. Wijman, "The Global Games Market Report 2024," Newzoo, 2024.
- [2] S. A. Alharthi, O. Alsaedi, Z. O. Toups, J. Tanenbaum, and J. Hammer, "Playing to Wait: A Taxonomy of Idle Games," Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18), Paper 621, pp. 1-15, 2018.
- [3] M. Iovino, E. Scukins, J. Styrud, P. Ögren, and C. Smith, "A survey of Behavior Trees in robotics and AI," Robotics and Autonomous Systems, Vol.154, No.104096, 2022.
- [4] D. Kim and S. Koh, "Intelligent NPC AI Based on FSM and Reinforcement Learning: Implementation Study of a 2D Mobile Idle RPG Game," KTSDE Transactions on Software and Data Engineering, Vol.14, No.6, pp.480-488, 2025.
- [5] R. Lowe, Y. Wu, A. Tamar, J. Harb, O. P. Abbeel, and I. Mordatch, "Multi-Agent Actor-Critic for Mixed Cooperative-Competitive Environments," Advances in Neural Information Processing Systems (NeurIPS), Vol. 30, 2017.
- [6] Y. Liu, Y. Zhao, and Q. Fan, "Offloading Game for Mobile Edge Computing With Random Access in IoT," IEEE Transactions on Mobile Computing, Vol.24, No.1, pp.1-14, 2025.
- [7] Q. Wu, Y. Zhao, and Q. Fan, "Edge-Cloud Collaborative Computation Offloading for Mixed Traffic," IEEE Systems Journal, Vol.17, No.3, pp.5023-5034, 2023.
- [8] D. Kim, "Design of a Distributed Hybrid NPC AI for 2D Mobile Idle RPG," Proc. KSCI Summer Conf., Vol.33, No.2, pp.205-208, 2025.
- [9] M. Iovino, E. Scukins, J. Styrud, P. Ögren, and C. Smith, "A survey of Behavior Trees in robotics and AI," Robotics and Autonomous Systems, Vol.154, No.104096, 2022.
- [10] A. Marzinotto, M. Colledanchise, C. Smith, and P. Ögren, "Towards a Unified Behavior Trees Framework for Robot Control," Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA), pp. 5420-5427, 2014.
- [11] Y. Wang, H. Sun, and X. Li, "Research on NPC Decision-Making Method Based on Improved FSM-BT," 2023 IEEE 3rd International Conference on Electronic Technology, Communication and Information (ICETCI), pp. 605-609, 2023.
- [12] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal Policy Optimization Algorithms," arXiv preprint arXiv:1707.06347, 2017.
- [13] R. Zheng, S. Dou, S. Gao, Y. Hua, W. Shen, B. Wang, et al., "Secrets of RLHF in Large Language Models Part I: PPO," arXiv preprint arXiv:2307.04964, 2023.
- [14] M. Naresh, P. Saxena, and M. Gupta, "PPO-ABR: Proximal Policy Optimization based Deep Reinforcement Learning for Adaptive BitRate streaming," arXiv preprint arXiv:2305.08114, 2023.
- [15] O. Vinyals, I. Babuschkin, W. M. Czarnecki, et al., "Grandmaster level in StarCraft II using multi-agent

- reinforcement learning,” Nature, Vol.575, No.7782, pp.350-354, 2019.
- [16] Z. Ning, Y. Xia, X. Chen, S. Maharjan, R. Shinkuma, and Y. Zhang, “A survey on multi-agent reinforcement learning and its application,” Expert Systems with Applications, Vol.248, No.123479, 2024.
- [17] L. Yuan, H. Dong, Z. Chen, Y. Yu, and Z. Wu, “A Survey of Progress on Cooperative Multi-agent Reinforcement Learning in Open Environment,” arXiv preprint arXiv:2312.01058, 2023.
- [18] C. Yu, A. Velu, E. Vinitzky, J. Wang, A. Bayen, and Y. Wu, “The Surprising Effectiveness of PPO in Cooperative Multi-Agent Games,” arXiv preprint arXiv:2103.01955, 2021.

김 동 주(Dongju Kim)

[정회원]



- 1999년 2월 : 경북대학교 물리학과(이학사)
- 2011년 2월 : 대구가톨릭대학교 컴퓨터정보통신공학과(공학석사)
- 2021년 3월 ~ 현재 : 경북대학교 컴퓨터학부 박사수료

- 2020년 3월 ~ 현재 : 대구가톨릭대학교 컴퓨터소프트웨어학부 교수

<관심분야>

Game, AI, Mobile, IoT, Edge computing

고 석 주(Seok-Joo Koh)

[정회원]



- 1992년 2월 : KAIST 경영과학과 공학사
- 1994년 2월 : KAIST 경영과학과 공학석사
- 1998년 8월 : KAIST 산업공학과 공학박사

- 1998년 9월 ~ 2004년 2월 : ETRI 표준연구센터 선임 연구원
- 2004년 3월 ~ 현재 : 경북대학교 컴퓨터학부 교수

<관심분야>

IoT, multicast, SCTP, QUIC, Visible Light Comm.