

실시간 공항 운영을 위한 PPO 기반 자원 배정 최적화 시스템 설계

김준규, 천민범, 안광모

주식회사 케파스

jkkim@kephas.kr, mbcheon@kephas.kr, gmahn@kephas.kr

Design of a PPO-Based Optimization System for Real-Time Airport Resource Allocation

Kim Jun Kyu, Cheon Min Beom, Ahn Kwang Mo

Kephas.

요약

전 세계 항공 운송 수요의 지속적 증가에 따라 공항 운영에서 자원의 효율적이고 실시간적인 배정은 핵심 과제로 부상하고 있다. 기존 혼합정수선형계획법(Mixed-Integer Linear Programming, MILP)이나 메타휴리스틱 기반 기법은 환경 변화에 대한 유연성이 낮고, 실시간 운영 상황에서의 적용에 한계가 존재한다. 본 연구에서는 수학적 모델링과 정책 기반 강화학습(Proximal Policy Optimization, PPO)을 결합한 실시간 자원 배정 시스템을 제안한다. 제안 시스템은 공항 데이터 입력과 API 연동을 통해 자원 배정 요청을 수신하고, 전처리된 데이터를 기반으로 MILP 모델을 구성한 후 사전 학습된 PPO 정책을 호출하여 실시간 배정 결과를 산출한다. 인천국제공항의 실제 항공편 데이터를 활용한 실험에서 배정률 100%, 1,000편 기준 평균 실행시간 258초, 제약 위반률 0.27%를 기록하였다. 또한 동일 조건에서 수행한 기존 연구의 A3C 기반 REGAPS 모델이 평균 실행시간 185초를 보인 것과 비교해, 제안 방식은 평균 15.1초로 현저히 빠른 응답 속도를 달성하였다. 이를 통해 본 시스템이 실시간 운영 환경에서 기존 접근법 대비 우수한 성능을 제공함을 입증하였으며, 향후 다양한 운영 시나리오와 다중 자원 배정 문제로의 확장 가능성을 제시한다.

I. 서론

전 세계 항공 운송 수요는 지속적으로 증가하고 있으며, 이에 따라 공항 자원 배정의 효율성과 실시간성은 운영 안정성을 결정짓는 핵심 요소로 부상하고 있다. 국제항공운송협회(IATA)는 2037년까지 전 세계 항공 여객 수가 약 82억 명에 이를 것으로 전망하며[1], 게이트, 주기장, 체크인 카운터, 수하물 처리 시스템 등 핵심 인프라 자원 수요 역시 급격히 증가할 것으로 예상된다. 특히 항공편 지연, 기종 변경, 긴급 착륙 등 예외 상황이 빈번한 실제 운영 환경에서는 단순 계획 수립을 넘어 실시간 유연한 자원 배정 체계가 필수적이다[2].

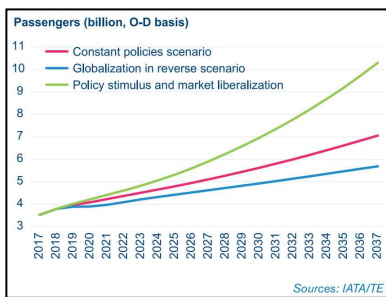


그림 1. IATA 시나리오별 항공 여객 수 증가 전망(2017-2037)

공항 자원 배정 문제는 시간 충돌, 기체 수용 조건, 자원 선호도 등 다양한 제약을 동시에 고려해야 하는 조합 최적화 문제이며, 일반적으로 NP-Hard에 속한다. 기존 연구들은 혼합정수선형계획법(Mixed-Integer Linear Programming, MILP), 유전 알고리즘(Genetic Algorithm, GA) 등 수치 기반 또는 메타휴리스틱 기법을 적용해 문제를 해결해왔다[3]. 그러나 이러한 접근은 대부분 정적 계획 기반으로 구성되어 있어 환경 변화 시 전체 계획을 반복적으로 재계산해야 하는 한계를 가진다. 특히 실시간성이 중요한 운영 환경에서는 계산 복잡도 증가, 수작업 개입, 낮은 자동화 수준 등으로 인해 실제 적용이 어렵다.

이러한 한계를 해결하기 위해서는 단순한 알고리즘 개선을 넘어 문제 정

의 -정책 실행- 결과 반환 전 과정을 통합적으로 처리할 수 있는 구조적 설계가 필요하다. 본 연구에서는 자원 배정 문제를 수학적으로 정의하고, **정책기반강화학습(Proximal Policy Optimization, PPO)**을 결합한 실시간 자원 배정 시스템을 설계하였다. 사용자의 요청을 입력받아 최적화 모델을 구성하고, 사전 학습된 정책을 호출하여 최적의 배정 결과를 실시간으로 도출하는 전 과정을 하나의 구조로 통합하였다. PPO는 정책 업데이트 안정성과 빠른 수렴 속도에서 강점을 가지며, 보상 구조를 통해 다양한 제약을 내재화할 수 있어 수리 모델과의 결합에도 효과적이다[4]. 제안 시스템은 입력, 최적화, 정책 호출, 결과 반환까지의 전 과정을 실시간으로 자동화함으로써, 공항 운영 환경의 실질적 문제를 해결할 수 있는 설계 구조를 제시한다.

II. 본론

본 연구에서는 실시간성과 자동화를 동시에 충족하기 위해, 공항 자원 배정 시스템을 모듈화된 구조로 설계하였다. 제안하는 시스템은 공항 데이터 입력부터 최적화 모델 생성, 정책 기반 배정 실행, 결과 반환까지의 전 과정을 통합된 프로세스로 구성하며, 그림 2와 같이 단계별 모듈이 유기적으로 연동되어 실시간 운영 환경에 대응할 수 있도록 설계하였다.

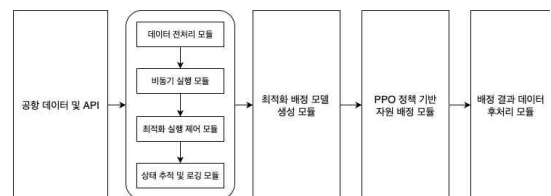


그림 2. PPO 기반 실시간 공항 자원 배정 시스템 구조도

시스템은 공항 데이터 입력 및 API를 통해 자원 배정 요청과 관련 데이터를 직접 입력받으며, 이를 기반으로 내부 처리 흐름을 시작한다. 이후 **데이터전처리모듈**이 입력된 데이터를 정제한다. 이 모듈은 입력 정보를 시계열 기반으로 재구성하고, 최적화 모델에 적용할 수 있도록 필요한 형식

과 구조로 가공한다. 다음으로 **비동기실행모듈**이 각 배정 작업을 독립적으로 병렬 실행하며, **최적화실행제어모듈**은 요청 상태를 관리하고, 필요 시 실행을 중단하거나 재요청할 수 있도록 제어한다. **상태추적및로깅모듈**은 각 단계의 진행 상황을 실시간으로 모니터링하고 기록하여 전체 흐름의 안정성과 투명성을 확보한다. 이렇게 준비된 데이터는 **최적화모델 생성모듈**을 통해 수학적으로 문제화된다. 이때 자원 충돌, 수용 조건, 선호도 등 다양한 제약을 포함한 배정 모델을 생성하고, 이어서 **정책기반 자원배정모듈**이 사전에 학습된 정책을 호출해 최적의 배정 결과를 도출한다. 마지막으로, **결과후처리모듈**이 배정 결과를 정제하고 응답 형태로 구성하여 공항 운영 시스템에 반환한다. 제안하는 시스템은 각 모듈이 독립적으로 구성되어 있어 유지보수가 용이하며, API 기반으로 연동되어 다양한 운영 환경에 유연하게 적용 가능하다. 특히 데이터 입력부터 결과 반환까지의 전 과정을 실시간으로 자동 처리함으로써, 기존의 정적 계획 방식보다 빠르고 유연한 대응이 가능하다. 제안하는 시스템의 핵심은 자원 배정 문제를 강화학습 환경으로 변환하고, 사전에 학습된 정책을 활용해 실시간으로 배정하는 구조다. 이를 위해 상태-행동-보상 구조의 마르코프 결정 과정(Markov Decision Process, MDP)을 기반으로 학습 환경을 정의하고, 정책 기반 강화학습 기법인 PPO로 자원 배정 정책을 학습하였다. 상태(state)는 항공편의 도착 시간, 기체 종류, 항공사 코드, 사용 가능한 자원 목록 등으로 구성되며, 행동(action)은 특정 항공편에 하나의 자원을 선택하는 방식이다. 보상(reward)은 유효한 배정 여부, 제약조건 충족, 선호도 자원 배정 등을 반영해 설계하여 다양한 제약을 내재화하도록 했다. PPO는 학습 안정성과 수렴 속도에서 우수하며, 실시간 추론 환경에 적합하다. 정책 업데이트 시 불안정을 방지하기 위해 클리핑(clipping) 기법을 도입해 이전 정책과 크게 벗어나지 않는 범위에서 점진적으로 성능을 향상시킨다. 아래 수식은 새로운 정책과 이전 정책의 비율에 기반해 어드밴티지 값을 최소화하는 PPO의 핵심 손실 함수다.

$$L^{CLIP}(\theta) = E_t [\min(r_t(\theta) A_t, clip(r_t(\theta), 1 - \epsilon, 1 + \epsilon) A_t)] \quad (1)$$

수식 (1)에서 $r_t(\theta)$ 는 새로운 정책과 이전 정책의 비율, A_t 는 어드밴티지 추정값, ϵ 은 클리핑 계수로 설정되며, 이 수식을 통해 과도한 정책 업데이트를 방지하고 안정적인 학습을 유도한다. 제안하는 학습 구조는 반복적 시뮬레이션을 통해 다양한 상태-행동 조합을 경험하고, 자원 배정 환경의 제약 조건을 보상 함수에 자연스럽게 녹여냄으로써 현실성과 효율성을 동시에 확보한다. 학습이 완료된 정책은 실제 운영 시스템 내에서 호출되어, 입력 정보에 기반한 자원 배정을 실시간으로 수행하게 된다.

실험에는 인천국제공항의 실제 항공편 데이터를 활용하였다. 국토교통부 항공정보포털(UBIKAIS)[5]에서 2025년 3월 21일부터 25일까지 수집한 5,000편의 운항 데이터를 기반으로, 총 228개의 주기장을 포함하는 환경을 구성하였다. 자원 배정 모델은 MILP 기반으로 수립되었으며, 시간 충돌 방지, 단일 자원 배정, 수용성, 선호도 등 네 가지 제약 조건을 함께 고려하였다.

정책 학습은 PPO 알고리즘을 기반으로 총 500 에피소드에 걸쳐 수행되었으며, 학습률 0.001, 할인율 0.99, GAE λ 값 0.95, 클리핑 계수 0.2 등의 하이퍼파라미터를 사용하였다. 실험은 5,000편의 운항 데이터를 1,000편 단위로 구분하여 총 5회 수행하였으며, 주요 평가지표는 배정률, 실행 시간, 제약 위반률로 설정하였다. 그 결과, 모든 실험에서 배정률 100%, 1,000편 기준 평균 실행 시간 258초, 제약 위반률 평균 0.27%를 기록하였다. 이는 제안한 정책이 제약 조건을 효과적으로 반영하면서도 실시간 대응 능력을 유지함을 보여준다.

특히 기존 연구[6]에서는 A3C 기반의 REGAPS 모델을 활용하여 124편의

항공편과 14개의 게이트로 구성된 실제 공항 데이터를 대상으로 게이트 배정 문제를 해결하였으나 평균 실행 시간이 약 185초로 측정되어 실시간 운영에 적용하기에는 한계가 있었다. 반면, 본 연구는 동일한 조건에서 PPO 정책을 적용하여 평균 실행시간을 15.1초로 단축하였으며, 기존 연구 대비 현저히 빠른 응답 속도를 달성하였다. 이를 통해 제안한 방식이 실시간 운영 환경에서 우수한 성능을 제공함을 입증하였다. 이는 단일 실행만으로 배정 결정을 수행할 수 있으며, 추론 속도와 실시간성 측면에서 강화된 자동화 구조를 제공한다는 점에서 차별성을 가진다.



그림 3. PPO를 활용한 항공편-주기장 실시간 배정 결과

III. 결론

본 연구는 공항 자원 배정 문제에 대해 정책 기반 강화학습(PPO)을 적용한 실시간 자동화 시스템을 설계·구현하였다. 제안 시스템은 사용자 입력부터 최적화, 정책 호출, 결과 반환까지의 전 과정을 통합 구조로 구성하여 기존 정적 계획 방식의 한계를 극복하였다. 인천국제공항 실제 데이터를 활용한 실험에서 제안 방식은 배정률 100%, 낮은 제약 위반률(0.27%)을 달성하며 실시간 대응이 가능함을 입증하였다. 또한 기존 A3C 기반 기법 대비 추론 속도와 자동화 수준에서 우수한 성능을 보였다. 본 연구는 향후 다양한 운영 시나리오와 예외 상황을 반영한 정책 고도화, 다른 강화 학습 알고리즘과의 비교, 다중 자원 동시 배정 문제로의 확장 등 지능형 공항 운영 시스템 발전을 위한 기반을 제시한다.

ACKNOWLEDGMENT

이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원 SW컴퓨팅산업원천기술개발사업의 지원을 받아 수행된 연구임(RS-2024-00398050).

참 고 문 헌

- [1] International Air Transport Association (IATA), “20-Year Air Passenger Forecast” 2018, (<https://www.iata.org/en/publications/store/20-year-passenger-forecast/>)
- [2] G. S. Daş, F. Gzara, and T. Stützle, “A Review on Airport Gate Assignment Problems: Single versus Multi Objective Approaches,” *Omega*, vol. 92, p. 102146, 2020.
- [3] C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*, Mineola, NY: Dover Publications, 1998.
- [4] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, “Proximal Policy Optimization Algorithms,” *arXiv preprint, arXiv:1707.06347*, 2017.
- [5] UBIKAIS: Korean Air Traffic Information System, (<https://ubikais.fois.go.kr:8030>)
- [6] H. Li, X. Wu, M. Ribeiro, B. Santos, and P. Zheng, “Deep reinforcement learning approach for real-time airport gate assignment,” *Operations Research Perspectives*, vol. 14, Article

100338, 2025.