강화 학습을 이용한 현장 전문가 의견 기반 주문량 결정 모델에 관한 연구

박형빈, 정봉주

연세대학교 산업공학과

hyungbin90@yonsei.ac.kr, bongju@yonsei.ac.kr

2021년 한국SCM학회 춘계컨퍼런스

이 논문은 2019년도 정부(과학기술정보통신부)의 재원으로 한국연구재단 기초연구사업의 지원을 받아 수행된 연구임 (NRF-2019R1F1A1061349)

Contents

1. 연구 배경
2. 기존 문헌 연구
3. 문제 정의
4. 주문량 결정 학습 모델
5. 결론 및 향후 연구 내용

1. 연구 배경

비안정적인 수요와 공급 환경

- 최근 기술의 발전으로 시장의 글로벌화는 제품수명주기를 단축시키고 제조 회사들의 경쟁이 심화됨에 따라 전반적인 고객 수요 불안정성을 증폭시키는 원인이 됨
- 또한, 온라인 시장의 다양한 제품 유통 채널이 발달하면서 고객 수요에 대한 제품 가격 탄력성이 높아져 수요 변동성이 확대
- 다양한 실제 환경에서는 요인들(불량, 장비 이상, 자연 재해 등)은 공급사슬상의 위험으로 인한 공급 손실률 발생
- 공급 손실률로 인한 고객 만족도 하락, 잠재적 고객 이탈, 판매량 감소 등 비용적 손실 발생

➢ **강화 학습을 이용해 시간에 따라 변화하는 고객 수요 분포를 고려한 환경에서 수요와 공급의 불확실성에 대응한 주문량 결정 모델 개발**

1. 연구 배경

인간과 인공지능 에이전트의 조화

- 기존의 주문량 결정 모델은 현장 전문가의 경험, 지식, 정보 등을 정성적인 데이터를 활용하는 것이 어려움
- 수요와 공급의 불확실성에 대한 현장 전문가의 의견을 묻고, 인공지능 에이전트는 피드백을 바탕으로 최적의 주문량을 결정

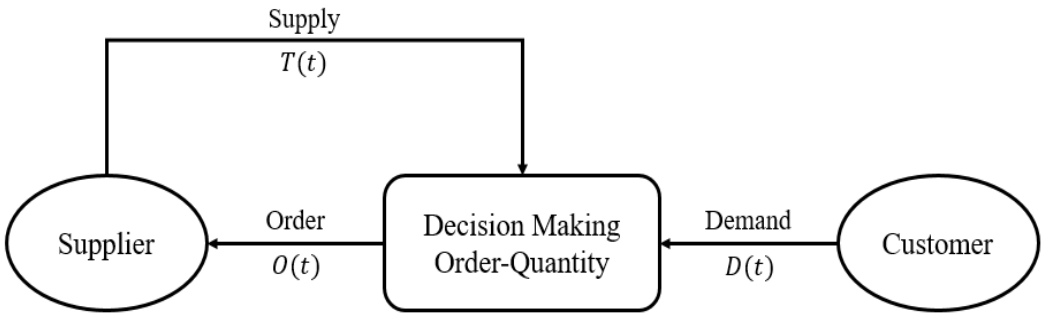
➢ **현장에서 얻을 수 있는 정성적인 현장 전문가의 의견을 반영하여 수요와 공급에서 발생하는 불확실성에 대응하는 주문량 결정 모델 개발**

2. 기존 문헌 연구

구분	저자 (연도)	요약
Reinforcement learning approach for inventory management	Kim et al. (2005)	• 불확실한 수요 환경에 적응하여 주문량을 결정하기 위한 목표 서비스 수준 조정 연구 진행
	Baek et al. (2005)	• JIT(Just-in-time)형태의 공급을 수행하는 가정에서 고객 수요의 불확실성을 고려하여 공급자, 소매자 2단계 공급사슬에서 적응형 재고관리 모델을 강화 학습을 이용하여 접근 • 현 재고수준을 상태로 보고 가장 이들이 되는 주문량과 주문 리드타임을 가진 공급자를 결정
	Kwon (2014)	• 시간에 따라 변화하는 수요 분포 환경에서 주어진 서비스 수준을 보장하는 재고 정책 도출 • 비안정적인 고객 수요 환경에서 사전에 정의된 서비스수준을 보장하기 위해 재고 수준을 관측하고 보상계수를 조정하는 방법 제안
Supply chain uncertainty	Ghotge et al. (2020)	• 에너지 수요 불확실성이 있는 상황에서 전기차 주차장의 최대 전력 수요를 줄이는 전략을 개발하기 위해 불확실성의 정도와 전기차 도착 시기를 고려한 연구 진행
	Zhao et al. (2017)	• 송유관 정체 및 천연가스 가격 변동에 따른 공급 불확실성 연구 • 두 가지 유형의 천연 가스 공급 조건에 미치는 영향을 분석하기 위한 2-stage 혼합 정수 선형 최적화 모델을 개발

- 본 연구에서는 환경과 상호작용 최적의 해를 도출하기 위해 학습을 진행하는 의사결정 기계학습 기법인 강화 학습을 이용해 최적 주문량 결정
- 정석적인 현장 전문가의 의견을 반영한 수요와 공급의 불확실성을 고려하여 주문량을 결정하는 지능형 주문량 결정 모델을 제안

3. 문제 정의



<그림 1> Order quantity decision process

- 정기 발주 형식으로 재고를 보충하는 Retailer의 관점에서 이윤 최대화를 목적으로 최적의 주문량을 결정하는 문제
- Retailer 간의 교차 공급(crossover delivery)와 추후 납품(backorder)는 없다고 가정
- $D(t)$= 주문 발주와 다음 발주 시점까지의 수요, $O(t)$= 해당 주문 시점에 공급자에게 발주하는 주문량, $T(t)$=주문 해당 시점에서 실제 공급되는 공급량

3. 문제 정의

수요와 공급의 불확실성

$$\boldsymbol{D(t) = \eta_d(t) * \varphi(t)}$$

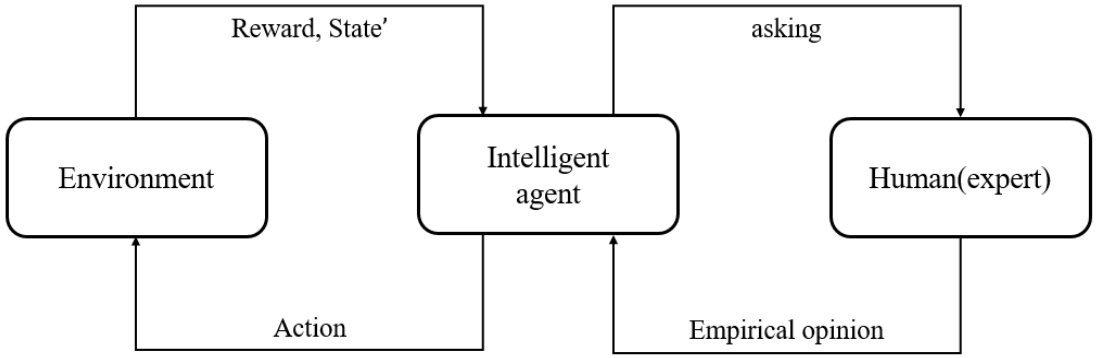
- $\boldsymbol{\eta_d(t)}$**:** 해당 시점에 지역적 시기적인 수요 불확실성 인자로 수요의 불확실성을 증가시키는 고객 수요의 일시적인 변동 반영
- $\boldsymbol{\varphi(t)}$**:** 시간에 따라 평균과 분산이 바뀌는 수요 분포를 따른다고 가정

$$\boldsymbol{T(t) = \eta_s(t) * O(t)}$$

- $\boldsymbol{\eta_s(t)}$**:** 공급량의 불확실성이 일어나지 않는다면 1, 불확실성이 커져 공급 손실이 커진다면 0에 가까워짐, [0, 1]
- $\boldsymbol{O(t)}$**:** t 시점의 주문량

3. 문제 정의

정성적인 정보의 활용

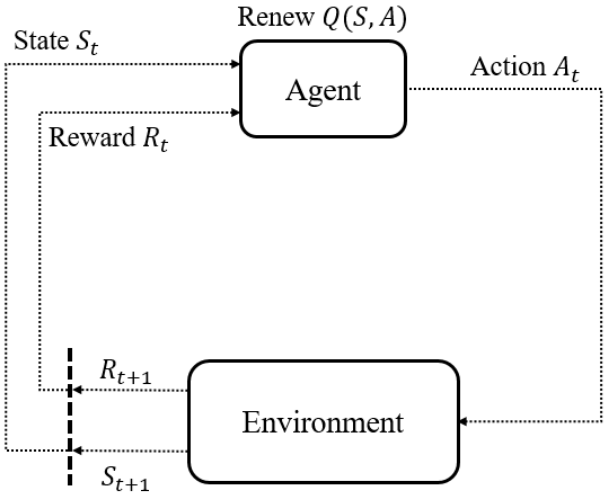


<그림 2> Opinion based learning process

- 과거로부터 축적된 발주 이력과 판매 실적을 활용한 자동 발주 시스템은 해당 매장의 매출에 직접적으로 영향을 주는 요소들을 반영하기 어려움
- 수요나 공급에 직접적으로 영향을 주는 지역적인 정보를 활용하는 연구 필요
- Retail에서의 발주를 담당하는 현장 전문가는 수요와 공급에 영향을 미치는 사건이나 요인에 대한 정성적인 정보 보유
- 주문량 결정을 하는 지능형 에이전트(Intelligent agent)는 현장 담당자(Expert)에게 의견을 묻고 피드백을 받아 현장의 정성적인 정보를 활용
- 현장 담당자의 의견과 현 재고 기반으로 주문량을 결정하는 모델을 강화 학습을 통해 제시

4. 주문량 결정 학습 모델

강화 학습(Reinforcement Learning)



<그림 3> Interaction of agent and environment (action policy learning process)

- State, action, reward로 학습 주체인 에이전트와 환경이 상호 작용하며 학습 진행
- t시점에서 에이전트는 환경의 state, S_t를 확인하고 그에 따른 best action, A_t를 취함
- A_t를 행한 결과로 에이전트는 수리적인 보상인 reward, R_{t+1}을 받고 새로운 state, S_{t+1}로 업데이트
- 단기적으로는 R_{t+1}가 그 시점의 action 선택 기준이 될 수 있지만, 장기적으로는 state들의 reward의 합을 사용하여 action 선택 기준으로 정함

4. 주문량 결정 학습 모델

지능형 주문량 결정 모델(Intelligent order quantity decision model)

- 내구성 있는 제품 재고 기반 모델과 부패하기 쉬운 제품의 수명 기반 모델을 제안
- 과거 데이터에 수요와 공급 불확실성에 대한 현장 전문가의 의견이 조합하여 모든 과정에 반영하여 학습 진행
- 현장 전문가의 정성적인 의견 반영
  - State를 현 재고수준, 전 시점의 공급 손실률, 수요에 대한 전문가의 의견, 공급에 대한 전문가의 의견, 의사 결정 시점으로 정의
  - 공급과 수요에 대한 전문가의 정성적인 의견을 state에 반영함으로써 불확실성에 대해 대응
  - 정성적인 의견의 정확도까지 고려한 지능형 에이전트의 학습을 통해 최적의 해 도출
-

5. 결론 및 향후 연구 내용

< 결론 >

- 본 연구에서는 시간이 지남에 따라 수요의 분포가 바뀌는 비안정적인 환경에서 수요와 공급의 불확실성을 고려한 주문량 결정 강화 학습 모델을 제안
- 수요와 공급의 일시적인 변동에 대한 발주 담당자와 같은 현장 전문가의 정성적인 의견을 활용하여 주문량 결정에 반영하는 연구 수행