



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2018년06월28일
(11) 등록번호 10-1860258
(24) 등록일자 2018년05월15일

- (51) 국제특허분류(Int. Cl.)
G06F 17/50 (2006.01) G06N 99/00 (2010.01)
- (52) CPC특허분류
G06F 17/5009 (2013.01)
G06N 99/005 (2013.01)
- (21) 출원번호 10-2017-0037719
- (22) 출원일자 2017년03월24일
심사청구일자 2017년03월24일
- (56) 선행기술조사문헌
JP2011073883 A
US7117130 B1
“A Markov Model for Inventory Level Optimization in Supply-Chain Management”, Institute for Information Technology -e-Business(pp. 133-144), 2005. 05.
“재 제조시스템의 가용도 분석모델과 평가척도”, 한국정밀공학회지 제27권 2호(pp. 78-85), 2010년 02월

- (73) 특허권자
한밭대학교 산학협력단
대전광역시 유성구 동서대로 125 (덕명동)
- (72) 발명자
김형준
대전광역시 유성구 동서대로 125, S3동 314호
김정일
서울특별시 마포구 독막로18길 5 102동 1403호 (상수동, 두산위브아파트)
- (74) 대리인
특허법인 아이퍼스

전체 청구항 수 : 총 15 항

심사관 : 박승철

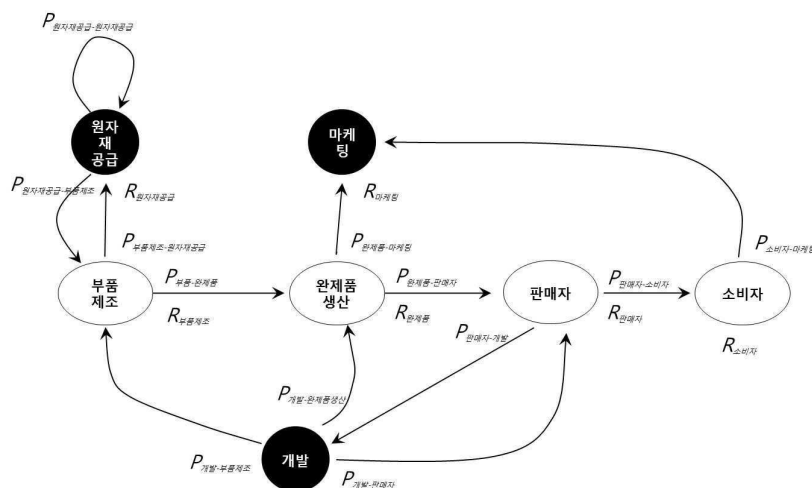
(54) 발명의 명칭 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 시스템 및 그 방법

(57) 요약

본 발명은 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 시스템에 대한 것이다.

보다 상세하게는 원자재를 이용한 완제품이 소비자에 이르기까지 상기 완제품의 생산 및 유통에 기여하는 적어도 (뒷면에 계속)

대표도



하나 이상의 상태(S) 또는 행동(A)이 가상의 체인-(Chain) 구조를 형성하고, 상태(S)에 따른 상기 원자재 또는 상기 원자재를 이용한 반제품 또는 상기 완제품의 재고량 또는 출하량 정보를 포함하는 현장보상값(R)를 생성하는 공급망을 최적화 하는 제조업 공급망 최적화 설계 시스템에 있어서, 상기 상태(S) 또는 상기 행동(A) 간의 전이확률(P, Transition Probability)이 $P[S_{t+1}|S_t] = P[S_{t+1}|S_1, \dots, S_t]$ 인 관계식을 만족하도록 상기 가상의 체인 구조를 구성하고 하기 수학식1을 만족하는 상태가치함수($v_{\pi}(s)$)와 하기 수학식2를 만족하는 행동가치함수($q_{\pi}(s,a)$) 및 하기 수학식3을 만족하는 정책(policy, $\pi(a|s)$)을 포함하는 MDP(Markov Decision Process) Factor 생성하는 공급망구성부와, 상기 현장보상값을 입력받는 현장보상값수집부 및, 현재시점에서 상기 현장보상값수집부를 통하여 입력 받은 상기 현장보상값과 상기 MDP Factor를 기반으로 특정된 미래시점에서 모든 상태(S)의 상기 상태가치함수의 총합이 특정된 미래목표값에 수렴하도록 정책예측값과 미래가중치예측값과 MDP Factor 예측값 및 현장보상값예측값을 포함하는 공급망예측제어정보를 산출하는 공급망학습부를 포함하는 공급망설계단; 및 상기 공급망예측제어정보를 기반으로 상기 공급망을 제어하는 공급망제어단;을 포함하는 것을 특징으로 하는 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 시스템:

[수학식]

$$\pi(a|s)=P(A_t=a|S_t=s)$$

$$\sum_{a \in A} \pi(a|s)=1$$

$$q_{\pi}(s,a)=R_s^a+\gamma \sum_{s' \in S} P_{ss'}^a v_{\pi}(s')$$

$$v_{\pi}=\sum_{a \in A} \pi(a|s) q_{\pi}(s,a)$$

상기 [수학식]에서

γ 는 미래가중치(discount factor)이고,

R는 현장보상값 이다.

에 관한 것이다.

명세서

청구범위

청구항 1

원자재를 이용한 완제품이 소비자에 이르기까지 상기 완제품의 생산 및 유통에 기여하는 적어도 하나 이상의 상태(S) 또는 행동(A)이 가상의 체인-(Chain) 구조를 형성하고, 상태(S)에 따른 상기 원자재 또는 상기 원자재를 이용한 반제품 또는 상기 완제품의 재고량 또는 출하량 정보를 포함하는 현장보상값(R)을 생성하는 공급망을 최적화 하는 제조업 공급망 최적화 설계 시스템에 있어서,

상기 상태(S) 또는 상기 행동(A) 간의 전이확률(P, Transition Probability)이 $P[S_{t+1}|S_t]=P[S_{t+1}|S_1,...,S_t]$ 인 관계식을 만족하도록 상기 가상의 체인 구조를 구성하고 하기 수학식을 만족하는 상태가치함수($v_{\pi}(s)$)와 하기 수학식을 만족하는 행동가치함수($q_{\pi}(s,a)$) 및 하기 수학식을 만족하는 정책(policy, $\pi(a|s)$)을 포함하는 MDP(Markov Decision Process) Factor 생성하는 공급망구성부와, 상기 현장보상값을 입력받는 현장보상값수집부 및, 현재시점에서 상기 현장보상값수집부를 통하여 입력 받은 상기 현장보상값과 상기 MDP Factor를 기반으로 특정한 미래시점에서 모든 상태(S)의 상기 상태가치함수의 총합이 특정한 미래목표값에 수렴하도록 정책예측값과 미래가중치예측값과 MDP Factor예측값 및 현장보상값예측값을 포함하는 공급망예측제어정보를 산출하는 공급망학습부를 포함하는 공급망설계단; 및

상기 공급망예측제어정보를 기반으로 상기 공급망을 제어하는 공급망제어단;을 포함하는 것을 특징으로 하는 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 시스템;

[수학식]

$$\pi(a|s)=P(A_t=a|S_t=s)$$

$$\sum_{a \in A} \pi(a|s)=1$$

$$q_{\pi}(s,a)=R_s^a+\gamma \sum_{s' \in S} P_{ss'}^a+v_{\pi}(s')$$

$$v_{\pi}=\sum_{a \in A} \pi(a|s)q_{\pi}(s,a)$$

상기 [수학식]에서

γ 는 미래가중치(discount factor)이고,

R는 현장보상값 이다.

청구항 2

제1항에 있어서,

상기 공급망설계단은 상기 현장보상값과 상기 MDP Factor과 상기 미래목표값 및 상기 공급망예측제어정보를 시계열적으로 저장하는 설계단저장부를 더 포함하는 것을 특징으로 하는 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 시스템.

청구항 3

제2항에 있어서,

상기 공급망설계단은 상기 설계단저장부에 시계열적으로 저장된 상기 현장보상값과 상기 현장보상값예측값을 비교하여 상기 MDP Factor과 상기 미래목표값 및 상기 공급망예측제어정보의 예측정확도를 산출하는 설계단평가부를 더 포함하는 것을 특징으로 하는 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 시스템.

청구항 4

제3항에 있어서,

상기 설계단저장부는, 상기 현장보상값과 상기 MDP Factor과 상기 미래목표값와 상기 공급망예측제어정보 및 상기 예측정확도 중에 적어도 하나 이상을 표시하는 설계단저장부표시유닛을 더 포함하는 것을 특징으로 하는 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 시스템.

청구항 5

제1항에 있어서,

상기 공급망설계단은 관리자로부터 정보를 입력받는 설계단입력부;을 더 구비하여,

상기 미래목표값과 상기 현장보상값과 상기 MDP Factor의 초기값(initial value)를 입력 받는 것을 특징으로 하는 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 시스템.

청구항 6

제5항에 있어서,

상기 공급망학습부는 마르코프 결정 프로세스(Markov Decision Process) 또는 벨만 방정식(Bellman Equation)을 기반으로 정책예측값과 미래가중치예측값와 MDP Factor예측값 및 현장보상값예측값을 포함하는 공급망예측제어정보를 산출하는 학습부계산유닛과, 상기 정책예측값과 미래가중치예측값와 MDP Factor예측값 및 현장보상값예측값을 포함하는 공급망예측제어정보를 상기 관리자에게 표시하는 학습부표시유닛을 더 포함하는 것을 특징으로 하는 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 시스템.

청구항 7

제1항에 있어서,

상기 공급망구성부는 관리자로부터 상기 체인 구조를 입력받는 구성부입력유닛을 더 포함하는 것을 특징으로 하는 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 시스템.

청구항 8

제7항에 있어서,

상기 공급망구성부는 공급자재고관리(VMI, Vendor Managed Inventory)와 즉시대응체제(QR, Quick Response)와 협력적 예약보충시스템(CPFR, Collaborative Planning, Forecasting and Replenishment)와 연속상품보충(CRP, Continuous Replenishment Program) 및 효율적소비자반응시스템(ECR, Efficient Consumer Response)을 포함하는 제조업기본SCM체인구조를 저장하는 제조업기본SCM저장부를 더 구비하여, 상기 관리자가 제조업기본SCM체인구조중 적어도 하나를 선택할 수 있는 것을 특징으로 하는 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 시스템.

청구항 9

원자재를 이용한 완제품이 소비자에 이르기까지 상기 완제품의 생산 및 유통에 기여하는 적어도 하나 이상의 상태(S) 또는 행동(A)이 가상의 체인(Chain) 구조를 형성하고, 상태(S)에 따른 상기 원자재 또는 상기 원자재를 이용한 반제품 또는 상기 완제품의 재고량 또는 출하량 정보를 포함하는 현장보상값을 생성하는 공급망을 최적화 하는 제조업 공급망 최적화 설계 방법에 있어서,

공급망구성부가, 상기 상태(S) 또는 행동(A)간의 전이확률(P, Transition Probability)이 $P[S_{t+1}|S_t]=P[S_{t+1}|S_1,...,S_t]$ 인 관계식을 만족하도록 상기 체인 구조를 구성하는 단계(S01);

공급망구성부가, 하기의 수학식을 만족하는 상태가치함수($v_{\pi}(s)$)와 하기의 수학식을 만족하는 행동가치함수($q_{\pi}(s,a)$) 및 하기의 수학식을 만족하는 정책(policy, $\pi(a|s)$)을 포함하는 MDP(Markov Decision Process) Factor 생성하는 단계(S02);

공급망학습부가, 현재시점에서 입력 받은 상기 현장보상값과 상기 MDP Factor를 기반으로 특정한 미래시점에서 모든 상태(S)의 상기 상태가치함수의 총합이 특정한 미래목표값에 수렴하도록 정책예측값과 미래가중치예측값과 MDP Factor 예측값 및 현장보상값예측값을 포함하는 공급망예측제어정보를 산출하는 단계(S03); 및

공급망제어단은, 상기 공급망예측제어정보를 기반으로 공급망제어부가 상기 공급망을 제어하는 단계(S04);를 포함하는 것을 특징으로 하는 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 방법;

[수학식]

$$\pi(a|s)=P(A_t=a|S_t=s)$$

$$\sum_{a \in A} \pi(a|s)=1$$

$$q_{\pi}(s,a)=R_s^a+\gamma \sum_{s' \in S} P_{ss'}^a v_{\pi}(s')$$

$$v_{\pi}=\sum_{a \in A} \pi(a|s) q_{\pi}(s,a)$$

상기 [수학식]에서

γ 는 미래가중치(discount factor)이고,

R는 현장보상값 이다.

청구항 10

제9항에 있어서,

상기 S03단계와 상기 S04단계 사이에,

상기 현장보상값과 상기 MDP Factor과 상기 미래목표값 및 상기 공급망예측제어정보가 설계단저장부에 시계열적으로 저장되는 단계(S05);를 더 포함하는 것을 특징으로 하는 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 방법.

청구항 11

제10항에 있어서,

상기 S04단계 이후에,

설계단평가부는, 상기 설계단저장부에 시계열적으로 저장된 상기 현장보상값 상기 현장보상값예측값을 비교하여 상기 MDP Factor과 상기 미래목표값 및 상기 공급망예측제어정보의 예측정확도를 산출하는 단계(S06);를 더 포함하는 것을 특징으로 하는 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 방법.

청구항 12

제11항에 있어서,

상기 S02 단계와 상기 S03단계 사이에,

설계단저장부표시유닛은, 상기 현장보상값과 상기 MDP Factor과 상기 미래목표값과 상기 공급망예측제어정보 및 상기 예측정확도 중에 적어도 하나 이상을 표시하는 단계(S07);를 더 포함하는 것을 특징으로 하는 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 방법.

청구항 13

제9항에 있어서,

상기 S02단계와 상기 S03단계 사이에,

설계부입력단을 통하여, 관리자로부터 상기 미래목표값과 상기 현장보상값과 상기 MDP Factor의 초기값(initial value)를 입력 받는 단계(S08);를 더 포함하는 것을 특징으로 하는 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 방법.

청구항 14

제13항에 있어서,

상기 S03단계는,

마르코프 결정 프로세스(Markov Decision Process) 또는 벨만 방정식(Bellman Equation)을 이용하여 정책예측값과 미래가중치예측값과 MDP Factor예측값 및 현장보상값예측값을 포함하는 공급망예측제어정보를 산출하는 것을 특징으로 하는 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 방법.

청구항 15

제9항에 있어서,

상기 S01단계는

관리자로부터 상기 체인 구조를 입력받는 단계(S10);를 더 포함하는 것을 특징으로 하는 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 방법.

발명의 설명

기술 분야

[0001] 본 발명은 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 시스템에 대한 것이다.

[0002] 보다 상세하게는, 다양한 SCM 구조를 시간에 따라 행동에 대한 보상 가치가 달라지는 것을 discount factor로 표시하는 마르코프 결정 프로세스(Markov Decision Process)로 구성하고, 강화학습 기반의 딥러닝을 적용함으로써, 최적화된 공급망 구조 및 이를 위한 공급망관리정책과 연동된 discount factor를 도출할 수 있는 마르코프

결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 시스템에 관한 것이다.

배경 기술

- [0004] 공급망(Supply Chain, SC)은 제품이나 서비스를, 원천에서 소비에 이르도록 공급하기 위해, 원재료 공급업체, 제조업체의 제조공장, 완제품 출하 및 유통, 소매할인점 및 고객을 연결하는 조직 및 비즈니스 프로세스의 네트워크이다.
- [0005] 이러한 공급망을 제공업자로부터 생산자, 배포자, 고객에 이르는 물류의 흐름을 하나의 가치사슬 관점에서 파악하고, 필요한 정보가 원활히 흐르도록 지원하는 시스템을 공급망관리(Supply Chain Management, SCM)이다. 기업 내에 부문별 최적화나 개별 기업단위의 최적화에서 탈피하여 공급망의 구성요소들 간에 이루어지는 전체프로세스 최적화를 달성하고자 하는 경영혁신기법중에 하나이다.
- [0006] 사회구조가 복잡해지면서, 그에 맞는 새로운 형태와 기능의 상품들이 개발되고 있으며, 기존에 인간 생활에서 사용되어오던 상품들의 그 특징과 기능에 큰 변화가 일어나고 있다.
- [0008] 의류 업계의 선두 주자인 Zara사와 UNIQLO사는 뛰어난 공급망 관리가 우수한 업체이지만, 두 브랜드는 정 반대의 공급망관리정책을 선택하고 있다.
- [0009] UNIQLO사는 기본 아이템의 소품종 대량생산의 제품과 생산전략을 선택함으로써 높은 재고 수준을 유지하는 재고 전략을 유지하며, 많은 제조 기능을 아웃소싱하며, 연령대 별로 고른 비율로 주요 고객층이 분포한다. 따라서 공급망관리에 있어서, 효율성을 최우선으로 한다.
- [0010] 이에 비하여ZARA사는 패션을 민감한 아이템의 다품종 소량생산의과 제품, 생산전략을 선택함으로써 낮은 재고 수준을 유지하는 재고 전략을 유지하며, 디자인부터 제조 및 판매까지 수직계열화하며, 주로 젊은 여성들이 고객층이 분포한다. 따라서 공급망관리에 있어서,속도들을 최우선으로 한다.
- [0012] 시장조사 및 컨설팅 기관인 가트너(Gartner)가 2013년도에 발표한 Supply Chain Top 25에 의하면, 스마트폰 업계의 국제적 선도기업인 Apple사는 종합점수 10점 만점에 9.51을 기록하여, 5.87에 불과한 종합점수를 기록한 세계적인 햄버거 프랜차이즈업체인 McDonald's 를 월등히 앞서 1위를 기록했다. 또한 SCM이라면 세계적 판매 유통업체인 Amazone.com도 5.86에 불과한 점수로 3위에 그친 것으로 보면, Apple 사의 세계적 시장 지배력은 제품 성능과 혁신뿐만이 아니라 공급망관리도 큰 역할을 하고 있다고 판단할 수 있다. 특히 현재 Apple사의 CEO를 맡고 있는 팀쿡(Tim Cook)이 SCM 전문가라는 점을 주지하면, 앞으로의 산업 생태계에서 공급망관리가 얼마나 중요한 기업의 경쟁력 요소인지를 잘 알 수 있다.
- [0014] 종래의 공급망관리는, 실제 비즈니스 환경과 제조업체 및 유통 구조 현황을 고려하지 않은 상황에서, 기업 내부에서 결정된 주요 정책(Policy)을 검증이나 확인 없이 적용하여 SCM을 수행해 왔다. 또한 산업별, 고객별로 다양한 SCM 정책이 사용되고 있지만, 이를 선택하기 위한 기준과 권한이 SCM을 수행하는 기업의 관계자들에게 제대로 제공되지 못하고 있다.
- [0015] 따라서 다양한 SCM 구조를 시간에 따라 행동에 대한 보상 가치가 달라지는 것을 discount factor로 표시하는 마르코프 결정 프로세스(Markov Decision Process)로 구성하고, 강화학습 기반의 딥러닝을 적용함으로써, 최적화된 공급망 구조 및 이를 위한 공급망관리정책과 연동된 discount factor를 도출할 수 있는 제조업 기반의 공급 체인 최적화 설계 시스템 개발의 필요성이 대두되고 있다.

선행기술문헌

특허문헌

- [0017] (특허문헌 0001) 공개특허 10-2016-0071776
- (특허문헌 0002) 공개특허 10-2016-0063006
- (특허문헌 0003) 공개특허 10-2009-0003488
- (특허문헌 0004) 등록특허 10-1573985

발명의 내용

해결하려는 과제

[0018] 따라서 본 발명은 상기와 같은 종래의 문제점을 해결하기 위하여 안출된 것으로서, 본 발명의 일 실시예에 따르면, 마르코프 결정 프로세스에서 규정하고 있는 상태(S), 행동(A) 및 상태 전이 확률(P)의 특징을 만족하도록 특정 기업의 상황에 맞는 공급망을 구성하고, 이렇게 구성된 공급망 구성에 해당 기업의 내외 상황을 고려한 정책(π)과 미래가중치(γ)를 입력하여 각 상태(S)에 따른 보상(R)의 예상치를 도출할 수 있으며, 모든 상태가치 함수의 총합이 미리 설정한 특정 목표값에 근접수렴 하도록, 입력으로 사용한 정책(π) 또는 미래가중치(γ)를 변화시키면서 공급망 구성을 최적화함으로써, 기업이 목표로 하는 가치에 도달하기 위해서 필요한 비즈니스 전략 수립에 활용할 수 있는 정책(π)과 현재와 미래가치의 중요도를 안내할 수 있는 미래가중치(γ)를 도출할 수 있는 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 시스템을 제공하는데 목적이 있다.

[0020] 한편, 본 발명에서 이루고자 하는 기술적 과제들은 이상에서 언급한 기술적 과제들로 제한되지 않으며, 언급하지 않은 또 다른 기술적 과제들은 아래의 기재로부터 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0022] 본 발명의 제1목적은 원자재를 이용한 완제품이 소비자에 이르기까지 상기 완제품의 생산 및 유통에 기여하는 적어도 하나 이상의 상태(S) 또는 행동(A)이 가상의 체인-(Chain) 구조를 형성하고, 상태(S)에 따른 상기 원자재 또는 상기 원자재를 이용한 반제품 또는 상기 완제품의 재고량 또는 출하량 정보를 포함하는 현장보상값(R)을 생성하는 공급망을 최적화 하는 제조업 공급망 최적화 설계 시스템에 있어서, 상기 상태(S) 또는 상기 행동(A) 간의 전이확률(P, Transition Probability)이 $P[S_{t+1}|S_t]=P[S_{t+1}|S_1,...,S_t]$ 인 관계식을 만족하도록 상기 가상의 체인 구조를 구성하고 하기 수학식1을 만족하는 상태가치함수($v_\pi(s)$)와 하기 수학식2를 만족하는 행동가치함수($q_\pi(s,a)$) 및 하기 수학식3을 만족하는 정책(policy, $\pi(a|s)$)을 포함하는 MDP(Markov Decision Process) Factor 생성하는 공급망 구성부와, 상기 현장보상값을 입력 받는 현장보상값 수집부 및, 현재시점에서 상기 현장보상값 수집부를 통하여 입력 받은 상기 현장보상값과 상기 MDP Factor를 기반으로 특정한 미래시점에서 모든 상태(S)의 상기 상태가치함수의 총합이 특정한 미래목표값에 수렴하도록 정책예측값과 미래가중치 예측값과 MDP Factor 예측값 및 현장보상값 예측값을 포함하는 공급망예측제어정보를 산출하는 공급망학습부를 포함하는 공급망설계단; 및 상기 공급망예측제어정보를 기반으로 상기 공급망을 제어하는 공급망제어단;을 포함하는 것을 특징으로 하는 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 시스템:

[0023] [수학식]

$$\pi(a|s)=P(A_t=a|S_t=s)$$

[0024]

$$\sum_{a \in A} \pi(a|s)=1$$

[0025]

$$q_\pi(s,a)=R_s^a+\gamma \sum_{s' \in S} P_{ss'}^a + v_\pi(s')$$

[0026]

$$v_\pi = \sum_{a \in A} \pi(a|s) q_\pi(s,a)$$

[0027]

[0028] 상기 [수학식]에서

[0029] γ 는 미래가중치(discount factor)이고,

[0030] R는 현장보상값 이다.

- [0031] 을 제공함으로써 달성될 수 있다.
- [0032] 또한, 상기 공급망설계단은 상기 현상보상값과 상기 MDP Factor와 상기 미래목표값 및 상기 공급망예측제어정보를 시계열적으로 저장하는 설계단저장부를 더 포함하는 것을 특징으로 할 수 있다.
- [0033] 그리고, 상기 공급망설계단은 상기 설계단저장부에 시계열로 저장된 상기 현상보상값 상기 현상보상예측값을 비교하여 상기 MDP Factor와 상기 미래목표값 및 상기 공급망예측제어정보의 예측정확도를 산출하는 설계단평가부를 더 포함하는 것을 특징으로 하는 할 수 있다.
- [0034] 또한, 상기 설계단저장부는, 상기 현상보상값과 상기 MDP Factor와 상기 미래목표값와 상기 공급망예측제어정보 및 상기 예측정확도 중에 적어도 하나 이상을 표시하는 설계단저장부표시유닛을 더 포함하는 것을 특징으로 할 수 있다.
- [0035] 그리고, 상기 공급망설계단은 관리자로부터 정보를 입력받는 설계단입력부;을 더 포함시켜, 상기 미래목표값과 상기 현상보상값과 상기 MDP Factor의 초기값(initial value)를 입력 받는 것을 특징으로 할 수 있다.
- [0036] 또한, 상기 공급망학습부는 마르코프 결정 프로세스(Markov Decision Process) 또는 벨만 방정식(Bellman Equation)을 기반으로 정책예측값과 미래가중치예측값과 MDP Factor예측값 및 현상보상값예측값을 포함하는 공급망예측제어정보를 산출하는 학습부계산유닛과, 상기 정책예측값과 미래가중치예측값과 MDP Factor예측값 및 현상보상값예측값을 포함하는 공급망예측제어정보를 상기 관리자에게 표시하는 학습부표시유닛을 더 포함하는 것을 특징으로 할 수 있다.
- [0037] 그리고, 상기 공급망구성부는 관리자로부터 상기 체인 구조를 입력받는 구성부입력유닛을 더 포함하는 것을 특징으로 할 수 있다.
- [0038] 또한, 상기 공급망구성부는 공급자재고관리(VMI, Vendor Managed Inventory)와 즉시대응체제(QR, Quick Response)와 협력적 예약보충시스템(CPFR, Collaborative Planning, Forecasting and Replenishment)와 연속상품보충(CRP, Continuous Replenishment Program) 및 효율적소비자반응시스템(ECR, Efficient Consumer Response) 등 제조업 기본 SCM policy 를 포함하는 구성부제조업기본SCM저장부를 더 구비하여, 상기 관리자가 제조업기본SCM체인구조중 적어도 하나를 선택할 수 있는 것을 특징으로 할 수 있다.
- [0040] 본 발명의 제2목적은, 원자재를 이용한 완제품이 소비자에 이르기까지 상기 완제품의 생산 및 유통에 기여하는 적어도 하나 이상의 상태(S) 또는 행동(A)이 가상의 체인(Chain) 구조를 형성하고, 상태(S)에 따른 상기 원자재 또는 상기 원자재를 이용한 반제품 또는 상기 완제품의 재고량 또는 출하량 정보를 포함하는 현상보상값을 생성하는 공급망을 최적화 하는 제조업 공급망 최적화 설계 방법에 있어서, 공급망구성부가, 상기 상태(S) 또는 행동(A)간의 전이확률(P, Transition Probability)이 $P[S_{t+1}|S_t]=P[S_{t+1}|S_1,...,S_t]$ 인 관계식을 만족하도록 상기 체인 구조를 구성하는 단계(S01); 공급망구성부가, 하기의 수학식1을 만족하는 상태가치함수($v(s)$)와 하기의 수학식2를 만족하는 행동가치함수($q_\pi(s,a)$) 및 하기의 수학식3을 만족하는 정책(policy, $\pi(a|s)$)을 포함하는 MDP(Markov Decision Process) Factor 생성하는 단계(S02); 공급망학습부가, 현재시점에서 입력 받은 상기 현상보상값과 상기 MDP Factor를 기반으로 특정한 미래시점에서 모든 상태(S)의 상기 상태가치함수의 총합이 특정한 미래목표값에 수렴하도록 정책예측값과 미래가중치예측값과 MDP Factor예측값 및 현상보상값예측값을 포함하는 공급망예측제어정보를 산출하는 단계(S03); 및 공급망제어단은, 상기 공급망예측제어정보를 기반으로 공급망제어부가 상기 공급망을 제어하는 단계(S04);를 포함하는 것을 특징으로 하는 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 방법:

[0041] [수학식]

[0042]
$$\pi(a|s)=P(A_t=a|S_t=s)$$

[0043]
$$\sum_{a \in A} \pi(a|s)=1$$
,

[0044]
$$q_\pi(s,a)=R_s^a+\gamma \sum_{s' \in S} P_{ss'}^a v_\pi(s')$$
,

$$v_{\pi} = \sum_{a \in A} \pi(a|s) q_{\pi}(s, a),$$

[0045]

[0046]

[0047]

[0048]

[0049]

[0050]

[0051]

[0052]

[0053]

[0054]

[0055]

상기 [수학식]에서

γ 는 미래가중치(discount factor)이고,

R는 현장보상값 이다.

을 제공함으로써 달성될 수 있다.

그리고, 상기 S03단계와 상기 S04단계 사이에, 상기 현장보상값과 상기 MDP Factor과 상기 미래목표값 및 상기 공급망예측제어정보가 설계단저장부에 시계열적으로 저장되는 단계(S05);를 더 포함하는 것을 특징으로 할 수 있다.

또한 상기 S04단계 이후에, 설계단평가부는, 상기 설계단저장부에 시계열적으로 저장된 상기 현장보상값 상기 현장보상예측값을 비교하여 상기 MDP Factor과 상기 미래목표값 및 상기 공급망예측제어정보의 예측정확도를 산출하는 단계(S06);를 더 포함하는 것을 특징으로 할 수 있다.

그리고, 상기 S02 단계와 상기 S03단계 사이에, 설계단저장부표시유닛은, 상기 현장보상값과 상기 MDP Factor과 상기 미래목표값과 상기 공급망예측제어정보 및 상기 예측정확도 중에 적어도 하나 이상을 표시하는 단계(S07);를 더 포함하는 것을 특징으로 할 수 있다.

또한, 상기 S02단계와 상기 S03단계 사이에, 설계부입력단을 통하여, 관리자로부터 상기 미래목표값과 상기 현장보상값과 상기 MDP Factor의 초기값(initial value)를 입력 받는 단계(S08);를 더 포함하는 것을 특징으로 할 수 있다.

그리고, 상기 S03단계는, 마르코프 결정 프로세스(Markov Decision Process) 또는 벨만 방정식(Bellman Equation)을 이용하여 정책예측값과 미래가중치예측값과 MDP Factor예측값 및 현장보상값예측값을 포함하는 공급망예측제어정보를 산출하는 것을 특징으로 할 수 있다.

또한, 상기 S01단계는 관리자로부터 상기 체인 구조를 입력받는 단계(S10);를 더 포함하는 것을 특징으로 할 수 있다.

발명의 효과

[0057]

본 발명의 일 실시예에 따르면, 마르코프 결정 프로세스에서 규정하고 있는 상태(S), 행동(A) 및 상태 전이 확률(P)의 특징을 만족하도록 특정 기업의 상황에 맞는 공급망을 구성하고, 이렇게 구성된 공급망 구성에 해당 기업의 내외 상황을 고려한 정책(π)과 미래가중치(γ)를 입력하여 각 상태(S)에 따른 보상(R)의 예상치를 도출할 수 있으며, 모든 상태가치함수의 총합이 미리 설정한 특정 목표값에 근접수렴하도록, 입력으로 사용한 정책(π) 또는 미래가중치(γ)를 변화시키면서 공급망 구성을 최적화함으로써, 기업이 목표로하는 가치에 도달하기 위해서 필요한 비즈니스 전략 수립에 활용할 수 있는 정책(π)과 현재와 미래가치의 중요도를 안내할 수 있는 미래가중치(γ)를 도출 할 수 있는 장점이 있다.

[0059]

한편, 본 발명에서 얻을 수 있는 효과는 이상에서 언급한 효과들로 제한되지 않으며, 언급하지 않은 또 다른 효과들은 아래의 기재로부터 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

[0061]

도 1a. 일반적인 강화학습 매커니즘 개념도.

도 1b. 마르코프 프로세스 개념도.

도 1c. 마르코프 보상 프로세스 개념도.

도 2. 본 발명의 일 실시예에 따른 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급

체인 최적화 설계 시스템에 적용되는 공급망 구성의 예시.

도 3. 본 발명의 일 실시예에 따른 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급체인 최적화 설계 시스템에 적용되는 공급망 구성을 활용하여 정책(π)과 미래가중치(γ) 예측값을 도출하기 위한 계산 방법.

도 4. 본 발명의 일 실시예에 따른 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급체인 최적화 설계 시스템의 블록 구성도.

발명을 실시하기 위한 구체적인 내용

- [0062] 이하 첨부된 도면을 참조하여 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자가 본 발명을 쉽게 실시할 수 있는 실시예를 상세히 설명한다. 다만, 본 발명의 바람직한 실시예에 대한 동작 원리를 상세하게 설명함에 있어 관련된 공지 기능 또는 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략한다.
- [0063] 또한, 도면 전체에 걸쳐 유사한 기능 및 작용을 하는 부분에 대해서는 동일한 도면 부호를 사용한다. 명세서 전체에서, 어떤 부분이 다른 부분과 연결되어 있다고 할 때, 이는 직접적으로 연결되어 있는 경우뿐만 아니라, 그 중간에 다른 소자를 사이에 두고, 간접적으로 연결되어 있는 경우도 포함한다. 또한, 어떤 구성요소를 포함한다는 것은 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라, 다른 구성요소를 더 포함할 수 있는 것을 의미한다.
- [0064] 이하에서는 본 발명의 일 실시예에 따른 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급체인 최적화 설계 시스템의 구성 및 기능에 대해 설명하도록 한다.
- [0065] 도 1a에는 일반적인 강화학습 매커니즘의 개념도를 도시하였다. 강화학습(reinforcement learning)은 귀납적 학습(inductive learning)인 기계학습(machine learning)의 한 종류로서, 행동(Action)에 대응하는 정확한 출력 정보를 제공하지는 않지만, 보상(Reward)라는 평가 정보는 주어지는 문제에 대해, 행동(Action)과 보상(Reward)의 연관성으로부터 각 상태의 최적 행동을 결정할 수 있다.
- [0066] 예를 들어, 임의의 공급망의 생산 상태에서 매출 증대를 목표로 제품의 증산 계획을 수립하여 수행하였을 경우에, 시장 수요가 감소하면 증산에 투입한 자재 및 인건비를 비롯한 생산 비용의 증가, 재고 증가로 인한 창고비 증가 등의 부정적 요인으로 인해, 생산중단 또는 생산량 감소와 같은 생산 상태에서의 손실요소가 발생할 수 있고, 결국 증산이라는 행동(Action)에 대해 생산중단 또는 생산량 감소라는 음의 실수값으로 표시되는 보상(Reward)이 주어진 것으로 간주할 수 있다.
- [0067] 반대로, 시장 수요가 증가하면 매출 증가, 이익 구조개선 등으로 생산량 유지 또는 증가와 같은 이익 요소가 발생할 수 있으며, 결국 증산이라는 행동(Action)에 대해 생산증가라는 양의 이익 양의 실수값으로 표시되는 보상(Reward)이 주어진 것으로 간주할 수 있다.
- [0068] 따라서 과거의 행동(Action)과 행동에 따른 보상(Reward)의 반복 경험을 통하여, 최적화된 미래의 행동을 예측할 수 있다.
- [0069] 그러나, 실제 공급망 관리에서는 도 1a와 같이, 단순한 구조의 행동(Action)과 보상(Reward)로 표현되지 않으며, 다수개의 상태(State)와 상태 전이 확률(state transition probability)이 복잡한 형태를 구성한다.
- [0070] 앞서 인용한 생산 상태에서 증산할 수도 있으며, 감산할 수도 있다. 따라서 생산 상태에서 가능한 행동은 두 가지이며, 각각은 확률로 표현될 수 있다. 만약에 감산할 확률이 10%라고 가정하면, 증산할 확률은 90%로 분할될 수 있다.
- [0071] 또한 상태(state)의 개수가 많아지면, 보상(reward)가 주어지기까지 여러 단계의 상태 전이를 거쳐야 한다. 앞서 인용한 생산 상태에서 증산 또는 감산의 행동(action)에 대한 보상(reward)은, 증산 또는 감산이라는 행동(action)에 대한 구매-생산-유통-판매로 연결되는 공급망 전체의 영향을 모두 고려하여야 한다.
- [0072] 따라서 강화 학습을 통하여, 현재 행동(action)에 의해 결정되는 상태(state)별 보상(reward)를 예측함으로써, 최적 행동(action)을 도출할 수 있는 공급망 구조를 설계할 수 있어 한다.
- [0073] 도 1b는 마르코프(Markov Process) 프로세스 개념도를 도시하였다.
- [0074] 도 1b를 참조하면, 각각의 상태(state)는 일년의 시간적 사건을 의미하며, 현재 단계의 상태에서 예상되는 다음

단계의 상태는 과거의 사건과 무관한 것으로 기술된다. 즉, 상태1은 상태2 또는 상태3에 영향을 미치지만, 상태4에는 영향을 미치지 않으며, 아래의 [수학식1]로 정의될 수 있다.

수학식 1

$$P[S_{t+1}|S_t]=P[S_{t+1}|S_1,...,S_t]$$

[수학식1]에서 $P[a|b]$ 는 상태b에서 상태a로의 상태 전이확률이다.

앞서 언급한 바와 같이 일반적인 실제 공급망 관리에서는 다수의 상태(state)와 상태와 상태 사이의 전이 확률(state transition probability)가 복잡하게 연결되어있다. 도 1b에 도시한 바와 같은 마르코프 프로세스를 활용하여 실제 공급망을 구성하면, 물류 또는 자금의 흐름에서 각 상태(state) 사이의 연관성을 명확히 규정할 수 있는 장점이 있다. 또한 수학식 1에서 확인할 수 있듯이, 마르코프 프로세스는 현재 단계 상태(state)에서 예상되는 다음 단계 상태는 과거사건과 무관하다고 전제하고 있기 때문에, 각 상태 사이의 연관성, 즉 상태 전이 확률(state transition probability) 산출이 명확해지며, 전체 프로세스 전체의 예측 기능을 가능하게 한다.

도 1c에는 마르코프 보상 프로세스를 도시하였다. 도 1b에 도시한 마르코프 프로세스에 보상(reward, R)과 미래가중치(discount factor, γ)를 적용하여 마르코프 보상 프로세스를 구성할 수 있으며 이때의 상태 가치 함수는 아래의 [수학식2]로 정의될 수 있다.

수학식 2

$$v(s)=R-\gamma+\sum_{s'\in S}P_{ss'}v(s')$$

도 1c에 도시한 마르코프 보상 프로세스는 도 1b에 도시한 마르코프 프로세스에, 마르코프 프로세스에 임의의 상태(state)에서 다음 단계로 이동하는 행동(action)을 취할 때, 마르코프 프로세스를 만족하도록 구성된 공급망으로부터 피드백 받는 보상(reward)를 포함한 것이다.

도 2에는 본 발명의 일 실시예에 따른 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급체인 최적화 설계 시스템에 적용되는 공급망 구성의 예시를 도시하였다.

도 2는 도 1c의 마르코프 보상프로세스에, 행동(action) 및 정책(π)을 포함함으로써 실제 공급망 구조를 모사한 구성의 예시를 도시한 것이다.

이때 상태(state)와 행동(action)을 연결해주는 함수로서 상태(state)가 주어졌을 때 행동(action)을 할 확률로 정의되는 정책(policy, π)를 규정할 수 있으며, 정책(π)은 수학식 3으로 표현될 수 있다.

수학식 3

$$\pi(a|s)=P(A_t=a|S_t=s)$$

$$\sum_{a\in A}\pi(a|s)=1$$

또한, 상태 s에서 행동 a를 취하고 정책 π 를 따랐을 때 기대할 수 있는 가치의 총합은 수학식 4와 같이 표현될 수 있다.

수학식 4

$$q_{\pi}(s,a)=R_s^a+\gamma \sum_{s' \in S} P_{ss'}^a v_{\pi}(s')$$

[0089]

[0090] 이때, [수학식 2]에 표시된 상태가치함수는 수학식 5와 같이 정책을 포함한 상태 가치함수로 수정될 수 있다.

수학식 5

$$v_{\pi}=\sum_{a \in A} \pi(a|s) q_{\pi}(s,a)$$

[0091]

[0092] 도 2에 도시된 바와 같이 본 발명의 일 실시예에 따른 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 시스템에서는, 원자재 공급부터 소비자에 이르기까지 부품제조, 완제품 제조, 및 판매자를 포함하는 여러 단계의 상태(state)와 구매, 마케팅, 출하, 공정개선, 개발과 같은 다양한 행동(action)이 서로 연결되어 공급망을 구성할 수 있다.

[0093] 이러한 공급망 구성은 산업과 고객, 기업 내외부의 환경 및 기업의 비즈니스 정책에 따라서 기업 고유의 공급망이 구성될 수 있다. 예를 들어 부품이나 원자재의 안정적인 공급을 위해서 완제품 생산 업체는 원자재나 부품 제조공급업체에 완제품 생산 계획이나 추진계획을 공유하는 공급보증관리 또는 선지급을 공급망 관리에 포함하는 경우가 있다. 또한 제품 개발이나 공정개선과 같은 행동들을 공급망 구조에 포함시키는 경우도 있다.

[0094] 앞서 언급한 바와 같이 공급망 구성은 여러 가지 요인에 의해서 결정될 수 있는데, 기업 내외부의 산업 환경뿐만 아니라, 고도의 기술력이 보편화되고 소비자의 생활 패턴, 양적 질적 발전이 가속화되어 지리적 정치적 한계가 거의 없어지고 있는 물류시스템의 영향으로, 개별 기업에 가장 적합한 공급망을 구성하는 경향이 두드러지고 있다.

[0095] 또한, 앞서 설명한 바와 같이 기업 내 외부의 비즈니스 환경은 복잡해지면서, 시장 선점에 시간이 무엇보다도 중요해지는 상황에서, 기업은 현재와 미래 가치를 명확히 하고 그에 따라서 경영 정책을 제시함으로써, 사업화의 성공여부가 결정된다. 그러므로, 도 2에 도시된 바와 같은 본 발명의 일 실시예에 따른 공급 체인 최적화 설계 시스템에서 정의하고 있는 정책(p)와 미래가중치(γ)를 변수로 하여, 임의의 시점의 임의의 단계에서 취할 수 있는 행동에 따른 보상을 예상할 수 있다.

[0096] 따라서, 도 2에 도시된 본 발명의 일 실시예에 따른 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 시스템에 적용된 공급망 구성을 활용하면, 마르코프 결정 프로세스에서 규정하고 있는 상태(S), 행동(A) 및 상태 전이 확률(P)의 특징을 만족하도록 특정 기업의 상황에 맞는 공급망을 구성하고, 이렇게 구성된 공급망 구성에 해당 기업의 내외 상황을 고려한 정책(π)과 미래가중치(γ)를 입력하여 각 상태(S)에 따른 보상(R)의 예상치를 도출할 수 있다. 이때, 모든 상태가치함수의 총합이 특정한 미리목표 값에 근접수렴하도록, 입력으로 사용한 정책(π) 또는 미래가중치(γ)를 변화시키면서 공급망 구성을 최적화함으로써, 기업이 목표로 하는 가치에 도달하기 위해서 필요한 비즈니스 전략 수립에 활용할 수 있는 정책(π)과 현재와 미래가치의 중요도를 안내할 수 있는 미래가중치(γ)를 도출 할 수 있다.

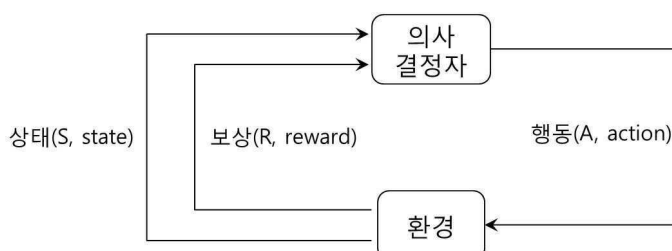
[0097] 도 3에는 본 발명의 일 실시예에 따른 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급 체인 최적화 설계 시스템에 적용되는 공급망 구성을 활용하여 정책(π)과 미래가중치(γ)를 도출하기 위한 계산 방법을 도시하였다.

[0098] 도 2와 수학식 5를 참조하면, 정책(π)과 미래가중치(γ)의 최적값을 도출하기 위해서는 정책(π)과 미래가중치(γ)를 변수로 하되, 각각의 상태(S)가 상호 연결되어있으며, 특히 보상(R)은 특정 단계에서 다음 단계로 이동하는 행동을 취할 때 피드백(feedback)받는 스칼라 실수값이기 때문에, 정책(π)과 미래가중치(γ)의 입력 조건에 따라서, 상태(S) 가치 함수가 수렴(convergence)하지 못하고 발산(divergence)는 상황이 도출 될 수 있다.

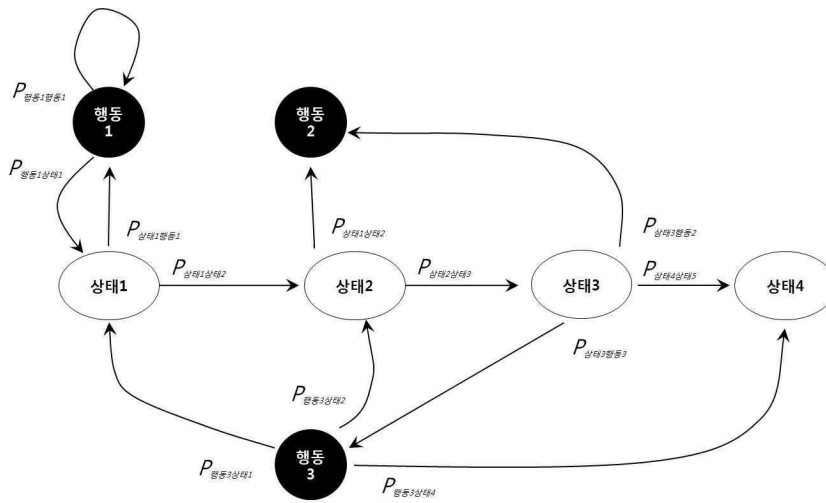
- [0099] 도 2 내지 도 3을 참조하면, 본 발명의 일 실시예에 따른 공급체인 최적화 설계 시스템은, 모든 상태(S)와 행동(A) 및 상태 전이 확률을 비롯한 공급망 구성에 대한 모든 정보를 알고 있기 때문에, 상태가치함수의 총합을 특정한 미래목표값에 수렴할 때까지 정책(π)과 미래가중치(γ)의 입력값 변경과 계산을 반복함으로써, 정책(π)과 미래가중치(γ)의 최적값이 도출될 수 있음을 알 수 있다.
- [0100] 이렇게 도출된 정책(π)과 미래가중치(γ) 정보를 기반으로 하여, 각각의 상태(S)에서 행동과 상태 전이를 선택할 수 있는데, 이때 선택된 행동이나 상태 전이를 취할 때 모든 상태(S)에서 보상(R)이 피드백 된다.
- [0101] 따라서 시계열적으로 구분되는 특정한 단계에서 도출된 최적 정책(π)과 미래가중치(γ)의 예측값은, 다음 단계에서의 피드백되는 상태(S)별 보상(R)값을 이용하여 그 정확도를 검증할 수 있다. 즉, 도출된 입력값인 정책(π)과 미래가중치(γ)을 기반으로 각 상태(S)에서 행동(A)를 결정하고, 이에 대한 평가정보가 주어짐으로써, 구성된 공급망의 환경 변수를 학습하여, 정책(π)과 미래가중치(γ)의 입력과 보상(R)사이의 함수관계를 귀납적으로 도출할 수 있다. 이러한 학습 방법은 출력에 대한 정보는 제공하지 않지만, 평가정보는 주어지는 문제에 대해 각 상태에서의 행동을 결정하는 강화학습(reinforcement learning)에 해당한다.
- [0102] 도 4에는 본 발명의 일 실시예에 따른 마르코프 결정 프로세스와 기계학습 알고리즘을 이용한 제조업 기반의 공급체인 최적화 설계 시스템의 블록 구성도를 도시하였다.
- [0103] 도 4에 도시한 바와 같이, 본 발명의 일 실시예에 따른 공급체인 최적화 설계 시스템은 업체 내외부의 환경에 맞게 구성된 공급망(supply chain)과 공급망의 각 구성인자들을 설계하는 공급망설계단, 그리고 공급망설계단에서 도출된 결과물을 기반으로 공급망을 제어하는 공급망제어단으로 구성됨을 알 수 있다.
- [0104] 도 2의 설명에서 언급한 바와 같이, 공급망의 상태(S)를 포함하는 모든 구성요소는, 산업별 고객별 특징에 따라서 업체에 적합하도록 구성될 수 있는데, 공급망 구성부는 수학적 1 내지 수학적 4를 만족하여, 공급망의 모든 구성이 마르코프 결정 프로세스 만족하도록 구성한다.
- [0105] 현장보상값수집부는, 공급망 구조에서 각 상태(S)별 보상값인 현장보상값을 입력 받으며, 공급망학습부는 최적 정책(π)과 미래가중치(γ) 예측값을 산출한다.
- [0106] 설계단 저장부는 반복계산에 따른 결과값과 예측된 정책(π)과 미래가중치(γ)등을 저장함으로써, 최적화 설계 시스템의 예측 정확도를 평가할 수 있으며, 이를 공급망 관리자에게 표시함으로써 공급망 설계의 효율성을 평가함과 동시에, 관리자로 하여금 설계시스템의 변경 또는 수정을 가능하도록 한다.
- [0107] 또한, 설계된 공급망 구조에서 상태(S) 및 행동(A)를 포함하는 일부 구성에 대한 정보가 없을 경우에는, 몬테-카를로 강화학습(Monte-Carlo Reinforcement Learning) 알고리즘을 적용할 수도 있다.
- [0109] 또한, 상기와 같이 설명된 장치 및 방법은 상기 설명된 실시예들의 구성과 방법이 한정되게 적용될 수 있는 것이 아니라, 상기 실시예들은 다양한 변형이 이루어질 수 있도록 각 실시예들의 전부 또는 일부가 선택적으로 조합되어 구성될 수도 있다.

도면

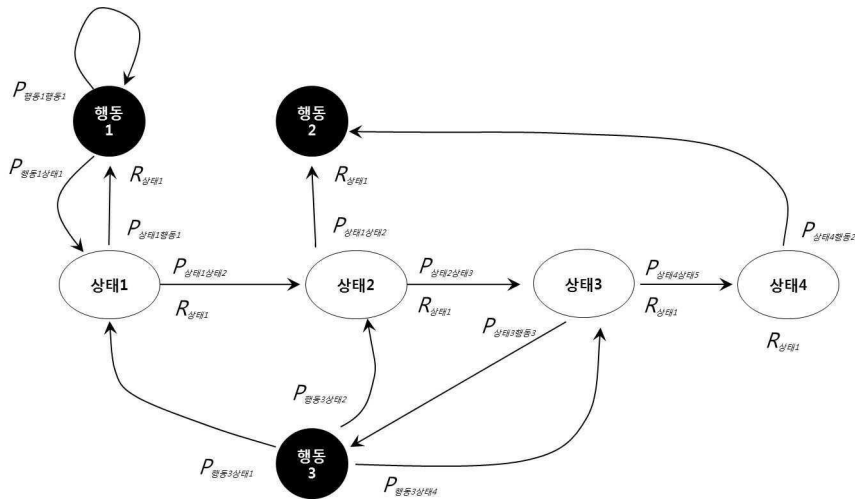
도면1a



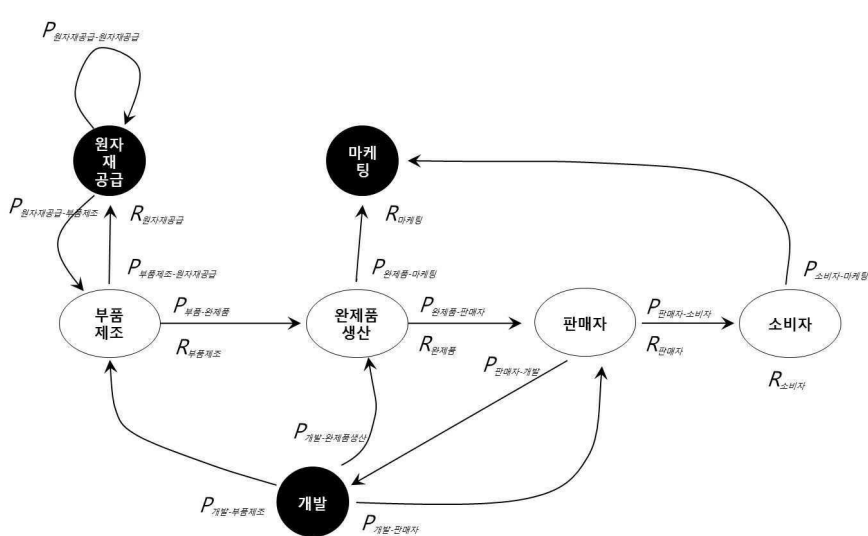
도면1b



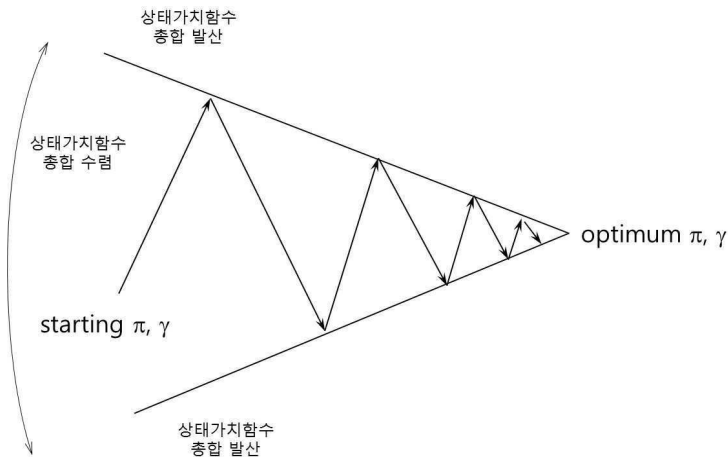
도면1c



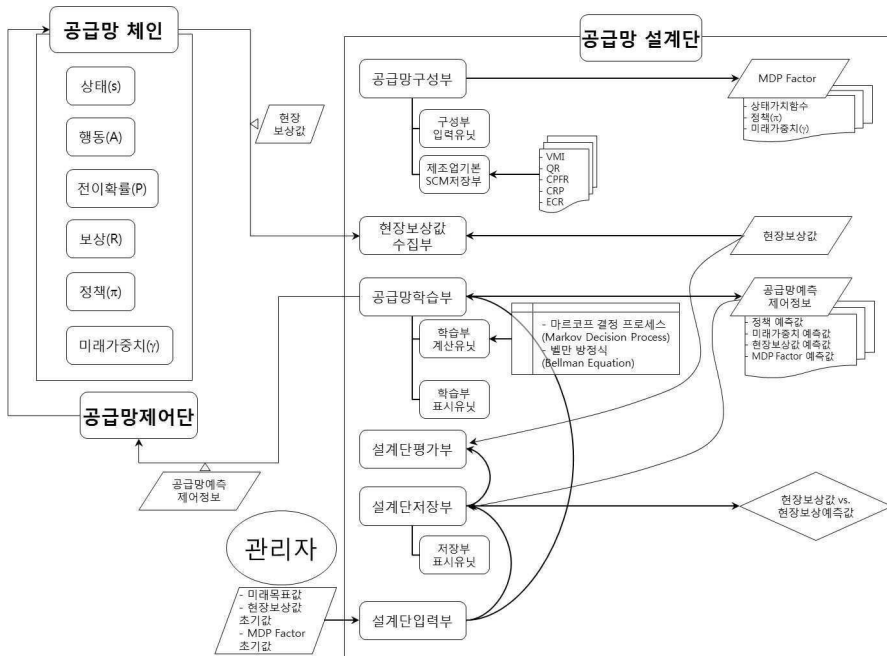
도면2



도면3



도면4



【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 청구항 1의 줄 4

【변경전】

상기 완제품의

【변경후】

상기 완제품의

【직권보정 2】

【보정항목】 청구범위

【보정세부항목】 청구항 1의 줄 8

【변경전】

하기 수학식1을 만족하는

【변경후】

하기 수학식을 만족하는

【식권보정 3】

【보정항목】 청구범위

【보정세부항목】 청구항 1의 줄 8

【변경전】

하기 수학식2를 만족하는

【변경후】

하기 수학식을 만족하는

【식권보정 4】

【보정항목】 청구범위

【보정세부항목】 청구항 1의 줄 9

【변경전】

하기 수학식3을 만족하는

【변경후】

하기 수학식을 만족하는

【식권보정 5】

【보정항목】 청구범위

【보정세부항목】 청구항 3의 줄 3

【변경전】

상기 현장보상예측값을 비교하여

【변경후】

상기 현장보상값예측값을 비교하여

【식권보정 6】

【보정항목】 청구범위

【보정세부항목】 청구항 15의 줄 5

【변경전】

공급 체인 최적화 설계 시스템 설계 방법.

【변경후】

공급 체인 최적화 설계 방법.

【식권보정 7】

【보정항목】 청구범위

【보정세부항목】 청구항 9의 줄 9

【변경전】

하기의 수학식1을 만족하는

【변경후】

하기의 수학식을 만족하는

【식권보정 8】

【보정항목】 청구범위

【보정세부항목】 청구항 9의 줄 10

【변경전】

하기의 수학식2를 만족하는

【변경후】

하기의 수학식을 만족하는

【직권보정 9】

【보정항목】 청구범위

【보정세부항목】 청구항 9의 줄 10

【변경전】

하기의 수학식3을 만족하는

【변경후】

하기의 수학식을 만족하는

【직권보정 10】

【보정항목】 청구범위

【보정세부항목】 청구항 11의 줄 4

【변경전】

상기 현장보상예측값을 비교하여

【변경후】

상기 현장보상값예측값을 비교하여

【직권보정 11】

【보정항목】 청구범위

【보정세부항목】 청구항 9의 줄 4

【변경전】

상기 완제품의

【변경후】

상기 완제품의