



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0074640
(43) 공개일자 2025년05월27일

(51) 국제특허분류(Int. Cl.)
G06Q 50/08 (2012.01) G06N 3/092 (2023.01)
G06Q 10/0633 (2023.01) G06Q 10/0637 (2023.01)
(52) CPC특허분류
G06Q 50/08 (2013.01)
G06N 3/092 (2023.01)
(21) 출원번호 10-2024-0166227
(22) 출원일자 2024년11월20일
심사청구일자 2024년11월20일
(30) 우선권주장
1020230161646 2023년11월20일 대한민국(KR)

(71) 출원인
인하대학교 산학협력단
인천광역시 미추홀구 인하로 100(용현동, 인하대학교)
(72) 발명자
신도형
서울특별시 강남구 삼성로 14, 404동 1701호 (개포동, 개포자이 프레지던스)
장민서
경기도 고양시 덕양구 화신로 36, 703동 1409호 (행신동, 소만마을7단지아파트)
(뒷면에 계속)
(74) 대리인
특허법인태백

전체 청구항 수 : 총 14 항

(54) 발명의 명칭 강화학습 모델을 기반으로 하는 공정계획 설계 장치 및 그 방법

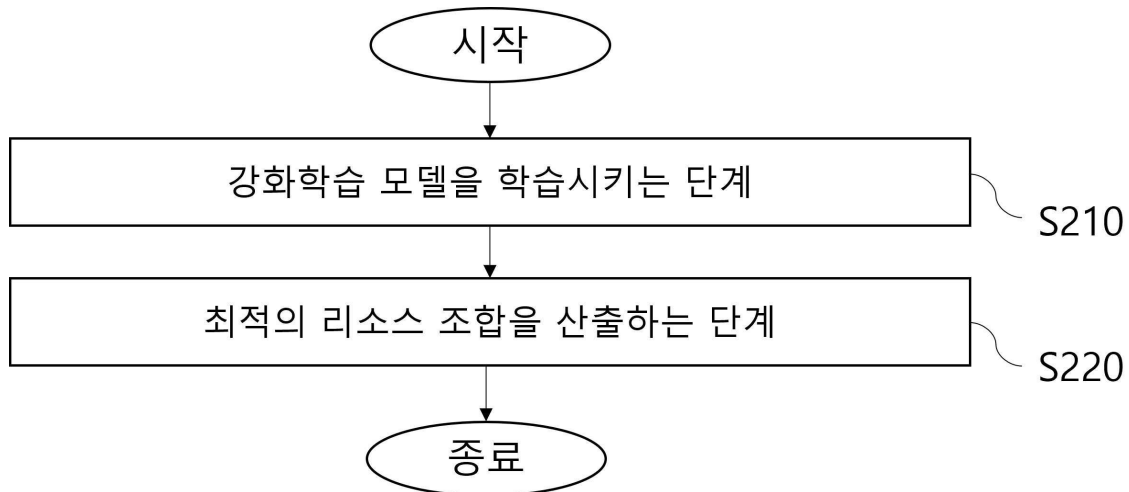
(57) 요약

본 발명은 강화학습 모델을 기반으로 하는 공정계획 설계 장치 및 그 방법에 관한 것이다.

본 발명에 따르면, 강화학습 모델을 기반으로 하는 공정계획 설계 방법에 있어서 건설 공정과 관련하여 건설 현장의 환경조건, 복수의 리소스 조합을 입력데이터로 설정하고 효율성을 출력데이터로 설정하여 강화학습 모델을

(뒷면에 계속)

대표도 - 도2



학습시키는 단계, 그리고 상기 강화학습 모델에 건설 현장의 환경조건을 입력하면 최적의 리소스 조합을 산출하는 단계를 포함하며, 상기 강화학습 모델을 학습시키는 단계는, 복수의 리소스 조합에 대하여 효율성 순위를 결정하고, 상위 n개의 효율성 순위를 가지는 리소스 조합에 대하여 순위별로 가중치를 차등 부여한다.

본 발명에 따르면, 본 발명의 일 실시예에 따르면, 강화학습 모델 기반의 공정계획을 최적화하는 과정에서 국부 최적화를 방지할 수 있고, 필요 반복 횟수의 감소로 빠른 최적 공기, 비용 도출로 인해 공정계획 최적화 시간을 단축할 수 있으며 최적화 성능을 향상시킬 수 있다.

(52) CPC특허분류

G06Q 10/0633 (2023.01)

G06Q 10/0637 (2023.01)

(72) 발명자

이제섭

인천광역시 미추홀구 용정공원로83번길 43, 102동 2405호 (용현동)

한제희

인천광역시 서구 청라한대로 40, 592동 2203호 (청라동, 청라 호수공원 한신더휴)

한승우

인천광역시 연수구 신송로82번길 6, 402동 803호 (송도동, 송도풍림아이원4단지아파트)

고용호

서울특별시 서초구 남부순환로 2567, 1402호 (서초동, 코네스트)

이 발명을 지원한 국가연구개발사업

과제고유번호	1615013119
과제번호	KA158708
부처명	국토교통부
과제관리(전문)기관명	국토교통과학기술진흥원
연구사업명	스마트건설기술개발사업
연구과제명	도로실증을 통한 스마트 건설기술 개발
과제수행기관명	한국도로공사
연구기간	2023.01.01 ~ 2023.12.31

명세서

청구범위

청구항 1

강화학습 모델을 기반으로 하는 공정계획 설계 장치에 의해 수행되는 공정계획 설계 방법에 있어서,

건설 공정과 관련하여 건설 현장의 환경조건, 복수의 리소스 조합을 입력데이터로 설정하고 효율성을 출력데이터로 설정하여 강화학습 모델을 학습시키는 단계, 그리고

상기 강화학습 모델에 건설 현장의 환경조건을 입력하면 최적의 리소스 조합을 산출하는 단계를 포함하며,

상기 강화학습 모델을 학습시키는 단계는,

복수의 리소스 조합에 대하여 효율성 순위를 결정하고, 상위 n 개의 효율성 순위를 가지는 리소스 조합에 대하여 순위별로 가중치를 차등 부여하는 공정계획 설계 방법.

청구항 2

제1항에 있어서,

상기 리소스 조합은,

공사 수행에 들어가는 인력, 장비, 자재 중에서 적어도 하나를 포함하는 공정계획 설계 방법.

청구항 3

제1항에 있어서,

상기 리소스 조합은,

덤프트럭과 굴삭기를 포함하는 공정계획 설계 방법.

청구항 4

제1항에 있어서,

상기 효율성은,

공사 수행에 따른 생산량에 비례하고, 비용에 반비례하는 공정계획 설계 방법.

청구항 5

제3항에 있어서,

상기 건설 현장의 환경조건은,

굴삭기의 용량, 덤프트럭의 용량, 작업효율, 적재 속도, 흙의 상태, 덤프 운반거리, 덤프 운반속도, 건설 현장의 면적, 건설 현장의 형태, 건설 현장의 진입로의 도로폭, 처리해야 하는 물량, 기상 상태 중 적어도 하나를 포함하는 공정계획 설계 방법.

청구항 6

제1항에 있어서,

상기 강화학습 모델을 학습시키는 단계는,

복수의 리소스 조합에 대하여 효율성 순위가 가장 높은 리소스 조합에 대하여 가중치를 10배로 적용하고, 효율성 순위가 2순위인 리소스 조합에 대하여 가중치를 4배로 적용하여 국부 최적화를 방지하는 공정계획 설계 방법.

청구항 7

제1항에 있어서,

상기 강화학습 모델을 학습시키는 단계는,

상기 강화학습 모델을 통해 리소스 조합이 선택될때까지 학습과정을 반복하며,

최종 결정 상태에 도달했을 때, 상기 강화학습 모델이 선택한 환경조건 별 리소스 조합에 따른 총 공기(工期) 및 비용을 산출하고, 이를 목표 공기 및 비용과 비교하며,

상기 산출된 리소스 조합에 따른 총 공기 및 비용과 목표 공기 및 비용을 비교한 결과를 토대로 보상을 제공하여 각 환경조건 별 리소스 선택에 대한 보상이 이루어지도록 상기 강화학습 모델을 학습시키는 공정계획 설계 방법.

청구항 8

강화학습 모델을 기반으로 하는 공정계획 설계 장치에 있어서,

건설 공정과 관련하여 건설 현장의 환경조건, 복수의 리소스 조합을 입력데이터로 설정하고 효율성을 출력데이터로 설정하여 강화학습 모델을 학습시키는 학습부, 그리고

상기 강화학습 모델에 건설 현장의 환경조건을 입력하여 최적의 리소스 조합을 산출하는 제어부를 포함하며,

상기 학습부는,

복수의 리소스 조합에 대하여 효율성 순위를 결정하고, 상위 n 개의 효율성 순위를 가지는 리소스 조합에 대하여 순위별로 가중치를 차등 부여하는 공정계획 설계 장치.

청구항 9

제8항에 있어서,

상기 리소스 조합은,

공사 수행에 들어가는 인력, 장비, 자재 중에서 적어도 하나를 포함하는 공정계획 설계 장치.

청구항 10

제8항에 있어서,

상기 리소스 조합은,

덤프트럭과 굴삭기를 포함하는 공정계획 설계 장치.

청구항 11

제8항에 있어서,

상기 효율성은,

공사 수행에 따른 생산량에 비례하고, 비용에 반비례하는 공정계획 설계 장치.

청구항 12

제10항에 있어서,

상기 건설 현장의 환경조건은,

굴삭기의 용량, 덤프트럭의 용량, 작업효율, 적재 속도, 흙의 상태, 덤프 운반거리, 덤프 운반속도, 상기 건설 현장의 면적, 상기 건설 현장의 형태, 상기 건설 현장의 진입로의 도로폭, 처리해야 하는 물량, 기상 상태 중에서 적어도 하나를 포함하는 공정계획 설계 장치.

청구항 13

제8항에 있어서,

상기 학습부는,

복수의 리소스 조합에 대하여 효율성 순위가 가장 높은 리소스 조합에 대하여 가중치를 10배로 적용하고, 효율성 순위가 2순위인 리소스 조합에 대하여 가중치를 4배로 적용하여 국부 최적화를 방지시키는 공정계획 설계 장치.

청구항 14

제8항에 있어서,

상기 학습부는,

상기 강화학습 모델을 통해 리소스 조합이 선택될때까지 학습과정을 반복하며,

최종 결정 상태에 도달했을 때, 상기 강화학습 모델이 선택한 환경조건 별 리소스 조합에 따른 총 공기(工期) 및 비용을 산출하고, 이를 목표 공기 및 비용과 비교하며,

상기 산출된 리소스 조합에 따른 총 공기 및 비용과 목표 공기 및 비용을 비교한 결과를 토대로 보상을 제공하여 각 환경조건 별 리소스 선택에 대한 보상이 이루어지도록 상기 강화학습 모델을 학습시키는 공정계획 설계 장치.

발명의 설명

기술 분야

[0001] 본 발명은 강화학습 모델을 기반으로 하는 공정계획 설계 장치 및 그 방법에 관한 것으로, 좀 더 상세하게는 강화학습 모델 기반의 공정계획을 최적화하는 과정에서 발생할 수 있는 국부 최적화를 방지하며 최적화 성능을 향상시키는 강화학습 모델을 기반으로 하는 공정계획 설계 장치 및 그 방법에 관한 것이다.

배경 기술

[0002] 일반적으로 공정계획 수립이란 설계, 시공, 원가 등의 지식을 기반으로 계약문서와 원가를 분석하여 공사의 목적물을 계약된 공기 내에 완성하기 위한 일정계획을 수립하는 것이다.

[0003] 하지만 기존 공정계획 기술의 경우, 주로 공정관리자의 경험이나 기업체의 능력을 따르게 되며, 실제 상황에서 일어날 수 있는 리스크 등의 상황이 제대로 반영되지 못한다는 특성이 있으며, 이로 인해, 기존의 공정계획을 통해 얻은 예측 공기와 비용은 그 정확도가 떨어진다는 한계가 존재한다.

[0004] 또한 건설 산업의 경우 생산성이 낮은 상황에서 고령화 및 숙련인력감소가 빠르게 진행되고 있어 디지털화 및 자동화 필요성이 증대되고 있다.

[0005] 특히, 1년간의 공정계획을 보여주는 연간공정표만 하더라도 수백 개 이상의 공정으로 구성되게 되는데, 이러한 공정들의 자원최적화에는 상당한 노력과 시간이 소요된다는 문제가 있다.

[0006] 이를 해결하기 위해서 국내외에서 다양한 공정계획 시뮬레이션이 개발되었다. [리스크 발생 확률 및 DEAS 데이터를 반영한 AI 기반 공정계획 최적화 방법, 출원번호 10-2021-0149409]에서는 리스크에 따른 공정계획 불확실성과 최적의 장비 조합 도출을 반영한 시뮬레이션 방법이 제시되었다.

[0007] 이 기술에서 공사 수행에 들어가는 인력, 장비, 자재 등의 리소스(resource)에 대한 행동(Action)을 취하는 과정[도 3a]에서 국부 최적화가 발생할 가능성이 있고 수많은 반복(iteration)으로 인하여 공정계획 최적화 시간이 길어질 수 있다는 한계가 있다.

[0008] 이러한 공정계획 기술과 관련된 선행기술로 공개특허 제10-2013-0049391호(참고문헌 1), 등록특허 제10-0191405호(참고문헌 2), 출원번호 10-2021-0149409(참고문헌 3)가 있다.

발명의 내용

해결하려는 과제

[0009] 본 발명은 이러한 문제점들을 해결하기 위한 것으로서 강화학습 모델 기반의 공정계획을 최적화하는 과정에서 발생할 수 있는 국부 최적화를 방지하며 최적화 성능을 향상시키는 강화학습 모델을 기반으로 하는 공정계획 설계 장치 및 그 방법을 제공하는데 그 목적이 있다.

과제의 해결 수단

[0010] 본 발명의 실시예에 따른, 강화학습 모델을 기반으로 하는 공정계획 설계 방법에 있어서 건설 공정과 관련하여 건설 현장의 환경조건, 복수의 리소스 조합을 입력데이터로 설정하고 효율성을 출력데이터로 설정하여 강화학습 모델을 학습시키는 단계, 그리고 상기 강화학습 모델에 건설 현장의 환경조건을 입력하면 최적의 리소스 조합을 산출하는 단계를 포함하며, 상기 강화학습 모델을 학습시키는 단계는, 복수의 리소스 조합에 대하여 효율성 순위를 결정하고, 상위 n개의 효율성 순위를 가지는 리소스 조합에 대하여 순위별로 가중치를 차등 부여한다.

[0011] 상기 리소스 조합은 공사 수행에 들어가는 인력, 장비, 자재 중에서 적어도 하나를 포함할 수 있다.

[0012] 상기 리소스 조합은 덤프트럭과 굴삭기를 포함할 수 있다.

[0013] 상기 효율성은 공사 수행에 따른 생산량에 비례하고, 비용에 반비례할 수 있다.

[0014] 상기 건설 현장의 환경조건은 굴삭기, 덤프트럭의 용량, 작업효율, 적재 속도, 흙의 상태, 덤프 운반거리, 덤프 운반속도, 상기 건설 현장의 면적, 상기 건설 현장의 형태, 상기 건설 현장의 진입로의 도로폭, 처리해야 하는 물량, 기상 상태 중에서 적어도 하나를 포함할 수 있다.

[0015] 상기 강화학습 모델을 학습시키는 단계는 복수의 리소스 조합에 대하여 효율성 순위가 가장 높은 리소스 조합에 대하여 가중치를 10배로 적용하고, 효율성 순위가 2순위인 리소스 조합에 대하여 가중치를 4배로 적용하여 국부 최적화를 방지할 수 있다.

[0016] 상기 강화학습 모델을 학습시키는 단계는 상기 강화학습 모델을 통해 리소스 조합이 선택될때까지 학습과정을 반복하며 최종 결정 상태에 도달했을 때, 상기 강화학습 모델이 선택한 환경조건 별 리소스 조합에 따른 총 공기(工期) 및 비용을 산출하고, 이를 목표 공기 및 비용과 비교하며, 상기 산출된 리소스 조합에 따른 총 공기 및 비용과 목표 공기 및 비용을 비교한 결과를 토대로 보상을 제공하여 각 환경조건 별 리소스 선택에 대한 보상이 이루어지도록 상기 강화학습 모델을 학습시킬 수 있다.

[0017] 본 발명의 다른 실시예에 따르면, 강화학습 모델을 기반으로 하는 공정계획 설계 장치에 있어서, 건설 공정과 관련하여 건설 현장의 환경조건, 복수의 리소스 조합을 입력데이터로 설정하고 효율성을 출력데이터로 설정하여 강화학습 모델을 학습시키는 학습부, 그리고 상기 강화학습 모델에 건설 현장의 환경조건을 입력하여 최적의 리소스 조합을 산출하는 제어부를 포함하며, 상기 학습부는 복수의 리소스 조합에 대하여 효율성 순위를 결정하고, 상위 n개의 효율성 순위를 가지는 리소스 조합에 대하여 순위별로 가중치를 차등 부여한다.

발명의 효과

[0018] 본 발명에 따르면, 강화학습 모델 기반의 공정계획을 최적화하는 과정에서 국부 최적화를 방지할 수 있고, 필요 반복 횟수의 감소로 빠른 최적 공기, 비용 도출로 인해 공정계획 최적화 시간을 단축할 수 있으며 최적화 성능을 향상시킬 수 있다.

도면의 간단한 설명

[0019] 도 1은 본 발명의 일 실시예에 따른 강화학습 모델을 기반으로 하는 공정계획 설계 장치의 구성도이다.

도 2는 본 발명의 일 실시예에 따른 강화학습 모델을 기반으로 하는 공정계획 설계 방법의 흐름도이다.

도 3은 강화학습 모델의 일반적 구조를 도시한 도면이다.

도 4는 본 발명에 활용된 강화학습 모델의 학습을 위한 공정표 상 강화학습 필요 요소의 정의 예를 도시한 도면이다.

도 5a 내지 도 5d는 본 발명에 따른 강화학습을 활용한 AI 기반 건설계획 시뮬레이션 모델 학습 수행 순서를 순차적으로 도시한 도면들이다.

도 6은 국부 최적화와 전역 최적화의 예시를 도시한 도면이다.

도 7은 국부 최적화 방지 방법 적용을 위한 효율(ratio) 및 순위(rank) 책정 방식 설명을 위해 도시한

도면이다.

발명을 실시하기 위한 구체적인 내용

- [0020] 아래에서는 첨부한 도면을 참조하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 본 발명의 실시 예를 상세히 설명한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.
- [0021] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.
- [0022] 첨부한 도면을 참고로 하여 본 발명의 실시 예에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다.
- [0023] 이하에서는 도 1을 참조하여 강화학습 모델을 기반으로 하는 공정계획 설계 장치에 대해 상세하게 설명한다.
- [0024] 도 1은 본 발명의 일 실시예에 따른 강화학습 모델을 기반으로 하는 공정계획 설계 장치의 구성도이다.
- [0025] 도 1에 도시된 바와 같이, 강화학습 모델을 기반으로 하는 공정계획 설계 장치(100)는 학습부(110) 및 제어부(120)를 포함한다.
- [0026] 학습부(110)는 건설 공정과 관련하여 건설 현장의 환경조건, 복수의 리소스 조합을 입력데이터로 설정하고 효율성을 출력데이터로 설정하여 강화학습 모델을 학습시킨다.
- [0027] 이때, 효율성은 공사 수행에 따른 생산량에 비례하고, 비용에 반비례할 수 있다.
- [0028] 좀 더 상세하게는, 학습부(110)는 강화학습 모델을 통해 리소스 조합이 선택될때까지 학습과정을 반복하며, 최종 결정 상태에 도달했을 때, 상기 강화학습 모델이 선택한 환경조건 별 리소스 조합에 따른 총 공기 및 비용을 산출하고, 이를 목표 공기 및 비용과 비교하며, 상기 산출된 리소스 조합에 따른 총 공기 및 비용과 목표 공기 및 비용을 비교한 결과를 토대로 보상을 제공하여 각 환경조건 별 리소스 선택에 대한 보상이 이루어지도록 강화학습 모델을 학습시킨다.
- [0029] 또한, 학습부(110)는 복수의 리소스 조합에 대하여 효율성 순위를 결정하고, 상위 n개의 효율성 순위를 가지는 리소스 조합에 대하여 순위별로 가중치를 차등 부여한다.
- [0030] 좀 더 상세하게는, 학습부(110)는 복수의 리소스 조합에 대하여 효율성 순위가 가장 높은 리소스 조합에 대하여 가중치를 10배로 적용하고, 효율성 순위가 2순위인 리소스 조합에 대하여 가중치를 4배로 적용하여 국부 최적화를 방지한다.
- [0031] 이때, 리소스 조합은 공사 수행에 들어가는 인력, 장비, 자재 중에서 적어도 하나를 포함할 수 있고, 이 중, 장비는 덤프트럭과 굴삭기를 포함할 수 있다.
- [0032] 또한, 건설 현장의 환경조건은 굴삭기, 덤프트럭의 용량(0.8m³, 1.0m³/15ton, 25ton), 작업효율, 적재 속도, 흙의 상태(모래or암상토), 덤프 운반거리, 덤프 운반속도, 건설 현장의 면적, 건설 현장의 형태, 건설 현장의 진입로의 도로폭, 처리해야 하는 물량, 기상 상태 중 적어도 하나를 포함할 수 있다.
- [0033] 도 2는 본 발명의 일 실시예에 따른 강화학습 모델을 기반으로 하는 공정계획 설계 방법의 흐름도이고, 도 3은 강화학습 모델의 일반적 구조를 도시한 도면이다.
- [0034] 도 2에 도시된 바와 같이, 본 발명의 일 실시예에 따른 강화학습 모델을 기반으로 하는 공정계획 설계 방법에서는, 먼저, 공정계획 설계 장치(100)는 건설 공정과 관련하여 건설 현장의 환경조건, 복수의 리소스 조합을 입력데이터로 설정하고 효율성을 출력데이터로 설정하여 강화학습 모델을 학습시킨다(S210).
- [0035] 이때, 효율성은 공사 수행에 따른 생산량에 비례하고, 비용에 반비례할 수 있다.
- [0036] 학습이 완료된 후, 공정계획 설계 장치(100)는 강화학습 모델에 건설 현장의 환경조건을 입력하면 최적의 리소스 조합을 산출한다(S220).
- [0037] 강화학습 모델에 대해 도 3을 참조하여 좀 더 상세하게 설명하면, 본 발명의 학습부(110)는 에이전트(Agent)(111)를 강화 학습시키는 방법으로 딥러닝을 수행한다.

- [0038] 에이전트(111)는 자유롭게 여러 가지 행동(action)을 취할 수 있으며, 해당 행동은 에이전트(111)가 처해진 환경(Environment)(112)의 상태(State)를 파악하여 결정된다.
- [0039] 여기서, 에이전트(111)가 취하는 행동에 따라 에이전트(111)는 보상(Reward) 또는 페널티(Penalty)를 받게 되며, 에이전트(111)가 보상이 높아지는 방향으로 학습하도록 학습의 방향을 결정시키게 되며, 이러한 강화학습 모델에 의한 학습방식은 매 상황별로 달라지는 시나리오(scenario)가 존재할 때 활용하는데 가장 적합하다.
- [0040] 한편, 도 4는 본 발명에 활용된 강화학습을 활용한 AI 기반 공정계획 시뮬레이션 모델의 학습을 위한 공정표 상 강화학습 필요 요소의 정의 예를 도시한 도면이다.
- [0041] 이에 의하면, 공정계획 시뮬레이션에서 공정표 및 공정의 흐름은 환경(112)에 해당한다고 볼 수 있으며, 공정표의 액티비티(Activity)는 환경(112)에서 나타나는 상태 정보로 활용할 수 있다.
- [0042] 에이전트(111)가 공정의 흐름에서 나타나는 액티비티에 들어갈 수 있는 공정 별 리소스 조합의 선택을 행동이라고 생각한다면, 강화학습의 에이전트(111)가 행동할 수 있는 조건은 모두 갖춰지게 된다.
- [0043] 본 발명에서는, 에이전트(111)를 강화학습 모델, 환경(112)을 공정계획 시뮬레이션, 행동을 강화학습 모델이 공정의 흐름에서 나타나는 액티비티에 들어갈 수 있는 공정 별 리소스 조합의 선택으로 하여 강화학습을 수행한다.
- [0044] 이때, 리소스 조합은 공사 수행에 들어가는 인력, 장비, 자재 중에서 적어도 하나를 포함할 수 있고, 이 중, 장비는 덤프트럭과 굴삭기를 포함할 수 있다.
- [0045] 또한, 건설 현장의 환경조건은 굴삭기의 용량, 덤프트럭의 용량(0.8m³, 1.0m³/15ton, 25ton), 작업효율, 적재 속도, 흙의 상태(모래or암상토), 덤프 운반거리, 덤프 운반속도, 건설 현장의 면적, 건설 현장의 형태, 건설 현장의 진입로의 도로폭, 처리해야 하는 물량, 기상 상태 중 적어도 하나를 포함할 수 있다.
- [0046] 이러한 특정 토공사의 공정조건과 장비조건 등에 따라, 굴삭기의 깎고 적재하는 속도가 달라지게 되며, 덤프트럭의 용량과 운반거리에 따라 굴삭기와 덤프트럭이 대기하거나 쉬는 시간이 발생하지 않도록 하는 장비 대수의 조합이 달라지게 된다.
- [0047] 또한, 실제 공정 관리자가 공정계획 시뮬레이션을 이용한다고 가정할 때, 목적 공기(工期) 및 비용의 정보는 취득 가능한 정보이므로, 하나의 시나리오가 끝난 뒤 예측된 공기 및 비용 정보와 목적 공기 및 비용 사이의 오차가 얻어질 수 있다.
- [0048] 그리고, 이러한 예측된 공기 및 비용 정보가 목적 공기 및 비용보다 줄어드는 수준에 따라 양(+)의 보상, 증가하는 수준에 따라 음(-)의 보상을 주게 된다면, 강화학습 모델이 행하는 행동에 따른 보상을 적절하게 정의할 수 있다.
- [0049] 이러한 공정계획 시뮬레이션을 위한 강화학습 모델이 수행될 경우, 실제 시공 중 예측하지 못한 이벤트(event)가 발생했을 때, 해당 이벤트 발생 시점 이후의 후속 공정에 대한 빠른 최적화 조치 또한 쉽게 가능해진다.
- [0050] 이하에서는 도 5a 내지 도 7을 참조하여 본 방법의 강화학습을 통하여 모델을 학습시키는 방법을 상세히 설명한다.
- [0051] 먼저, 도 5a 내지 도 5d는 본 발명에 따른 강화학습을 활용한 AI 기반 공정계획 시뮬레이션 모델 학습 수행 순서를 순차적으로 도시한 도면들이다.
- [0052] 여기서, 도 5a 내지 도 5d는 시뮬레이션 모델의 학습 구동 사항의 이해를 돕기 위하여 공사 수행에 들어가는 인력, 장비, 자재 등의 리소스에 대한 행동만을 취하는 것으로 나타낸 것이다.
- [0053] 이에 의하면 도 5a 내지 도 5c에서와 같이, 상태가 '1'에서 '3'으로 이동함에 따라 강화학습 모델은 액티비티 별 행동 리스트(Action list)에서 적절한 행동을 고르도록 학습된다.
- [0054] 즉, 모델은 공정계획 시뮬레이션에서 공정표의 상태에 따른 액티비티 별 행동 리스트에서 가장 높은 보상을 얻을 수 있는 행동을 고르도록 학습된다.
- [0055] 좀 더 상세하게는, 먼저, 모델은 공정표의 상태에 따른 액티비티 별 행동 리스트에서 해당 상태일 때 가장 적절한 리소스 조합을 고르도록 학습하며, 이를 모든 상태의 리소스 조합이 선택될 때까지 반복한다.
- [0056] 예를 들어, 도 5a는 상태가 '1'로서 액티비티 #1의 행동 리스트 'A1'은 "1) 1 굴삭기 + 2 덤프트럭", "2) 2 굴

삭기 + 5 덤프트럭" 및 "3) 3 굴삭기 + 8 덤프트럭"임을 알 수 있다.

- [0057] 여기서, 강화학습 모델은 "상태=1"에서 "행동=3"을 선택하면 액티비티 #1의 행동 리스트 'A1'의 행동 "3) 3 굴삭기 + 8 덤프트럭"을 적절한 행동으로 고르도록 학습한다.
- [0058] 다음으로 도 5b는 상태가 '2'로서 액티비티 #2의 행동 리스트 'A2'는 "1) 1 굴삭기 + 2 덤프트럭", "2) 2 굴삭기 + 5 덤프트럭" 및 "3) 3 굴삭기 + 8 덤프트럭"임을 알 수 있다.
- [0059] 여기서, 강화학습 모델은 "상태=2"에서 "행동=2"를 선택하면 액티비티 #2의 행동 리스트 'A2'의 행동 "2) 2 굴삭기 + 5 덤프트럭"을 적절한 행동으로 고르도록 학습한다.
- [0060] 그리고, 도 5c는 상태가 '3'으로서 액티비티 #3의 행동 리스트 'A3'는 "1) 1 굴삭기 + 2 덤프트럭", "2) 2 굴삭기 + 5 덤프트럭" 및 "3) 3 굴삭기 + 8 덤프트럭"임을 알 수 있다. 여기서 강화학습 모델은 "상태=3"에서 "행동=2"를 선택하면 액티비티 #3의 행동 리스트 'A3'의 행동 "2) 2 굴삭기 + 5 덤프트럭"을 적절한 행동으로 고르도록 학습한다.
- [0061] 그리고, 마지막 상태인 도 5d는 상태가 '4(End)'임을 알 수 있다.
- [0062] 여기서, 강화학습 모델이 선택한 액티비티 별 리소스는 액티비티 #1의 행동 리스트 'A1'의 행동 "3) 3 굴삭기 + 8 덤프트럭", 액티비티 #2의 행동 리스트 'A2'의 행동 "2) 2 굴삭기 + 5 덤프트럭", 액티비티 #3의 행동 리스트 'A3'의 행동 "2) 2 굴삭기 + 5 덤프트럭"임을 알 수 있다.
- [0063] 즉, 강화학습 모델은 최종 결정 상태에 도달했을 때 강화학습 모델이 선택한 환경조건 별 리소스 조합에 따른 총 공기(工期) 및 비용을 산출하고, 이를 목표 공기 및 비용과 비교한다.
- [0064] 다음으로, 산출된 리소스 조합에 따른 총 공기 및 비용과 목표 공기 및 비용을 비교한 결과를 토대로 보상을 제 공하여 모델이 각 액티비티 별 리소스 선택에 대한 보상을 통해 학습하게 한다.
- [0065] 이와 같은 학습 과정에서, 강화학습 모델은 액티비티 별 적절한 리소스 조합을 도출할 수 있게 되므로, 액티비 티에 따라 최적화된 리소스가 반영된 공정표를 제공할 수 있다.
- [0066] 한편, 이와 같은 학습 과정에서 국부 최적화가 발생할 수 있다.
- [0067] 도 6은 국부 최적화와 전역 최적화의 예시를 도시한 도면이다.
- [0068] 국부 최적화란, 도 6을 참고하여 설명하면 표시된 세 점은 각 구간에서 가장 좋은 국부 최적해(Local Optimum)이며 가장 오른쪽 점은 전체에서 가장 좋은 전역 최적해(Global Optimum)이다.
- [0069] 만약 반복이 적다면 에이전트가 전역 최적해 구간을 경험하지 못하고 국부 최적해 구간만을 계속 학습하게 되어 국부 최적화가 발생할 가능성이 생긴다.
- [0070] 이러한 경우, 충분히 많은 반복횟수로 국부 최적화를 벗어날 수 있지만, 고려해야 하는 공정의 개수가 많거나 장비 조합의 경우가 많게 되면 필요로 하는 반복횟수는 기하급수적으로 증가하게 된다.
- [0071] 따라서, 본 발명에서는 국부 최적화를 방지하면서 반복횟수를 줄이기 위해 복수의 리소스 조합에 대하여 효율성 순위를 결정하고, 상위 n개의 효율성 순위를 가지는 리소스 조합에 대하여 순위별로 가중치를 차등 부여하여 효 율성이 높은 리소스 조합이 학습 과정에서 제외되지 않도록 한다.
- [0072] 도 7은 국부 최적화 방지 방법 적용을 위한 효율(ratio) 및 순위(rank) 책정 방식 설명을 위해 도시한 도면이다.
- [0073] 도 7을 참고하여 좀 더 상세하게 설명하면, 각 리소스 별로 단위 비용(total cost), 생산성(productivity) 정보 가 있다.
- [0074] 여기서 생산성은 높을수록 효율적이며 단위 비용은 낮을수록 효율적이므로 생산성과 단위비용의 관계를 통해 효 율(ratio)을 도출한다.
- [0075] 그 후, 리소스들 간의 효율 순위(rank)를 내림차순으로 책정하고, 순위에 따라 가중치를 차등 적용한다.
- [0076] 예를 들어, 상위 순위(예를 들어, 1, 2위)의 리소스의 확률에 가장 높은 가중치(예를 들어, 10배)를 적용하고, 중간 순위(예를 들어, 3~5위)의 리소스의 확률에 상위 순위의 가중치보다 낮은 배수의 가중치(예를 들어, 4배)를 적용할 수 있다.

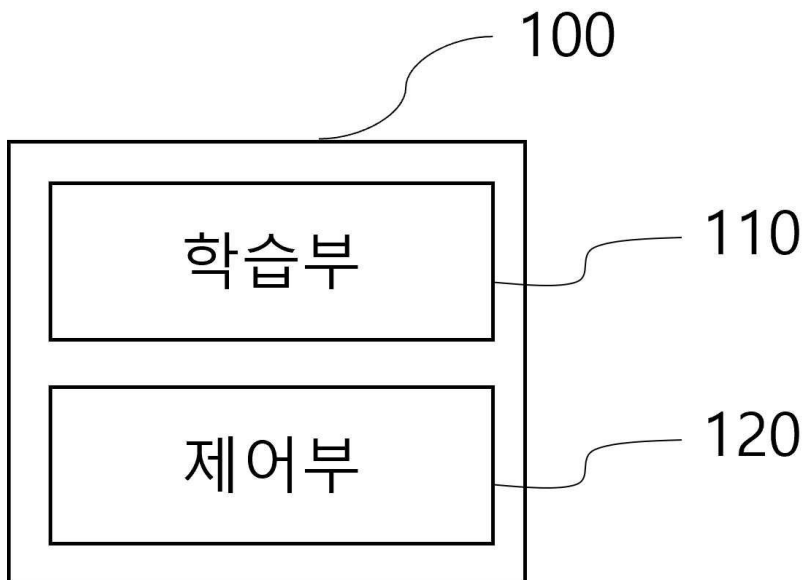
- [0077] 기존에는 처음에 모든 리소스들이 선택될 확률이 같았지만 이와 같은 방법을 활용하면 초기에 효율이 높은 리소스들이 선택될 확률이 높아지며 적은 반복으로도 전역 최적화를 달성할 수 있다.
- [0078] 도 5a 내지 도 5d에서 상기와 같은 방법을 적용해보면 다음과 같다.
- [0079] 액티비티 별로 행동 리스트를 나열하고 각 리소스 조합 별 효율을 분석하고 순위를 매긴다.
- [0080] 예를 들어 3) 3 굴삭기 + 8 덤프트럭이 1위, 2) 2 굴삭기 + 5 덤프트럭이 2위, 1) 1 굴삭기 + 2 덤프트럭이 3위로 책정된 경우, 3) 3 굴삭기 + 8 덤프트럭의 확률을 10배, 2 굴삭기 + 5 덤프트럭의 확률을 4배 적용시킨다.
- [0081] 본 발명의 일 실시예에 따르면, 강화학습 모델 기반의 공정계획을 최적화하는 과정에서 국부 최적화를 방지할 수 있고, 필요 반복 횟수의 감소로 빠른 최적 공기, 비용 도출로 인해 공정계획 최적화 시간을 단축할 수 있으며 최적화 성능을 향상시킬 수 있다.
- [0082] 본 발명은 도면에 도시된 실시 예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 다른 실시 예가 가능하다는 점을 이해할 것이다. 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 특허청구범위의 기술적 사상에 의하여 정해져야 할 것이다.

부호의 설명

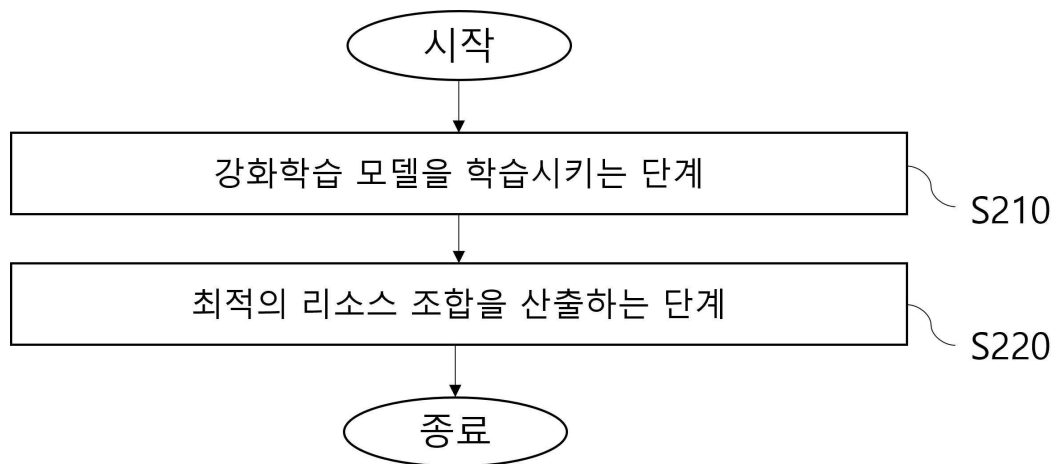
- [0083] 100: 공정계획 설계 장치
- 110: 학습부
- 111: 에이전트
- 112: 환경
- 120: 제어부

도면

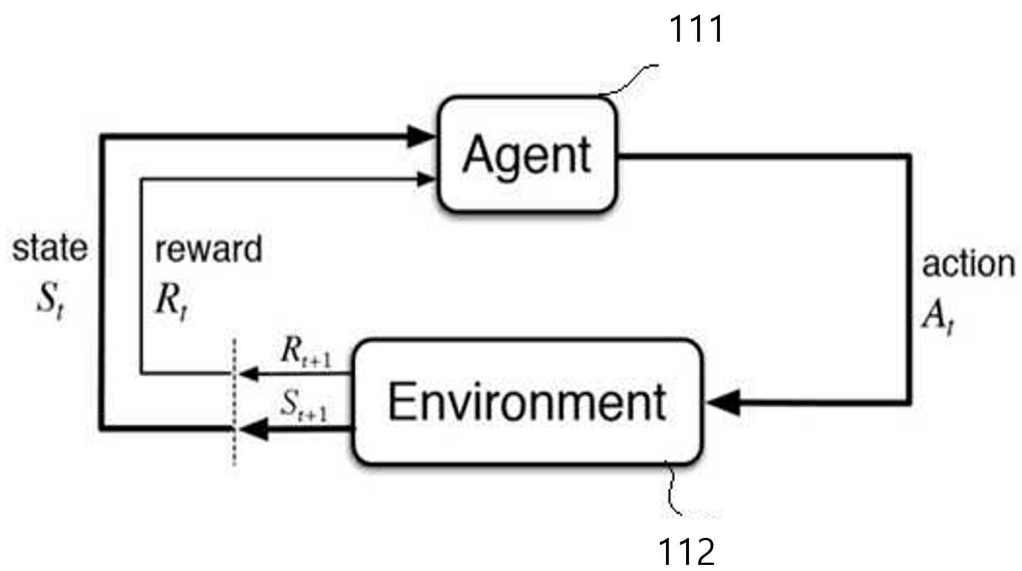
도면1



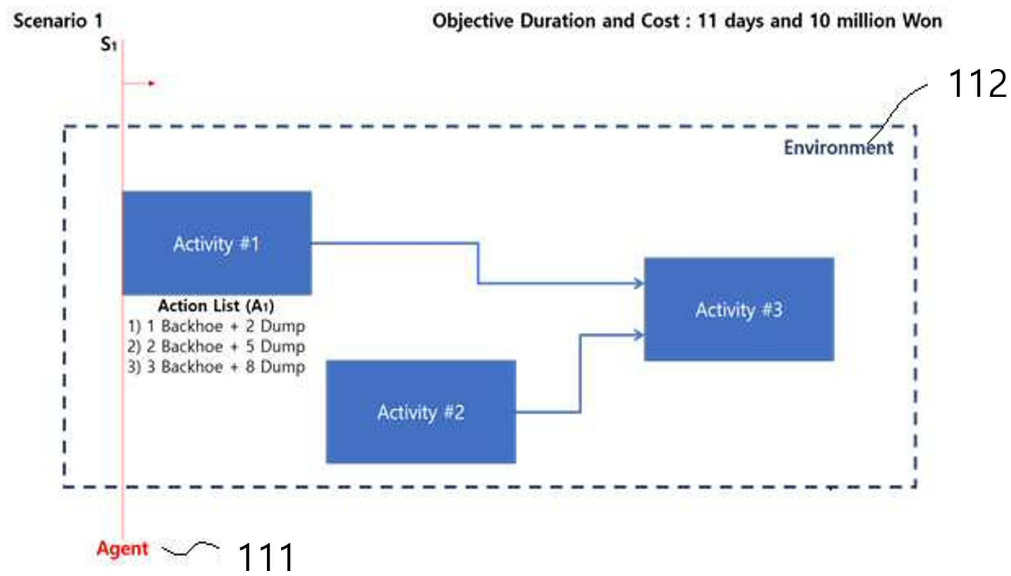
도면2



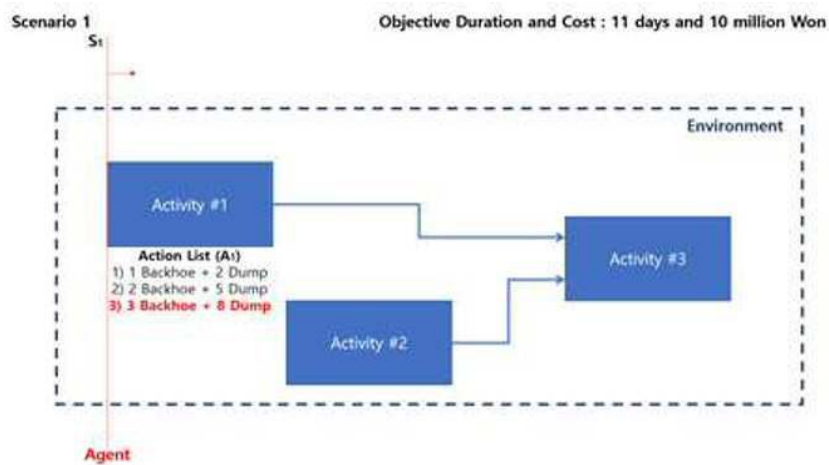
도면3



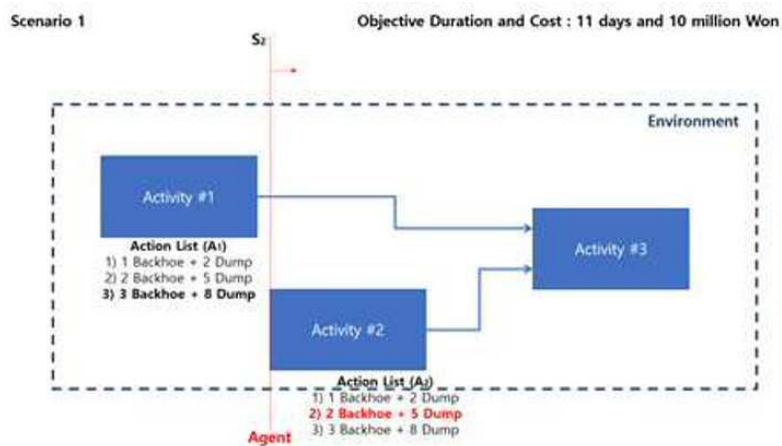
도면4



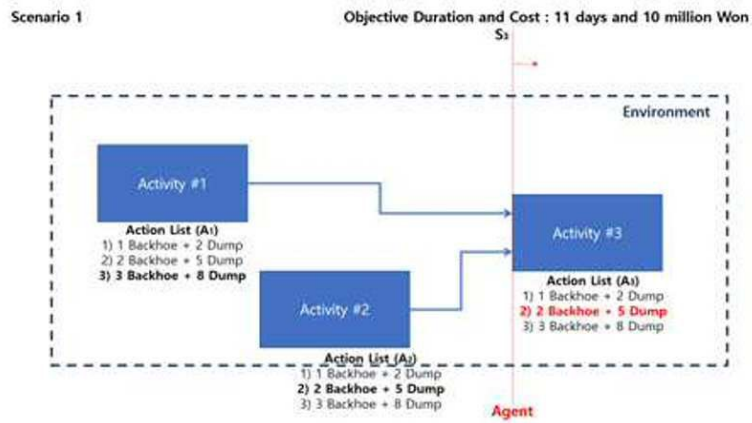
도면5a



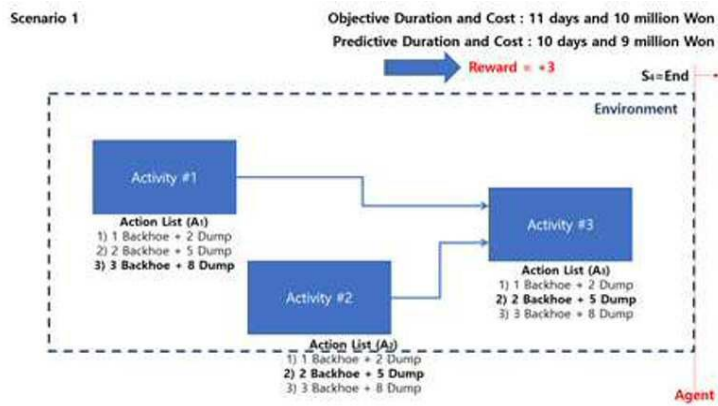
도면5b



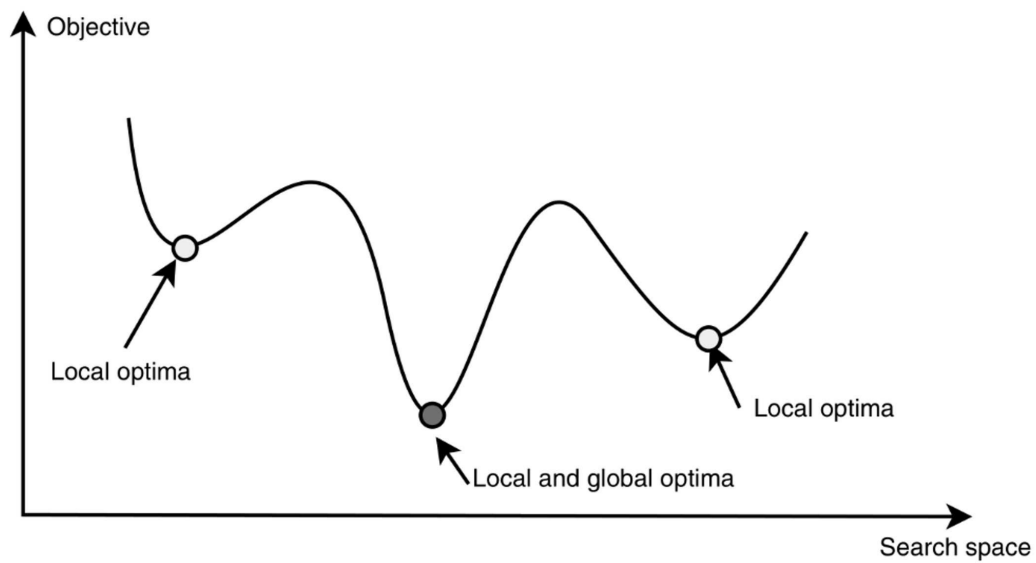
도면5c



도면5d



도면6



도면7

	Resource	Productivity [m3/day]	Total_cost [KRW/m3]	Ratio	Rank
Activity ##	0	678.72	8688.48	0.0781	7
	1	678.24	10017.25	0.0677	9
	2	1340.4	5389.61	0.2487	1
	3	1334.96	6083.54	0.2194	3
	4	678.24	10016.26	0.0677	8
	5	689.2	11158.54	0.0618	11
	6	1339.44	6062.69	0.2209	2
	7	1349.44	6682.52	0.2019	4
	8	690.72	11133.02	0.0620	10
	9	678.72	12651.53	0.0536	12
	10	1345.44	6701.89	0.2008	5
	11	1338.96	7404.28	0.1808	6