(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0175870
(43) 공개일자 2024년12월23일

(51) 국제특허분류(Int. Cl.)
G06Q 50/06 (2024.01) *G06N 3/092* (2023.01)
G06Q 10/06 (2012.01)
(52) CPC특허분류
G06Q 50/06 (2024.01)
G06N 3/092 (2023.01)
(21) 출원번호 10-2023-0076101
(22) 출원일자 2023년06월14일
심사청구일자 2023년06월14일

(71) 출원인
세종대학교산학협력단
서울특별시 광진구 능동로 209 (군자동, 세종대학교)
(72) 발명자
이현석
서울특별시 성동구 상원길 63, 107동 602호(성수동1가, 쌍용아파트)
이재훈
경기도 남양주시 오남읍 진건오남로 564
박지윤
광주광역시 광산구 수완로33번길 76, 101동 1001호 (수완동, 은빛마을 모아엘가)
(74) 대리인
특허법인위더피플

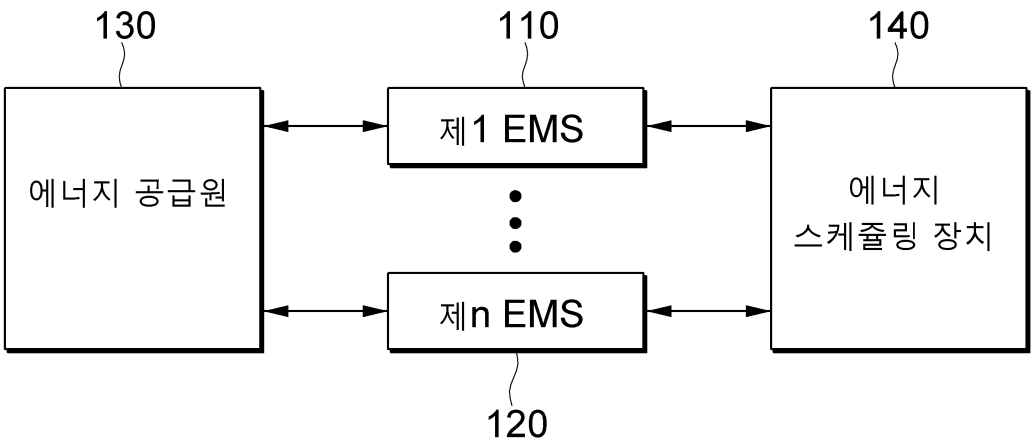
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 **연합강화학습을 이용한 다수 가구의 에너지 스케쥴링 방법 및 장치**

(57) 요 약

본 발명의 일 실시예에 따르면, 에너지 관리 장치는 에너지 관리 시스템(energy management system, EMS)의 에너지 수요량을 관리하도록 구성되는 수요량 관리부, 상기 EMS 내 재생 에너지원에서의 에너지 생산량을 관리하도록 구성되는 생산량 관리부, 상기 EMS 내 에너지 저장소의 에너지 잔여량을 관리하도록 구성되는 잔여량 관리부, 및 상기 EMS에 대한 상태 정보를 기반으로 에너지 스케쥴링 정책 모델에 대한 심층강화학습을 수행하여 에너지 공급원으로부터 공급받는 에너지 공급량을 관리하도록 구성되는 공급량 관리부를 포함하되, 상기 상태 정보는 단위 시간에서 상기 EMS의 에너지 수요량, 상기 단위 시간에서 상기 EMS의 에너지 생산량, 상기 단위 시간에서 상기 EMS의 에너지 잔여량, 상기 단위 시간에서 상기 에너지 공급원의 TOU(Time of Use) 요금 정보, 및 상기 EMS에 대한 DC(Demand Charge) 값을 포함할 수 있다.

대 표 도 - 도1



(52) CPC특허분류
G06Q 10/0631 (2023.01)
G06Q 10/0637 (2023.01)
Y04S 40/20 (2020.08)
Y04S 50/16 (2018.05)

이 발명을 지원한 국가연구개발사업
과제고유번호 1711180929
과제번호 2019M3F2A1073179
부처명 과학기술정보통신부
과제관리(전문)기관명 한국연구재단
연구사업명 에너지클라우드기술개발(과기정통부)(R&D)
연구과제명 효과적인 에너지 수요 패턴 및 요인 분석을 위한 앙상블 기법 기반의 XAI 에너지 플랫폼 개발
기 여 율 80/100
과제수행기관명 세종대학교
연구기간 2023.01.01 ~ 2023.12.31
이 발명을 지원한 국가연구개발사업
과제고유번호 1711193909
과제번호 IITP-2023-2021-0-01816
부처명 과학기술정보통신부
과제관리(전문)기관명 정보통신기획평가원
연구사업명 대학ICT연구센터육성지원사업
연구과제명 메타버스 자율트윈 핵심기술 연구
기 여 율 20/100
과제수행기관명 세종대학교 산학협력단
연구기간 2023.01.01 ~ 2023.12.31

명 세 서

청구범위

청구항 1

에너지 관리 장치에 있어서,

에너지 관리 시스템(energy management system, EMS)의 에너지 수요량을 관리하도록 구성되는 수요량 관리부;

상기 EMS 내 재생 에너지원에서의 에너지 생산량을 관리하도록 구성되는 생산량 관리부;

상기 EMS 내 에너지 저장소의 에너지 잔여량을 관리하도록 구성되는 잔여량 관리부; 및

상기 EMS에 대한 상태 정보를 기반으로 에너지 스케쥴링 정책 모델에 대한 심층강화학습을 수행하여 에너지 공급원으로부터 공급받는 에너지 공급량을 관리하도록 구성되는 공급량 관리부를 포함하되,

상기 상태 정보는,

단위 시간에서 상기 EMS의 에너지 수요량, 상기 단위 시간에서 상기 EMS의 에너지 생산량, 상기 단위 시간에서 상기 EMS의 에너지 잔여량, 상기 단위 시간에서 상기 에너지 공급원의 TOU(Time of Use) 요금 정보, 및 상기 EMS에 대한 DC(Demand Charge) 값을 포함하는, 에너지 관리 장치.

청구항 2

제1항에 있어서,

상기 DC 값은,

현재 주기에서 현재 시간까지 상기 에너지 공급원으로부터 공급받은 최대 에너지 공급량을 기반으로 도출되는, 에너지 관리 장치.

청구항 3

제1항에 있어서,

상기 공급량 관리부는,

제1 타임 슬롯에서 상기 에너지 공급원으로부터 공급받을 에너지 공급량을 선택하고, 상기 선택된 에너지 공급량과 상기 제1 타임 슬롯에서의 상기 EMS에 대한 상태 정보를 기반으로 상기 에너지 스케쥴링 정책 모델을 이용하여 상기 제1 타임 슬롯에서의 에너지 요금에 대한 정보 및 제2 타임 슬롯에서의 상기 EMS에 대한 상태 정보를 획득하는, 에너지 관리 장치.

청구항 4

제3항에 있어서,

상기 공급량 관리부는,

입실론 그리디(epsilon greedy) 정책을 기반으로 상기 제1 타임 슬롯에서 상기 에너지 공급원으로부터 공급받을 에너지 공급량을 선택하는, 에너지 관리 장치.

청구항 5

제3항에 있어서,

상기 공급량 관리부는,

상기 제1 타임 슬롯에서의 상기 EMS에 대한 상태 정보, 상기 제1 타임 슬롯에서 상기 에너지 공급원으로부터 공급받을 에너지 공급량, 상기 제1 타임 슬롯에서의 에너지 요금에 대한 정보 및 상기 제2 타임 슬롯에서의 상기 EMS에 대한 상태 정보를 기반으로 상기 에너지 스케쥴링 정책 모델을 업데이트하는, 에너지 관리 장치.

청구항 6

제3항에 있어서,

상기 에너지 스케쥴링 정책 모델에 대한 로컬 가중치 파라미터의 변화량 정보를 에너지 스케쥴링 장치로 전송하고, 상기 에너지 스케쥴링 장치로부터 상기 에너지 스케쥴링 장치의 중앙 에너지 스케쥴링 정책 모델에 대한 중앙 가중치 파라미터를 수신하도록 구성되는 통신부를 더 포함하고,

상기 공급량 관리부는,

상기 중앙 가중치 파라미터를 이용하여 상기 로컬 가중치 파라미터를 업데이트하는, 에너지 관리 장치.

청구항 7

제1항에 있어서,

상기 에너지 스케쥴링 정책 모델은,

에너지 스케쥴링 정책의 심층강화학습을 위한 메인 심층신경망 및 목표 에너지 요금의 도출을 위한 타켓 심층신경망을 포함하고,

상기 타켓 심층신경망에 대한 가중치 파라미터는,

상기 메인 심층신경망에 대한 가중치 파라미터를 기반으로 업데이트되는, 에너지 관리 장치.

청구항 8

에너지 관리 장치에 의한 에너지 관리 방법에 있어서,

제1 타임 슬롯에서의 에너지 관리 시스템(energy management system, EMS)에 대한 상태 정보를 획득하고, 상기 제1 타임 슬롯에서 에너지 공급원으로부터 공급받을 에너지 공급량을 선택하는 단계;

상기 선택된 에너지 공급량과 상기 제1 타임 슬롯에서의 상기 EMS에 대한 상태 정보를 기반으로 에너지 스케쥴링 정책 모델을 이용하여 상기 제1 타임 슬롯에서의 에너지 요금에 대한 정보 및 제2 타임 슬롯에서의 상기 EMS에 대한 상태 정보를 획득하는 단계; 및

상기 제1 타임 슬롯에서의 상기 EMS에 대한 상태 정보, 상기 제1 타임 슬롯에서 상기 에너지 공급원으로부터 공급받을 에너지 공급량, 상기 제1 타임 슬롯에서의 에너지 요금에 대한 정보 및 상기 제2 타임 슬롯에서의 상기 EMS에 대한 상태 정보를 기반으로 상기 에너지 스케쥴링 정책 모델을 업데이트하는 단계를 포함하되,

상기 상태 정보는,

해당 타임 슬롯에서 상기 EMS의 에너지 수요량, 해당 타임 슬롯에서 상기 EMS 내 재생 에너지원의 에너지 생산량, 해당 타임 슬롯에서 상기 EMS 내 에너지 저장소의 에너지 잔여량, 해당 타임 슬롯에서 상기 에너지 공급원의 TOU(Time of Use) 요금 정보, 및 상기 EMS에 대한 DC(Demand Charge) 값을 포함하는, 에너지 관리 방법.

청구항 9

제8항에 있어서,

상기 DC 값은,

현재 주기에서 현재 시간까지 상기 에너지 공급원으로부터 공급받은 최대 에너지 공급량을 기반으로 도출되는, 에너지 관리 방법.

청구항 10

제8항에 있어서,

상기 선택하는 단계는,

입실론 그리디(epsilon greedy) 정책을 기반으로 상기 제1 타임 슬롯에서 상기 에너지 공급원으로부터 공급받을 에너지 공급량을 선택하는 단계를 포함하는, 에너지 관리 방법.

청구항 11

제8항에 있어서,

상기 업데이트하는 단계 이후에,

상기 에너지 스케쥴링 정책 모델에 대한 로컬 가중치 파라미터의 변화량 정보를 에너지 스케쥴링 장치로 전송하는 단계; 및

상기 에너지 스케쥴링 장치로부터 상기 에너지 스케쥴링 장치의 중앙 에너지 스케쥴링 정책 모델에 대한 중앙 가중치 파라미터를 수신하는 단계를 더 포함하고,

상기 에너지 스케쥴링 정책 모델에 대한 로컬 가중치 파라미터는,

상기 중앙 가중치 파라미터를 기반으로 업데이트되는, 에너지 관리 방법.

청구항 12

제8항에 있어서,

상기 에너지 스케쥴링 정책 모델은,

에너지 스케쥴링 정책의 심층강화학습을 위한 메인 심층신경망 및 목표 에너지 요금의 도출을 위한 타켓 심층신경망을 포함하고,

상기 타켓 심층신경망에 대한 가중치 파라미터는,

상기 메인 심층신경망에 대한 가중치 파라미터를 기반으로 업데이트되는, 에너지 관리 방법.

청구항 13

에너지 스케쥴링 장치에 있어서,

복수개의 에너지 관리 시스템(energy management system, EMS) 각각으로부터 로컬 에너지 스케쥴링 정책 모델에 대한 로컬 가중치 파라미터의 변화량 정보를 수신하도록 구성되는 수신부;

상기 로컬 가중치 파라미터의 변화량 정보를 기반으로 중앙 에너지 스케쥴링 정책 모델의 중앙 가중치 파라미터를 업데이트하도록 구성되는 업데이트부; 및

상기 업데이트된 중앙 가중치 파라미터를 상기 복수개의 EMS로 송신하도록 구성되는 송신부를 포함하되,

상기 로컬 에너지 스케쥴링 정책 모델은,

단위 시간에서 각 EMS의 에너지 수요량, 상기 단위 시간에서 상기 각 EMS 내 재생 에너지원의 에너지 생산량, 상기 단위 시간에서 상기 EMS 내 에너지 저장소의 에너지 잔여량, 상기 단위 시간에서 상기 각 EMS에 대한 에너지 공급원의 TOU(Time of Use) 요금 정보, 및 상기 각 EMS에 대한 DC(Demand Charge) 값을 기반으로 학습되는, 에너지 스케쥴링 장치.

청구항 14

제13항에 있어서,

상기 DC 값은,

현재 주기에서 현재 시간까지 상기 각 EMS가 에너지 공급원으로부터 공급받은 최대 에너지 공급량을 기반으로 도출되는, 에너지 스케쥴링 장치.

청구항 15

제13항에 있어서,

상기 로컬 가중치 파라미터는,

상기 업데이트된 중앙 가중치 파라미터로 업데이트되는, 에너지 스케쥴링 장치.

청구항 16

에너지 스케쥴링 장치에 의한 에너지 스케쥴링 방법에 있어서,

복수개의 에너지 관리 시스템(energy management system, EMS) 각각으로부터 로컬 에너지 스케쥴링 정책 모델에 대한 로컬 가중치 파라미터의 변화량 정보를 수신하는 단계;

상기 로컬 가중치 파라미터의 변화량 정보를 기반으로 중앙 에너지 스케쥴링 정책 모델의 중앙 가중치 파라미터를 업데이트하는 단계; 및

상기 업데이트된 중앙 가중치 파라미터를 상기 복수개의 EMS로 송신하는 단계를 포함하되,

상기 로컬 에너지 스케쥴링 정책 모델은,

단위 시간에서 각 EMS의 에너지 수요량, 상기 단위 시간에서 상기 각 EMS 내 재생 에너지원의 에너지 생산량, 상기 단위 시간에서 상기 EMS 내 에너지 저장소의 에너지 잔여량, 상기 단위 시간에서 상기 각 EMS에 대한 에너지 공급원의 TOU(Time of Use) 요금 정보, 및 상기 각 EMS에 대한 DC(Demand Charge) 값을 기반으로 학습되는, 에너지 스케쥴링 장치.

청구항 17

제16항에 있어서,

상기 DC 값은,

현재 주기에서 현재 시간까지 상기 각 EMS가 에너지 공급원으로부터 공급받은 최대 에너지 공급량을 기반으로 도출되는, 에너지 스케쥴링 방법.

청구항 18

제16항에 있어서,

상기 로컬 가중치 파라미터는,

상기 업데이트된 중앙 가중치 파라미터로 업데이트되는, 에너지 스케쥴링 방법.

청구항 19

컴퓨터 판독 가능한 기록매체에 저장되어 있는 컴퓨터 프로그램으로서,

상기 컴퓨터 프로그램은, 프로세서에 의해 실행되면,

제1 타임 슬롯에서의 에너지 관리 시스템(energy management system, EMS)에 대한 상태 정보를 획득하는 단계와,

상기 제1 타임 슬롯에서 에너지 공급원으로부터 공급받을 에너지 공급량을 선택하는 단계와,

상기 선택된 에너지 공급량과 상기 제1 타임 슬롯에서의 상기 EMS에 대한 상태 정보를 기반으로 에너지 스케쥴링 정책 모델을 이용하여 상기 제1 타임 슬롯에서의 에너지 요금에 대한 정보 및 제2 타임 슬롯에서의 상기 EMS에 대한 상태 정보를 획득하는 단계와,

상기 제1 타임 슬롯에서의 상기 EMS에 대한 상태 정보, 상기 제1 타임 슬롯에서 상기 에너지 공급원으로부터 공급받을 에너지 공급량, 상기 제1 타임 슬롯에서의 에너지 요금에 대한 정보 및 상기 제2 타임 슬롯에서의 상기 EMS에 대한 상태 정보를 기반으로 상기 에너지 스케쥴링 정책 모델을 업데이트하는 단계를 포함하는 에너지 관리 방법을 상기 프로세서가 수행하도록 하기 위한 명령어를 포함하되,

상기 상태 정보는,

해당 타임 슬롯에서 상기 EMS의 에너지 수요량, 해당 타임 슬롯에서 상기 EMS 내 재생 에너지원의 에너지 생산량, 해당 타임 슬롯에서 상기 EMS 내 에너지 저장소의 에너지 잔여량, 해당 타임 슬롯에서 상기 에너지 공급원의 TOU(Time of Use) 요금 정보, 및 상기 EMS에 대한 DC(Demand Charge) 값을 포함하는, 컴퓨터 프로그램.

청구항 20

컴퓨터 판독 가능한 기록매체에 저장되어 있는 컴퓨터 프로그램으로서,

상기 컴퓨터 프로그램은, 프로세서에 의해 실행되면,

복수개의 에너지 관리 시스템(energy management system, EMS) 각각으로부터 로컬 에너지 스케쥴링 정책 모델에 대한 로컬 가중치 파라미터의 변화량 정보를 수신하는 단계와,

상기 로컬 가중치 파라미터의 변화량 정보를 기반으로 중앙 에너지 스케쥴링 정책 모델의 중앙 가중치 파라미터를 업데이트하는 단계와,

상기 업데이트된 중앙 가중치 파라미터를 상기 복수개의 EMS로 송신하는 단계를 포함하는 에너지 스케쥴링 방법을 상기 프로세서가 수행하도록 하기 위한 명령어를 포함하되,

상기 로컬 에너지 스케쥴링 정책 모델은,

단위 시간에서 각 EMS의 에너지 수요량, 상기 단위 시간에서 상기 각 EMS 내 재생 에너지원의 에너지 생산량, 상기 단위 시간에서 상기 EMS 내 에너지 저장소의 에너지 잔여량, 상기 단위 시간에서 상기 각 EMS에 대한 에너지 공급원의 TOU(Time of Use) 요금 정보, 및 상기 각 EMS에 대한 DC(Demand Charge) 값을 기반으로 학습되는, 컴퓨터 프로그램.

발명의 설명

기 술 분 야

[0001] 본 발명은 에너지 스케쥴링 방법 및 장치에 관한 것으로, 연합강화학습(federated reinforcement learning)을 이용하여 다수 가구의 에너지 수요를 충족시키면서 총 에너지 비용이 최소화되도록 에너지를 관리하고 스케쥴링하는 방법 및 장치에 관한 것이다.

배 경 기 술

[0002] 최근, 에너지 사용량의 급격한 증가, 연료 기반 에너지 생산량의 감소, 재생에너지(renewable energy)에 대한 의존도 증가 등으로 인해 안정적으로 에너지를 공급하기 어려워지고 있다. 이에 따라 정전을 예방하기 위해서는 피크(peak) 시간 대의 에너지 사용량을 줄이는 것이 매우 중요하다. 이를 위해 에너지 사업자는 TOU(Time of Use) 요금제, DC(Demand Charge) 요금제 등 다양한 요금제를 채택하고 있다. TOU 요금제는 피크 시간과 오프 피크(off peak) 시간에 서로 다른 전기 단가를 적용하여 에너지 수요자가 피크 시간 대의 에너지 소비를 줄이도록 유도한다. DC 요금제는 특정 주기(예를 들어, 하루, 한달 등) 동안의 최대 순간 에너지 수요를 기준으로 비용을 부과하여 에너지 수요자가 특정 시간 동안 집중적으로 에너지를 사용하지 않도록 유도한다. 여기서 순간 에너지 수요는 일정 기간(예를 들어, 30분, 1시간 등)에 걸쳐서 계산된다.

[0003] 이러한 요금제들은 에너지 제공자 및 에너지 관리자의 관점에서 피크 시간 동안의 에너지 수요를 피크 시간이 아닌 시간에 걸쳐 분산시켜 피크 에너지 수요를 줄일 수 있다는 장점이 있다. 그러나 에너지 수요자의 관점에서, 이러한 요금제들은 비용으로 인해 에너지 수요를 조정해야 하는 불편함이 있다. 이러한 불편함을 감소시키기 위해 에너지 수요자가 에너지를 저장한 후 나중에 사용할 수 있는 에너지 저장 장치와 재생에너지원(renewable energy source)을 갖춘 에너지 관리 시스템(energy management system, EMS)이 사용될 수 있다. EMS는 각 가구의 에너지 관리에 유연성을 제공하여 결과적으로 해당 가구의 에너지 수요를 충족하면서 에너지 비용을 최소화할 수 있도록 한다. 따라서 EMS를 위한 에너지 스케쥴링 방법이 광범위하게 연구되고 있다.

[0004] 그러나, 컨벡스 최적화(convex optimization) 및 선형 프로그래밍(linear programming)과 같은 최적화 기법들은 매우 높은 의사결정 복잡도를 가지는 에너지 관리 문제를 해결하기 어렵다. 입자 군집 최적화(particle swarm optimization) 및 유전자 알고리즘(genetic algorithm)과 같은 메타휴리스틱(metaheuristic) 기법들은 복잡도가 높은 문제를 해결할 수 있지만, 최적화 성능을 희생해야 한다. 모델 예측 제어(model predictive control) 기법은 에너지 관리 문제를 해결할 수 있지만, 재생 에너지 및 에너지 수요 분포의 확률 분포나 예측 정보와 같은 시스템 불확실성에 대한 사전 정보(또는 사전 예측 정보)가 반드시 필요하다. 또한, 알고리즘의 성능이 사전 정보의 정확성에 크게 의존하기 때문에 정확한 사전 정보를 얻기 위해 많은 시간과 비용이 요구된다.

[0005] 이러한 문제를 해결하기 위해 딥러닝(deep learning)을 이용하는 데이터 기반(data driven) 에너지 스케쥴링 방법이 연구되고 있다. 그러나, ADP(Approximate Dynamic Programming)와 같은 데이터 기반 기법들은 시스템 불

확실성 예측, 사전 정보 획득 등에 대한 리스크와 비용 문제를 해결할 수 있지만, 딥러닝 모델을 학습시키기 위한 충분한 데이터 수집에 많은 노력과 비용이 요구되며, 개별 가구의 전력 사용 데이터에 대한 보안 문제로 직접적인 데이터 수집에 한계가 있다.

[0006] 따라서, 높은 의사결정 복잡도를 가진 에너지 관리 문제를 효과적으로 해결할 수 있고, 에너지 소비량, 태양광 발전량 등에 대한 사전 정보가 필요하지 않으며, 각 EMS가 서로 다른 환경을 가지더라도 모든 EMS에서 사용 가능하고, TOU 요금제와 DC 요금제를 고려한 에너지 스케쥴링 방법이 요구된다.

선행기술문헌

특허문헌

[0007] (특허문헌 0001) 대한민국 등록특허공보 제10-2234270호 (등록일: 2021년 03월 25일)

발명의 내용

해결하려는 과제

[0008] 본 발명의 기술적 과제는 미래의 에너지 소비량, 재생 에너지 생산량 등에 대한 예측 정보 없이도 효과적으로 다수 가구에 대한 에너지 스케쥴링을 수행할 수 있는 방법 및 장치를 제공함에 있다.

[0009] 본 발명의 다른 기술적 과제는 TOU(Time Of Use) 요금과 DC(Demand Charge) 요금을 모두 고려한 에너지 스케쥴링 방법 및 장치를 제공함에 있다.

[0010] 본 발명의 또 다른 기술적 과제는 모든 EMS(Energy Management System)에서 사용 가능한 범용 에너지 스케쥴링 방법 및 장치를 제공함에 있다.

[0011] 본 발명이 해결하고자 하는 과제는 이상에서 언급된 것으로 제한되지 않으며, 언급되지 않은 또 다른 해결하고자 하는 과제는 아래의 기재로부터 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0012] 본 발명의 일 실시예에 따르면, 에너지 관리 장치는 에너지 관리 시스템(energy management system, EMS)의 에너지 수요량을 관리하도록 구성되는 수요량 관리부, 상기 EMS 내 재생 에너지원에서의 에너지 생산량을 관리하도록 구성되는 생산량 관리부, 상기 EMS 내 에너지 저장소의 에너지 잔여량을 관리하도록 구성되는 잔여량 관리부, 및 상기 EMS에 대한 상태 정보를 기반으로 에너지 스케쥴링 정책 모델에 대한 심층강화학습을 수행하여 에너지 공급원으로부터 공급받는 에너지 공급량을 관리하도록 구성되는 공급량 관리부를 포함하되, 상기 상태 정보는 단위 시간에서 상기 EMS의 에너지 수요량, 상기 단위 시간에서 상기 EMS의 에너지 생산량, 상기 단위 시간에서 상기 EMS의 에너지 잔여량, 상기 단위 시간에서 상기 에너지 공급원의 TOU(Time of Use) 요금 정보, 및 상기 EMS에 대한 DC(Demand Charge) 값을 포함할 수 있다.

[0013] 여기서, 상기 DC 값은 현재 주기에서 현재 시간까지 상기 에너지 공급원으로부터 공급받은 최대 에너지 공급량을 기반으로 도출될 수 있다.

[0014] 또한, 상기 공급량 관리부는 제1 타임 슬롯에서 상기 에너지 공급원으로부터 공급받을 에너지 공급량을 선택하고, 상기 선택된 에너지 공급량과 상기 제1 타임 슬롯에서의 상기 EMS에 대한 상태 정보를 기반으로 상기 에너지 스케쥴링 정책 모델을 이용하여 상기 제1 타임 슬롯에서의 에너지 요금에 대한 정보 및 제2 타임 슬롯에서의 상기 EMS에 대한 상태 정보를 획득할 수 있다.

[0015] 또한, 상기 공급량 관리부는 입실론 그리디(epsilon greedy) 정책을 기반으로 상기 제1 타임 슬롯에서 상기 에너지 공급원으로부터 공급받을 에너지 공급량을 선택할 수 있다.

[0016] 또한, 상기 공급량 관리부는 상기 제1 타임 슬롯에서의 상기 EMS에 대한 상태 정보, 상기 제1 타임 슬롯에서 상기 에너지 공급원으로부터 공급받을 에너지 공급량, 상기 제1 타임 슬롯에서의 에너지 요금에 대한 정보 및 상기 제2 타임 슬롯에서의 상기 EMS에 대한 상태 정보를 기반으로 상기 에너지 스케쥴링 정책 모델을 업데이트할 수 있다.

[0017] 또한, 상기 에너지 관리 장치는 상기 에너지 스케쥴링 정책 모델에 대한 로컬 가중치 파라미터의 변화량 정보를 에너지 스케쥴링 장치로 전송하고, 상기 에너지 스케쥴링 장치로부터 상기 에너지 스케쥴링 장치의 중앙 에너지 스케쥴링 정책 모델에 대한 중앙 가중치 파라미터를 수신하도록 구성되는 통신부를 더 포함하고, 상기 공급량 관리부는 상기 중앙 가중치 파라미터를 이용하여 상기 로컬 가중치 파라미터를 업데이트할 수 있다.

[0018] 또한, 상기 에너지 스케쥴링 정책 모델은 에너지 스케쥴링 정책의 심층강화학습을 위한 메인 심층신경망 및 목표 에너지 요금의 도출을 위한 타켓 심층신경망을 포함하고, 상기 타켓 심층신경망에 대한 가중치 파라미터는 상기 메인 심층신경망에 대한 가중치 파라미터를 기반으로 업데이트될 수 있다.

[0019] 본 발명의 다른 실시예에 따르면, 에너지 관리 장치에 의한 에너지 관리 방법은 제1 타임 슬롯에서의 EMS에 대한 상태 정보를 획득하고, 상기 제1 타임 슬롯에서 에너지 공급원으로부터 공급받을 에너지 공급량을 선택하는 단계, 상기 선택된 에너지 공급량과 상기 제1 타임 슬롯에서의 상기 EMS에 대한 상태 정보를 기반으로 에너지 스케쥴링 정책 모델을 이용하여 상기 제1 타임 슬롯에서의 에너지 요금에 대한 정보 및 제2 타임 슬롯에서의 상기 EMS에 대한 상태 정보를 획득하는 단계, 및 상기 제1 타임 슬롯에서의 상기 EMS에 대한 상태 정보, 상기 제1 타임 슬롯에서 상기 에너지 공급원으로부터 공급받을 에너지 공급량, 상기 제1 타임 슬롯에서의 에너지 요금에 대한 정보 및 상기 제2 타임 슬롯에서의 상기 EMS에 대한 상태 정보를 기반으로 상기 에너지 스케쥴링 정책 모델을 업데이트하는 단계를 포함하되, 상기 상태 정보는 해당 타임 슬롯에서 상기 EMS의 에너지 수요량, 해당 타임 슬롯에서 상기 EMS 내 재생 에너지원의 에너지 생산량, 해당 타임 슬롯에서 상기 EMS 내 에너지 저장소의 에너지 잔여량, 해당 타임 슬롯에서 상기 에너지 공급원의 TOU 요금 정보, 및 상기 EMS에 대한 DC 값을 포함할 수 있다.

[0020] 본 발명의 또 다른 실시예에 따르면, 에너지 스케쥴링 장치는 복수개의 EMS 각각으로부터 로컬 에너지 스케쥴링 정책 모델에 대한 로컬 가중치 파라미터의 변화량 정보를 수신하도록 구성되는 수신부, 상기 로컬 가중치 파라미터의 변화량 정보를 기반으로 중앙 에너지 스케쥴링 정책 모델의 중앙 가중치 파라미터를 업데이트하도록 구성되는 업데이트부, 및 상기 업데이트된 중앙 가중치 파라미터를 상기 복수개의 EMS로 송신하도록 구성되는 송신부를 포함하되, 상기 로컬 에너지 스케쥴링 정책 모델은 단위 시간에서 각 EMS의 에너지 수요량, 상기 단위 시간에서 상기 각 EMS 내 재생 에너지원의 에너지 생산량, 상기 단위 시간에서 상기 EMS 내 에너지 저장소의 에너지 잔여량, 상기 단위 시간에서 상기 각 EMS에 대한 에너지 공급원의 TOU 요금 정보, 및 상기 각 EMS에 대한 DC 값을 기반으로 학습될 수 있다.

[0021] 본 발명의 또 다른 실시예에 따르면, 에너지 스케쥴링 장치에 의한 에너지 스케쥴링 방법은 복수개의 EMS 각각으로부터 로컬 에너지 스케쥴링 정책 모델에 대한 로컬 가중치 파라미터의 변화량 정보를 수신하는 단계, 상기 로컬 가중치 파라미터의 변화량 정보를 기반으로 중앙 에너지 스케쥴링 정책 모델의 중앙 가중치 파라미터를 업데이트하는 단계, 및 상기 업데이트된 중앙 가중치 파라미터를 상기 복수개의 EMS로 송신하는 단계를 포함하되, 상기 로컬 에너지 스케쥴링 정책 모델은 단위 시간에서 각 EMS의 에너지 수요량, 상기 단위 시간에서 상기 각 EMS 내 재생 에너지원의 에너지 생산량, 상기 단위 시간에서 상기 EMS 내 에너지 저장소의 에너지 잔여량, 상기 단위 시간에서 상기 각 EMS에 대한 에너지 공급원의 TOU 요금 정보, 및 상기 각 EMS에 대한 DC 값을 기반으로 학습될 수 있다.

[0022] 본 발명의 또 다른 실시예에 따르면, 컴퓨터 판독 가능한 기록매체에 저장되어 있는 컴퓨터 프로그램은, 프로세서에 의해 실행되면, 제1 타임 슬롯에서의 EMS에 대한 상태 정보를 획득하는 단계와, 상기 제1 타임 슬롯에서 에너지 공급원으로부터 공급받을 에너지 공급량을 선택하는 단계와, 상기 선택된 에너지 공급량과 상기 제1 타임 슬롯에서의 상기 EMS에 대한 상태 정보를 기반으로 에너지 스케쥴링 정책 모델을 이용하여 상기 제1 타임 슬롯에서의 에너지 요금에 대한 정보 및 제2 타임 슬롯에서의 상기 EMS에 대한 상태 정보를 획득하는 단계와, 상기 제1 타임 슬롯에서의 상기 EMS에 대한 상태 정보, 상기 제1 타임 슬롯에서 상기 에너지 공급원으로부터 공급받을 에너지 공급량, 상기 제1 타임 슬롯에서의 에너지 요금에 대한 정보 및 상기 제2 타임 슬롯에서의 상기 EMS에 대한 상태 정보를 기반으로 상기 에너지 스케쥴링 정책 모델을 업데이트하는 단계를 포함하는 에너지 관리 방법을 상기 프로세서가 수행하도록 하기 위한 명령어를 포함하되, 상기 상태 정보는 해당 타임 슬롯에서 상기 EMS의 에너지 수요량, 해당 타임 슬롯에서 상기 EMS 내 재생 에너지원의 에너지 생산량, 해당 타임 슬롯에서 상기 EMS 내 에너지 저장소의 에너지 잔여량, 해당 타임 슬롯에서 상기 에너지 공급원의 TOU 요금 정보, 및 상기 EMS에 대한 DC 값을 포함할 수 있다.

[0023] 본 발명의 또 다른 실시예에 따르면, 컴퓨터 판독 가능한 기록매체에 저장되어 있는 컴퓨터 프로그램은, 프로세서에 의해 실행되면, 복수개의 EMS 각각으로부터 로컬 에너지 스케쥴링 정책 모델에 대한 로컬 가중치 파라미터

의 변화량 정보를 수신하는 단계와, 상기 로컬 가중치 파라미터의 변화량 정보를 기반으로 중앙 에너지 스케쥴링 정책 모델의 중앙 가중치 파라미터를 업데이트하는 단계와, 상기 업데이트된 중앙 가중치 파라미터를 상기 복수개의 EMS로 송신하는 단계를 포함하는 에너지 스케쥴링 방법을 상기 프로세서가 수행하도록 하기 위한 명령어를 포함하되, 상기 로컬 에너지 스케쥴링 정책 모델은 단위 시간에서 각 EMS의 에너지 수요량, 상기 단위 시간에서 상기 각 EMS 내 재생 에너지원의 에너지 생산량, 상기 단위 시간에서 상기 EMS 내 에너지 저장소의 에너지 잔여량, 상기 단위 시간에서 상기 각 EMS에 대한 에너지 공급원의 TOU 요금 정보, 및 상기 각 EMS에 대한 DC 값을 기반으로 학습될 수 있다.

발명의 효과

[0024] 본 발명의 일 실시예에 따르면, 각 가구의 환경 정보를 모두 전송하지 않고 에너지 스케쥴링 정책 모델의 파라미터만을 전송하기 때문에 네트워크 비용 및 데이터 보안 문제를 해결할 수 있고, 미래의 에너지 소비량, 재생에너지 생산량 등에 대한 예측 정보 없이도 높은 의사결정 복잡도를 가지는 다수 가구에 대한 에너지 관리 문제를 효과적으로 해결할 수 있다.

[0025] 본 발명의 일 실시예에 따르면, TOU(Time Of Use) 요금 및 DC(Demand Charge) 요금을 모두 고려하여 에너지를 스케쥴링할 수 있다.

[0026] 본 발명의 일 실시예에 따르면, 각 EMS(Energy Management System)의 에너지 스케쥴링 정책 모델의 파라미터를 취합하여 모든 EMS에서 사용 가능한 범용 에너지 스케쥴링 정책 모델의 파라미터를 생성하고 이를 각 EMS에 재분배하기 때문에 모델 학습 측면에서 데이터 효율성이 향상될 수 있다.

도면의 간단한 설명

[0027] 도 1은 본 발명의 일 실시예에 따른 에너지 스케쥴링 시스템을 나타내는 도면이다.

도 2는 본 발명의 일 실시예에 따른 에너지 관리 시스템을 나타내는 도면이다.

도 3은 본 발명의 일 실시예에 적용되는 심층강화학습을 나타내는 도면이다.

도 4는 본 발명의 일 실시예에 따른 에너지 스케쥴링 장치를 나타내는 도면이다.

도 5는 본 발명의 일 실시예에 따른 에너지 관리 방법을 나타내는 도면이다.

도 6는 본 발명의 일 실시예에 따른 에너지 스케쥴링 방법을 나타내는 도면이다.

발명을 실시하기 위한 구체적인 내용

[0028] 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은, 첨부되는 도면과 함께 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하고, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 따라서, 몇몇 실시예에서, 잘 알려진 처리 단계들, 잘 알려진 소자 구조 및 잘 알려진 기술들은 본 발명이 모호하게 해석되는 것을 피하기 위하여 구체적으로 설명되지 않는다.

[0029] 본 명세서에서 사용되는 용어는 본 발명에서의 기능을 고려하면서 가능한 현재 널리 사용되는 일반적인 용어들 중에서 선택되었으나, 이는 당 분야에 종사하는 기술자의 의도 또는 판례, 새로운 기술의 출현 등에 따라 달라질 수 있다. 또한, 특정한 경우는 출원인이 임의로 선정한 용어도 있으며, 이 경우 해당되는 발명의 설명 부분에서 상세히 그 의미를 기재할 것이다. 따라서 본 명세서에서 사용되는 용어는 단순한 용어의 명칭이 아닌, 그 용어가 가지는 의미와 본 발명의 전반에 걸친 내용을 토대로 해석되어야 한다.

[0030] 본 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 있음을 의미한다.

[0031] 또한, 본 명세서에서 사용되는 '부', '모듈' 등의 용어들은 소프트웨어 또는 FPGA(Field Programmable Gate Array)나 ASIC(Application Specific Integrated Circuit)과 같은 하드웨어 구성요소를 의미할 수 있으며, '부', '모듈' 등은 적어도 하나의 기능이나 동작을 수행한다. 그렇지만 '부', '모듈' 등은 소프트웨어 또는 하드웨어에 한정되는 의미는 아니다. '부', '모듈' 등은 어드레싱할 수 있는 저장 매체에 있도록 구성될 수도 있고, 하나 또는 그 이상의 프로세서들을 재생시키도록 구성될 수도 있다. 따라서, 일 예로서 '부', '모

듈' 등은 소프트웨어 구성요소들, 객체지향 소프트웨어 구성요소들, 클래스 구성요소들 및 태스크 구성요소들과 같은 구성요소들과, 프로세스들, 함수들, 속성들, 프로시저들, 서브루틴들, 프로그램 코드의 세그먼트들, 드라이버들, 펌웨어, 마이크로 코드, 회로, 데이터, 데이터베이스, 데이터 구조들, 테이블들, 어레이들 및 변수들을 포함한다. 본 발명의 구성요소들과 '부', '모듈' 등에서 제공되는 기능은 더 작은 수의 구성요소들, '부', '모듈' 등으로 결합되거나 추가적인 구성요소들, '부', '모듈' 등으로 더 분리될 수 있다.

[0032] 또한, 본 명세서에서 제1, 제2, 제3 등의 용어는 다양한 구성 요소들을 설명하는데 사용될 수 있지만, 이러한 구성 요소들은 상기 용어들에 의해 한정되는 것은 아니다. 상기 용어들은 하나의 구성 요소를 다른 구성 요소들로부터 구별하는 목적으로 사용된다.

[0033] 이하에서는 첨부한 도면을 참고하여 본 발명의 실시예에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략한다. 명세서 전체에 걸쳐 동일한 참조 부호는 동일한 구성 요소를 지칭한다.

[0034] 도 1은 본 발명의 일 실시예에 따른 에너지 스케쥴링 시스템을 나타내는 도면이다.

[0035] 도 1을 참조하면, 본 발명의 일 실시예에 따른 에너지 스케쥴링 시스템은 복수개의 에너지 관리 시스템(energy management system, EMS) 및 에너지 스케쥴링 장치를 포함한다.

[0036] 각 EMS(110, 120)는 태양광, 태양열, 풍력, 지열 등을 이용하여 에너지(또는 전력)를 생산하는 재생 에너지원(reusable energy source), 에너지 공급원(130) 으로부터 전력망을 통해 공급되는 에너지와 상기 재생 에너지원에서 생산되는 재생 에너지를 저장하는 에너지 저장소(energy storage), 및 해당 EMS(110, 120)가 구비된 가구의 에너지를 관리하는 에너지 관리 장치를 포함할 수 있다.

[0037] 각 EMS(110, 120) 내 에너지 관리 장치는 해당 EMS(110, 120)가 구비된 가구에서 요구 또는 소비되는 에너지의 양(이하, 에너지 수요량이라 함), 해당 EMS에 구비된 재생 에너지원에서 생산 또는 수확(harvest)되는 에너지의 양(이하, 에너지 생산량이라 함), 해당 EMS에 구비된 에너지 저장소에 저장된 잔여 에너지의 양(이하, 에너지 잔여량), 적어도 하나의 에너지 공급원(130)으로부터 공급받는 에너지의 양(이하, 에너지 공급량이라 함) 등을 관리할 수 있다.

[0038] 에너지 공급원(130)은 에너지를 공급받는 시간(예를 들어, off-peak time, low-peak time, mid-peak time, high-peak time 등)에 따라 요금이 변화하는 TOU(Time Of Use) 요금제 및 특정 주기(예를 들어, 하루, 한주, 한달 등) 중 특정 기간(예를 들어, 한시간, 두시간 등) 동안의 최대 에너지 공급량을 기준으로 요금을 부과하는 DC(demand Charge) 요금제를 기반으로 해당 EMS(또는 가구)로 공급한 에너지의 양에 대한 비용을 책정한다.

[0039] 각 EMS(110, 120) 내 에너지 관리 장치는 상기 에너지 수요량, 상기 에너지 생산량, 상기 에너지 잔여량 등과 같은 환경 정보를 기반으로 해당 EMS(또는 가구)의 에너지 수요는 만족하고 요금을 최소화할 수 있도록 에너지 공급원(130)으로부터의 에너지 공급량을 결정할 수 있다. 이를 위하여 일 예로, 각 EMS(110, 120) 내 에너지 관리 장치는 심층강화학습(deep reinforcement learning, DRL)을 기반으로 최적의 에너지 공급량을 결정하는 에너지 스케쥴링 정책 모델을 생성 또는 구축할 수 있다. 상기 에너지 스케쥴링 정책 모델에 대한 정보(예를 들어, 해당 모델의 파라미터)는 해당 에너지 스케쥴링 정책 모델의 학습 과정에서 주기적으로 에너지 스케쥴링 장치(140)로 전송될 수 있다.

[0040] 에너지 스케쥴링 장치(140)는 각 EMS(110, 120)로부터 수신한 해당 에너지 스케쥴링 정책 모델에 대한 정보를 모두 취합하고 연합강화학습(Federated Reinforcement Learning)을 수행하여 통합된 에너지 스케쥴링 정책 모델을 생성하고, 해당 모델에 대한 정보를 각 EMS(110, 120)로 재분배할 수 있다. 에너지 스케쥴링 장치(140)는 코디네이터(coordinator), 중앙 에너지 관리 장치 등의 용어로 불릴 수 있다.

[0041] 이하, 도 2 및 도 3을 참조하여 본 발명의 일 실시예에 따른 EMS에 대해 보다 상세히 설명한다.

[0042] 도 2는 본 발명의 일 실시예에 따른 에너지 관리 시스템을 나타내는 도면이고, 도 3은 본 발명의 일 실시예에 적용되는 심층강화학습을 나타내는 도면이다.

[0043] 도 2를 참조하면, 본 발명의 일 실시예에 따른 EMS는 에너지 관리 장치(200), 재생 에너지원(225) 및 에너지 저장소(235)를 포함할 수 있으며, 에너지 관리 장치(220)는 수요량 관리부(210), 생산량 관리부(220), 잔여량 관리부(230), 공급량 관리부(240), 메모리(250) 및 통신부(260)를 포함할 수 있다.

[0044] 수요량 관리부(210)는 해당 가구에서 소비되는 에너지의 양을 모니터링하고 이를 기반으로 EMS(200)의 에너지

수요량을 관리하도록 구성될 수 있다.

[0045] 생산량 관리부(210)는 재생 에너지원(225)에서 생산 또는 수확되는 에너지의 양을 모니터링하여 EMS에서의 에너지 생산량을 관리하도록 구성될 수 있다.

[0046] 잔여량 관리부(230)는 배터리와 같은 에너지 저장소(235)에 저장된 에너지의 양을 모니터링하여 EMS의 에너지 잔여량을 관리하도록 구성될 수 있다.

[0047] 공급량 관리부(230)는 각 단위 시간에서의 해당 가구에 대한 환경 정보를 기반으로 상기 각 단위 시간에서의 EMS에 대한 상태 정보를 구성하고, 상기 EMS에 대한 상태 정보를 기반으로 에너지 스케쥴링 정책 모델에 대한 심층강화학습을 수행하여 에너지 공급원으로부터의 에너지 공급량을 관리하도록 구성될 수 있다. 여기서, 상기 상태 정보는 해당 단위 시간에서 EMS의 에너지 수요량, 상기 단위 시간에서 EMS의 에너지 생산량, 상기 단위 시간에서 EMS의 에너지 잔여량, 상기 단위 시간에서 상기 에너지 공급원의 TOU 요금 정보, 및 상기 에너지 공급원의 EMS에 대한 DC 값을 포함할 수 있다. 공급량 관리부(230)는 현재 주기에서 현재 시간까지 에너지 공급원으로부터 공급받은 최대 에너지 공급량을 기반으로 상기 EMS에 대한 DC 값을 도출할 수 있다. 이하에서, DC 값은 현재 주기에서 현재 시간까지 에너지 공급원으로부터 공급받은 최대 에너지 공급량을 의미할 수 있고, DC 단위 요금은 해당 DC 값에 따라 에너지 공급자가 부과하는 단위 요금을 의미할 수 있다. DC 요금은 DC 값과 DC 단위 요금에 따라 부과되는 요금을 의미할 수 있다.

[0048] 일 예로, 공급량 관리부(230)는 입실론 그리디(epsilon greedy) 정책을 기반으로 제1 타임 슬롯에서 에너지 공급원으로부터 공급받을 에너지 공급량을 선택하고, 상기 선택된 에너지 공급량과 제1 타임 슬롯에서의 EMS(200)에 대한 상태 정보를 기반으로 에너지 스케쥴링 정책 모델을 이용하여 제1 타임 슬롯에서의 에너지 요금에 대한 정보 및 제2 타임 슬롯에서의 EMS(200)에 대한 상태 정보를 획득할 수 있다. 그리고, 공급량 관리부(230)는 제1 타임 슬롯에서의 EMS(200)에 대한 상태 정보, 제1 타임 슬롯에서 에너지 공급원으로부터 공급받을 에너지 공급량, 제1 타임 슬롯에서의 에너지 요금에 대한 정보 및 제2 타임 슬롯에서의 EMS(200)에 대한 상태 정보를 기반으로 경험 정보를 생성하고, 이를 기반으로 에너지 스케쥴링 정책 모델을 업데이트할 수 있다.

[0049] 메모리(250)는 에너지 수요량에 대한 정보, 에너지 생산량에 대한 정보, 에너지 잔여량에 대한 정보, 에너지 공급량에 대한 정보, 공급량 관리부(230)에 의해 생성된 경험 정보, 에너지 스케쥴링 정책 모델 등을 저장할 수 있다.

[0050] 통신부(260)는 에너지 스케쥴링 정책 모델(로컬 에너지 스케쥴링 정책 모델) 에 대한 가중치 파라미터(로컬 가중치 파라미터) 및/또는 상기 가중치 파라미터의 변화량 정보를 에너지 스케쥴링 장치로 전송하고, 해당 에너지 스케쥴링 장치로부터 중앙 에너지 스케쥴링 정책 모델에 대한 중앙 가중치 파라미터를 수신하도록 구성될 수 있다. 공급량 관리부(240)는 상기 중앙 가중치 파라미터를 이용하여 상기 로컬 가중치 파라미터를 업데이트할 수 있다. 여기서, 각 에너지 스케쥴링 정책 모델은 에너지 스케쥴링 정책의 심층강화학습을 위한 메인 심층신경망 및 목표 에너지 요금의 도출을 위한 타켓 심층신경망을 포함할 수 있다. 이 경우, 상기 타켓 심층신경망에 대한 가중치 파라미터는 상기 메인 심층신경망에 대한 가중치 파라미터를 기반으로 업데이트될 수 있다.

[0051] 이하, N개(여기서, N은 자연수)의 EMS들이 하나의 에너지 공급원으로부터 에너지를 공급받는 경우, 각 EMS가 에너지 스케쥴링을 수행하는 과정에 대해 설명한다.

[0052] 각 EMS의 인덱스 n은 $n \in N = \{1, 2, \ldots, N\}$으로 표현될 수 있다. 모든 EMS들은 동일한 시간 간격으로 분할된 타임 슬롯을 기반으로 운영될 수 있다. 각 EMS는 각 타임 슬롯마다 해당 EMS의 에너지 수요량, 에너지 생산량, 에너지 잔여량 등을 고려하여 에너지 공급량을 결정하고, 상기 에너지 공급원으로부터 결정된 양만큼의 에너지를 공급받을 수 있다. 여기서, 타임 슬롯의 인덱스 t는 $t \in T = \{1, 2, \ldots, T\}$으로 표현될 수 있다. 하루가 H개의 시간으로 분할되는 경우, 시스템 시간은 인덱스 $h^t = mod(t, H)$, 시스템의 d번째 날은 $d(t) = \lfloor \frac{t}{H} \rfloor$로 표현될 수 있다. d번째 날의 타임 슬롯 집합 T(d)는 (Hd, Hd+1, ..., H(d+1)-1)로 표현될 수 있다.

[0053] 한편, 타임 슬롯 t에서 n번째 EMS의 에너지 수요량은 L^t_n, 에너지 생산량은G^t_n, 에너지 잔여량은 B^t_n, 에너지 공급량은 C^t_n로 표현될 수 있다. 또한, 각 값은 L^{max}, G^{max}, B^{max}, C^{max}의 최대값 제한이 존재한다. 타임 슬롯 t에서 T_f까지의 에너지 수요량 및 에너지 생산량은 각각 $\mathbf{L}^t_n = (L^t_n, L^{t+1}_n, \ldots, L^{t+T_f}_n)$ 및

$\mathbf{G}_n^t = (G_n^t, G_n^{t+1}, ..., G_n^{t+T_f})$로 정의될 수 있다. 여기서, T_f는 예측된(forecasted) 타임 슬롯의 개수를 나타낸다. n번째 EMS는 최대 에너지 저장 용량 B^{max}_n와 최대 충/방전량 제한 B^{rate}_n을 가지며, n번째 EMS의 에너지 잔여량은 다음의 수학식 1을 만족한다.

수학식 1

[0054] $$-B_n^{rate} \le B_n^{t+1} - B_n^t \le B_n^{rate}$$

[0055] 또한, n번째 EMS의 에너지 수요를 보장하기 위해서는 시간대별 가용 에너지의 총량이 다음의 수학식 2와 같이 에너지 수요량보다 커야 한다.

수학식 2

[0056] $$L_n^t \le \min(B_n^t, B_n^{rate}) + G_n^t + C_n^t$$

[0057] 각 타임 슬롯에서 각 EMS의 에너지 잔여량은 에너지 수요량, 에너지 생산량, 및 해당 타임 슬롯에서의 에너지 공급량을 기반으로 결정된다. 에너지 생산량과 에너지 공급량의 합보다 에너지 수요량이 많으면, 에너지 저장소에 저장된 에너지가 사용된다. 그렇지 않으면 최대 에너지 저장 용량 B_n^{max}이 허용하는 한 잉여 에너지가 에너지 저장소에 저장된다. 따라서, 다음 타임 슬롯 t+1에서 n번째 EMS의 에너지 잔여량은 다음의 수학식 3과 같이 표현될 수 있다.

수학식 3

[0058] $$B_n^{t+1} = \min\left(B_n^t + \min(G_n^t + C_n^t - L_n^t, B_n^{rate}), B_n^{max}\right)$$

[0059] 한편 전력망에서 소비되는 에너지의 경우 즉, 에너지 공급원으로부터 에너지를 공급받는 경우, 각 EMS(또는 각 가구)는 TOU 요금과 DC 요금에 따라 부과되는 비용을 지불해야 한다. 타임 슬롯 t에서 TOU 요금은 $p^t_{T,n}$로 표현된다. TOU 요금의 단가는 일반적으로 에너지 공급자가 미리 발표하므로, 모든 EMS는 각 타임 슬롯에서의 TOU 요금의 단가를 알 수 있다. 따라서, 타임 슬롯 t에서 n번째 EMS에 대한 TOU 요금은 다음의 수학식 4와 같이 정의될 수 있다.

수학식 4

[0060] $$c_{T,n}^t = C_n^t p_{T,n}^t$$

[0061] DC 요금은 타임 슬롯 t가 하루의 마지막 기간(타임 슬롯)인 경우, 다음의 수학식 5와 같이 n번째 EMS에 대한 최대 에너지 공급량에 의해 결정된다.

수학식 5

[0062] $$c_{D,n}^t = \begin{cases} \max_{t \in \mathcal{T}(d(t))} C_n^t p_{D,n}, & \text{if } \mathrm{mod}(t,H) = H-1 \\ 0, & \text{otherwise} \end{cases}$$

[0063] 여기서, $P_{D,n}$은 n번째 EMS에 대한 DC 요금의 단가를 나타낸다. 따라서, 타임 슬롯 t에서 n번째 EMS의 총 에너지 비용은 다음의 수학식 6과 같이 정의될 수 있다.

수학식 6

[0064] $$c_n^t = c_{T,n}^t + c_{D,n}^t$$

[0065] 그러므로 타임 슬롯 t에서 모든 EMS에 대한 총 에너지 비용은 다음의 수학식 7과 같이 정의될 수 있다.

수학식 7

[0066] $$c^t = \sum_{n \in N} c_n^t$$

[0067] 한편, 각 EMS의 에너지 수요를 보장하면서 모든 EMS들에 대한 에너지 요금의 합을 최소화하기 위하여 다수 가구에 대한 에너지 스케쥴링 문제를 다음과 같이 공식화할 수 있다.

[0068] 먼저 타임 슬롯 t의 상태는 $s^t \in S$로 정의될 수 있다. 여기서, S는 상태 공간(state space)를 나타낸다. s^t는 로컬 상태(local state)와 공통 상태(common state)를 포함한다. 로컬 상태의 경우 각 EMS마다 서로 다른 정보(에너지 수요량 L^t_n, 에너지 생산량 G^t_n, 에너지 잔여량 B^t_n, 에너지 공급량 C^t_n)를 포함하고, 공통 상태의 경우 모든 EMS들의 서로 공유하는 기간(period) 정보인 h^t를 포함할 수 있다. 타임 슬롯 t에서 N 번째 EMS의 행동(action)은 $a^t_n \in A_n$으로 정의될 수 있다. 여기서, 상기 행동은 에너지 공급원으로 공급받는 에너지의 양인 C^t_n를 의미한다. 따라서, n번째 EMS의 행동 공간(action space) A_n은 C를 의미할 수 있다. 타임 슬롯 t에서 모든 EMS들의 에너지 스케쥴링 결정은 at = (a^t_1, 쩐, a^t_N)로 정의될 수 있고, 행동 공간은 $A = C^N$으로 정의될 수 있다. 따라서, 수학식 1 및 2를 만족시키기 위하여 상태 s_n을 가진 n번째 EMS이 사용 가능한 에너지 스케쥴링 결정 집합은 다음의 수학식 8과 같이 결정될 수 있다.

수학식 8

[0069] $$\mathcal{A}_n(s_n) = \{a \in \mathcal{A}_n : L_n^t \le \min(B_n^t, B_n^{rate}) + G_n^t + a\}$$

[0070] 모든 EMS들이 상태 s를 가질 때 가능한 에너지 스케쥴링 결정 집합은 $\mathcal{A}(s) = \prod_{n \in N} \mathcal{A}_n(s)$와 같이 결정된다. 따라서, 모든 EMS들에서의 에너지 스케쥴링 정책은 $\pi : \mathcal{S} \rightarrow \mathcal{A}$와 같이 정의되며, 에너지 스케쥴링 정책 π는 상태 s와 행동 a 간의 매핑 함수로 나타낼 수 있다. 따라서, 에너지 스케쥴링 문제를 다음의 수학식 9와 같이 MDP(Markov decision process) 형태로 설계할 수 있다.

수학식 9

[0071] $$(\mathrm{P}) \ \underset{\pi}{\text{minimize}} \ \mathbb{E}\left[\sum_{t=0}^{\infty} (\gamma)^t c(s^t, \pi(s^t))\right]$$

[0072] 여기서 $\gamma \in [0, 1]$은 할인 계수(discount factor)이고, c(s^t, π(s^t))는 수학식 7의 타임슬롯 t에서의 총 에

너지 비용이다. 수학식 7의 비용을 수학식 6의 각 EMS 비용으로 분해할 수 있고, 각 EMS의 스케쥴링 결정은 다른 EMS에 영향을 미치지 않으므로 수학식 9의 에너지 스케쥴링 문제 (P)를 다음과 같이 단일(또는 개별) EMS의 에너지 스케쥴링 문제로 분해할 수 있다. 이를 위해 타임 슬롯 t에서 n번째 EMS의 증강 상태(augmented state)를 $\bar{s}_n^t = (s_n^t, s_C^t) \in \bar{\mathcal{S}}_n$ 로 정의한다. 여기서 $\bar{\mathcal{S}}_n$ 는 n번째 EMS의 증강 상태의 상태 공간을 나타낸다. n번째 EMS의 에너지 스케쥴링 정책은 다음의 수학식 10과 같이 상태 공간 $\bar{\mathcal{S}}_n$ 와 결정 공간 C 사이의 함수로 나타낼 수 있다.

수학식 10

[0073] $$\pi_n : \bar{\mathcal{S}}_n \rightarrow \mathcal{C}$$

[0074] 증강 상태를 사용하여 수학식 9의 에너지 스케쥴링 문제 (P)를 재정의하면 다음의 수학식 11과 같다.

수학식 11

[0075] $$(\bar{\mathrm{P}}) \quad \underset{(\pi_n)_{n\in\mathcal{N}}}{\text{minimize}} \quad \sum_{n\in\mathcal{N}} \mathbb{E}\left[\sum_{t=0}^{\infty} (\gamma)^t c_n(\bar{s}_n^t, \pi_n(\bar{s}_n^t))\right]$$

[0076] 여기서, $c_n(\bar{s}^t \pi(\bar{s}^t))$ 는 수학식 6의 타임 슬롯 t에서 n번째 EMS의 총 에너지 비용을 나타낸다. 그러면, 수학식 9의 에너지 스케쥴링 문제 (P)는 다음의 수학식 12와 같이 분산 방식으로 해결될 수 있다.

수학식 12

[0077] $$(\bar{\mathrm{P}}_n) \quad \underset{\pi_n}{\text{minimize}} \quad \mathbb{E}\left[\sum_{t=0}^{\infty} (\gamma)^t c_n(\bar{s}_n^t, \pi_n(\bar{s}_n^t))\right]$$

[0078] 에너지 스케쥴링 문제 ($\bar{\mathrm{P}}_n$)의 최적값 함수는 다음의 수학식 13과 같이 얻을 수 있다.

수학식 13

[0079] $$V_n^*(\bar{s}_n) = \min_{\pi_n} \mathbb{E}\left[\sum_{t=0}^{\infty} (\gamma)^t c_n(\bar{s}_n^t, \pi(\bar{s}_n^t)) \mid \bar{s}_n^0 = \bar{s}\right]$$

[0080] 수학식 13으로부터 최적 행동값 함수는 다음의 수학식 14와 같이 얻을 수 있다.

수학식 14

[0081] $$Q_n^*(\bar{s}, a) = \mathbb{E}\left[c(\bar{s}^t, a^t) + \gamma V^*(\bar{s}^{t+1}) | \bar{s}^t = \bar{s}, a^t = a\right]$$

[0082] 따라서, 최적의 에너지 스케쥴링 정책은 다음의 수학식 15와 같이 정의될 수 있다.

수학식 15

[0083] $$\pi_n^* = \underset{a \in \mathcal{A}_n(\bar{s})}{\operatorname{argmin}} Q_n^*(\bar{s}, a), \ \forall \bar{s} \in \bar{\mathcal{S}}_n$$

[0084] 수학식 12의 n번째 EMS의 에너지 스케쥴링 문제도 MDP이다. MDP는 일반적으로 잘 알려진 값 반복 및 정책 반복 알고리즘과 같은 동적 프로그래밍을 통해 최적의 정책을 찾아 해결할 수 있다. 그러나 이러한 동적 계획법은 에너지 수요 및 재생 에너지에 대한 확률분포와 같은 시스템 불확실성에 대한 완벽한 정보를 요구하기 때문에 현실적이지 않다. 따라서 본 발명에서는 수학식 12의 MDP를 해결하기 위하여 사전 예측 정보가 필요하지 않은 DRL 방식이 이용될 수 있다.

[0085] 이하에서는 수학식 12의 단일 EMS에 대한 에너지 스케쥴링 문제를 해결하기 위한 TOU 요금 및 DC 요금을 고려한 에너지 스케쥴링(TOU and DC aware energy scheduling, TDAS) 정책에 대해 설명한다.

[0086] 타임 슬롯 t에서의 에너지 스케쥴링을 위한 공통 상태 정보는 $s^t_C = (h^t)$로 정의된다. 타임 슬롯 t에서 각 EMS의 상태 정보는 에너지 스케쥴링과 관련하여 다음 정보를 포함한다. 타임 슬롯 t, 타임 슬롯 t에서의 에너지 수요량 L^t, 타임 슬롯 t에서의 에너지 생산량 G^t, 타임 슬롯 t에서의 에너지 잔여량 B^t, 단위 요금 $P^t_{T,n}$ 및 $P_{D,n}$. 상기 상태 정보는 에너지 스케쥴링 정책이 향후 T_f 타임 슬롯 동안 TOU로 인한 비용을 종합적으로 고려하도록 구성될 수 있다. 예를 들어, 다음 T_f 타임 슬롯의 에너지 수요가 동일한 기간의 재생 에너지보다 크고 다음 Tf 타임 슬롯의 TOU 단가가 상대적으로 높다면, EMS는 현재 타임 슬롯에서 에너지 공급원으로부터 에너지를 구매하고 이를 에너지 저장소에 저장함으로써 예상 비용을 줄일 수 있다.

[0087] 또한 비용 효율적인 에너지 스케쥴링을 위하여, 수학식 5와 같이 하루의 마지막 기간이 끝날 때 하루에 한 번 부과되는 DC 요금도 고려하여 에너지 스케쥴링 정책을 수립할 수 있다. 이를 위해서는 하루 중 최대 에너지 소비량을 고려해야 한다. 그러나 DC 요금은 타임 슬롯 별 에너지 스케쥴링 결정에 따라 하루가 끝나는 시점에 결정되기 때문에 미리 알 수 없다. 이 문제를 해결하기 위해 본 발명에서는 다음의 수학식 16과 같이 하루 d(t) 중 첫 번째 기간부터 h^t - 1까지의 최대 에너지 공급량을 기반으로 타임 슬롯 t에서 에너지 공급량이 정의될 수 있다.

수학식 16

[0088] $$C_d^t = \begin{cases} 0, & \text{if } h^t = 0 \\ \max_{t-h^t \le \tau < t} C^\tau, & \text{otherwise} \end{cases}$$

[0089] DC 비용의 정의에 따라, DC 요금을 고려한 에너지 공급량 C^t_d보다 적은 에너지 공급량을 선택해도 추가 DC 비용은 발생하지 않는다. 따라서, 수학식 10의 에너지 스케쥴링 정책을 기반으로 수학식 16과 로컬 상태를 사용함으로써 에너지 스케쥴링 시 추가 DC 비용을 유발하지 않는 최대 에너지 공급량을 결정할 수 있다. 일 예로, TDAS 정책의 타임 슬롯 t에서의 로컬 상태는 $s^t = (L^t, G^t, B^t, P^t_T$ 및 $C^t_d)$로 정의될 수 있다. 여기서 $P^t_T = (P^t_T, P^{t+1}_T, \ldots, P^{t+H-1}_T)$[2]이고, 상태 공간은 $\mathcal{S} = \mathcal{L}^{T_f+1} \times \mathcal{G}^{T_f+1} \times \mathcal{B} \times \mathcal{C} \times \mathcal{P}_T^H$ 이다. 여기서 P_T는 TOU 단위 요금의 집합을 나타낸다. 타임 슬롯 t에서 TDAS 정책의 증강 상태는 다음의 수학식 17과 같이 정의될 수 있다.

수학식 17

[0090] $$\bar{s}^t = (\mathbf{L}^t, \mathbf{G}^t, B^t, \mathbf{p}_T^t, C_d^t, h^t) \in \bar{\mathcal{S}}$$

[0091] 여기서 $\bar{S}$ 는 증강 상태 공간을 나타낸다. TDAS 정책은 수학식 17을 기반으로 TOU 요금뿐만 아니라 에너지 스케쥴링 결정으로 인한 추가 DC 비용 문제를 효과적으로 해결할 수 있다. TDAS 정책은 수학식 10과 같이 정의되지만, 수학식 17의 증강 상태와 타임 슬롯 t에서 n번째 EMS의 에너지 스케쥴링 결정을 사용할 수 있다. 따라서, 타임 슬롯 t에서 n번째 EMS는 $\pi_n(\bar{s}_n^t)$ 로 표시되는 TDAS 정책을 사용하여 해당 타임 슬롯에서 주어진 증강 상태에 따라 에너지 스케쥴링 결정을 수행할 수 있다.

[0092] 이하에서는 TDAS 정책을 이용하여 에너지 스케쥴링 결정을 수행하면서 최적의 TDAS 정책을 학습하는 DRL 기반 에너지 스케쥴링 알고리즘에 대해 설명한다.

[0093] 도 3을 참조하면, DRL은 주어진 상태(state)에서 어떤 행동(action)을 했을 때 적절한 보상(reward)이 발생하는 환경에서 반복적으로 시행 착오를 겪으면서 미래의 기대 보상(Q-value)이 최대화되도록 에이전트(agent)를 학습시키는 방법이다. 환경(environment)은 에이전트에게 행동을 결정하는 근거가 되는 상태 정보를 제공하며, 에이전트는 환경으로부터 획득한 상태 정보를 기반으로 행동을 결정할 수 있다.

[0094] 임의의 시점 t에서 에이전트는 환경으로부터 획득한 상태 정보 S_t에 기초하여 가능한 행동들의 확률 분포에 따라 임의의 행동 A_t을 결정하고, 환경으로부터 업데이트된 상태 정보 S_{t+1}와 보상 R_{t+1}을 받는다. 이러한 상호 작용에 기반하여 에이전트는 주어진 환경에서 누적된 보상값을 최대화하는 정책을 학습한다. 여기서, 정책이란 에이전트가 특정 상태에 대해 특정 행동을 할 확률에 관한 집합을 의미할 수 있다. 본 발명에서 상기 에이전트는 에너지 스케쥴링 정책 모델이라 불릴 수 있다.

[0095] 본 발명의 일 실시예에 따른 TDAS 알고리즘(또는 TDAS 모델)은 DRL을 기반으로 학습될 수 있다. 이를 위해 일 예로 DRL 방식 중 하나인 DQL(Deep Q-Learning)이 사용될 수 있다. DQL은 상태-행동 값 함수 Q(s, a)를 표현하기 위해 심층신경망(deep neural network, DNN)을 사용한다. 따라서 이하에서는 DNN을 Q-네트워크라 한다. Q-네트워크는 경험(또는 경험 정보)을 사용하여 최적의 상태-행동 값 함수 Q^*(s, a)를 근사화하도록 훈련된다. 이 절차를 통해 DQL 방법을 사용하면 시스템 불확실성에 대한 사전 정보 없이 복잡한 문제를 해결할 수 있다.

[0096] 본 발명의 일 실시예에 따른 EMS는 일 예로 표 1과 같은 DRL 기반 TDAS 알고리즘을 기반으로 해당 가구에 대한 에너지 스케쥴링을 수행할 수 있다.

표 1

Algorithm 1 DRL-Based TDAS Algorithm

Initialize replay memory $\mathcal{D}$
Initialize the weights $\boldsymbol{\theta}$ randomly; $\boldsymbol{\theta}^{-} \leftarrow \boldsymbol{\theta}$
Observe initial state s^0
for timeslot $t \in \mathcal{T} = \{0, 1, \ldots\}$ **do**
    Run DQL($t, \boldsymbol{\theta}, \boldsymbol{\theta}^{-}, \mathcal{D}$)
end for
procedure DQL($t, \boldsymbol{\theta}, \boldsymbol{\theta}^{-}, \mathcal{D}$)
    With probability ϵ select a random action a^t
    otherwise select $a^t = \min_{a \in \mathcal{A}(s^t)} Q(\bar{s}^t, a; \boldsymbol{\theta})$
    Operate action a^t and observe cost c^t and the next state $\bar{s}^{t+1}$
    Store experience $(\bar{s}^t, a^t, c^t, \bar{s}^{t+1})$ in $\mathcal{D}$
    if $h^t = H - 1$ **then**
        Sample random minibatch of experiences $(\bar{s}^i, a^i, c^i, \bar{s}^{i+1})$ from $\mathcal{D}$
        Set target $y^i(\boldsymbol{\theta}^{-}) = c^i + \gamma \min_{a' \in \mathcal{A}(s^{i+1})} Q(\bar{s}^{i+1}, a'; \boldsymbol{\theta}^{-})$
        Set a loss function $L(\boldsymbol{\theta}, \boldsymbol{\theta}^{-}) = (y^i(\boldsymbol{\theta}^{-}) - Q(\bar{s}^i, a^i; \boldsymbol{\theta}))^2$
        Perform a gradient descent step on L w.r.t weights $\boldsymbol{\theta}$ using learning rate α
    end if
    if $\mathrm{mod}(t, T_{target}) = 0$ **then**
        Copy Q-network weights from $\boldsymbol{\theta}$ to $\boldsymbol{\theta}^{-}$
    end if
end procedure

[0097]

[0098] DQL 방법을 적용하여 최적의 TDAS 정책을 학습하려면, 먼저 TDAS 정책의 상태 $\bar{s}^t$를 DRL에 적합한 형식으로 변환해야 한다. 구체적으로, 각 날짜에서 현재 기간을 지정하기 위한 타임 슬롯 t의 상태에서 기간 h^t은 H = {0, 1, ..., H-1}의 값 중 하나를 갖는다. 그러나 기간 h^t를 이러한 수치변수로 생각하면, 알고리즘은 일반적으로 존재하지 않는 h^t의 크기와 비용 사이의 관계를 찾으려고 한다. 따라서 알고리즘에서 기간은 범주형 변수로 간주될 수 있다. 이를 위해 기간 h^t에 원-핫 인코딩(one-hot encoding)이 적용하여 기간 h^t를 원-핫 인코딩된 벡터 h^t_o로 변환할 수 있다. 예를 들어 h^t = 2는 원-핫 인코딩된 벡터 h^t_o = (0, 0, 1, 0, ..., 0)으로 변환될 수 있다.

[0099] 표 1을 참조하면, EMS는 추후에 사용하기 위한 과거 경험을 저장하는 재생 메모리를 초기화할 수 있다. TDAS 알고리즘에서 두 개의 Q-네트워크(메인 Q-네트워크 및 타겟 Q-네트워크)가 구성될 수 있다. 타겟 Q-네트워크는 나중에 타겟 Q-값을 계산할 때 사용될 수 있다. 메인 Q-네트워크에 대한 가중치 파라미터는 Θ로 나타낼 수 있고, 타겟 Q-네트워크의 가중치 파라미터는 Θ^{-}로 나타낼 수 있다. 초기 상태는 s^0로 관찰된다. EMS는 타임 슬롯 t에서 메인 DQL 절차(메인 Q-네트워크에 대한 DQL 절차)를 수행할 수 있다. 메인 DQL 절차에서 TDAS 정책 a^t의 동작(즉, 에너지 스케줄링 결정)은 입실론 그리디(epsilon greedy) 정책을 기반으로 결정될 수 있다. EMS는 확률이 ϵ인 경우 사용 가능한 에너지 스케줄링 결정 집합 $A(s^t)$에서 무작위로 행동(에너지 공급량)을 선택하고, 확률 1 - ϵ인 경우 증강 상태 $\bar{s}^t$에서 가장 큰 Q 값을 갖는 행동 $a^t = \min_{a \in \mathcal{A}(s^t)} Q(\bar{s}^t, a; \boldsymbol{\theta})$을 선택할 수 있다. 이후, EMS는 행동 a^t를 수행하고 그에 따른 비용 c^t과 다음 증강 상태 $\bar{s}^{t+1}$을 관찰하여 경험 정보 $(\bar{s}^t, a^t, c^t, \bar{s}^{t+1})$를 생성하고, 이를 재생 메모리 D에 저장할 수 있다. 이 때, 타임 슬롯 t의 기간이 하루의 마지막 기간인 경우 즉, h^t = H - 1인 경우, EMS는 재생 메모리에 저장된 경험 정보를 배치 사이즈만큼 샘

플링하고, 샘플링된 정보와 입실론 그리디 정책을 사용하여 학습을 진행할 수 있다. 이 때, EMS는 샘플링된 경험 정보와 타겟 Q-네트워크를 기반으로 다음의 수학식 18과 같이 타겟 Q-값을 계산할 수 있다.

수학식 18

[0100] $$y^i(\boldsymbol{\theta}^-) = c^i + \gamma \min_{a' \in \mathcal{A}(s^{i+1})} Q(\bar{s}^{i+1}, a'; \boldsymbol{\theta}^-)$$

[0101] 여기서 i는 미니 배치(minibatch)에서 샘플링된 경험 정보의 인덱스를 나타낸다. 손실 함수는 타겟 Q-값을 기반으로 다음의 수학식 19와 같이 설정될 수 있다.

수학식 19

[0102] $$L(\boldsymbol{\theta}, \boldsymbol{\theta}^-) = (y^i(\boldsymbol{\theta}^-) - Q(\bar{s}^i, a^i; \boldsymbol{\theta}))^2$$

[0103] EMS는 학습률 α를 사용하여 가중치 파라미터 Θ에 대한 손실 함수에 대한 경사 하강법(gradient descent method)으로 목표 오차가 최소화되도록 메인 Q-네트워크를 훈련시켜 메인 Q-네트워크의 가중치 파라미터 Θ를 업데이트할 수 있다. 타켓 Q-네트워크의 가중치 파라미터 Θ^-는 주기적으로 메인 Q-네트워크의 가중치 파라미터 Θ로 업데이트될 수 있다.

[0104] 재생 메모리의 무작위 미니 배치는 Q-네트워크를 훈련시킬 때 경험 간의 상관 관계를 줄일 수 있다. 또한, 메인 Q-네트워크와 타겟 Q-네트워크의 분리는 타겟 Q-값의 변화로 인한 메인 Q-네트워크 훈련 중 비정상성을 완화할 수 있다.

[0105] 도 4는 본 발명의 일 실시예에 따른 에너지 스케쥴링 장치를 나타내는 도면이다.

[0106] 도 4를 참조하면, 에너지 스케쥴링 장치(400)는 수신부(410), 업데이트부(420), 및 송신부(430)를 포함한다.

[0107] 수신부(410)는 복수개의 EMS들 각각으로부터 로컬 에너지 스케쥴링 정책 모델에 대한 로컬 가중치 파라미터의 변화량 정보를 수신하도록 구성될 수 있다. 여기서, 상기 로컬 에너지 스케쥴링 정책 모델은 단위 시간에서 각 EMS의 에너지 수요량, 상기 단위 시간에서 상기 각 EMS 내 재생 에너지원의 에너지 생산량, 상기 단위 시간에서 상기 EMS 내 에너지 저장소의 에너지 잔여량, 상기 단위 시간에서 상기 각 EMS에 대한 에너지 공급원의 TOU 요금 정보, 상기 각 EMS에 대한 DC 값 등을 기반으로 학습될 수 있다.

[0108] 업데이트부(420)는 수신부(410)에서 수신된 로컬 가중치 파라미터의 변화량 정보를 기반으로 연합강화학습(federated reinforcement learning)을 수행하여 중앙 에너지 스케쥴링 정책 모델의 중앙 가중치 파라미터를 업데이트하도록 구성될 수 있다.

[0109] 송신부(430)는 업데이트부(420)에서 업데이트된 중앙 가중치 파라미터를 복수개의 EMS들로 송신 또는 브로드캐스팅(broadcasting)하도록 구성될 수 있다. 각 EMS는 송신부(430)로부터 수신한 중앙 가중치 파라미터로 로컬 가중치 파라미터를 업데이트할 수 있다.

[0110] 에너지 스케쥴링 장치(400)는 FRL을 이용하여 각 EMS가 다양한 EMS의 환경을 고려하여 효과적으로 TDAS 정책을 학습하도록 할 수 있다. 그러나 각 EMS는 단일 가구에서의 에너지 사용에 초점을 맞추고 있기 때문에 TDAS 정책 구조를 FRL에 그대로 사용하는 것은 적절하지 않다. 따라서 다수 가구용 TDAS에 FRL을 적용하기 위해서는 모든 EMS에 사용할 수 있는 범용 TDAS 정책을 구축할 필요가 있다. 이를 위하여 모든 EMS들의 상태를 나타낼 수 있는 범용 상태 공간 S_U이 정의될 수 있다. 각 EMS의 상태는 서로 동일한 유형의 정보를 가지므로, 각 EMS의 증강된 상태 공간이 범용 상태 공간으로서 사용될 수 있다. 그러나, 각 EMS마다 TOU 요금 단가 $p^H_{T,n}$와 DC 요금 단가 $p_{D,n}$는 서로 다를 수 있다. 따라서, 이러한 문제를 해결하기 위해 $p^t_{T,n}$와 $p_{D,n}$을 로컬 상태 정보로 간주하여 다음의 수학식 20과 같이 범용 상태 공간(범용 상태 정보)을 정의할 수 있다.

수학식 20

[0111] $$\mathcal{S}_U = \mathcal{L}_U^{T_f+1} \times \mathcal{G}_U^{T_f+1} \times \mathcal{B}_U \times \mathcal{C} \times \mathcal{P}_T^H \times \mathcal{P}_D \times \mathcal{H}$$

[0112] 여기서, $L_U = [0, L^{max}_U]$, $G_U = [0, G^{max}_U]$, $B_U = [0, B^{max}_U]$이며, P^H_T는 TOU 단가 집합(각 MES에 대한 TOU 단위 요금들의 집합)을 나타내고, P_D는 DC 요금의 단가 집합(각 EMS에 대한 DC 단위 요금들의 집합)을 나타낸다. 행동 공간의 경우, 모든 EMS가 후보 에너지 공급양의 집합인 C를 행동 공간으로 사용하기 때문에 새로운 범용 행동 공간을 도입할 필요가 없다. 범용 상태 공간과 행동 공간으로 범용 TDAS 정책은 $\pi_U : \mathcal{S}_U \rightarrow \mathcal{C}$로 정의된다.

[0113] 범용 TDAS 정책 구조로 상태-행동 가치 함수를 나타내는 DNN(범용 에너지 스케쥴링 모델)은 FRL을 기반으로 훈련될 수 있으며, 각 EMS의 DNN(로컬 에너지 스케쥴링 모델)은 DRL을 기반으로 분산 방식으로 훈련될 수 있다.

[0114] 예를 들어, 에너지 스케쥴링 장치(이하, FRL 코디네이터라 함)는 중앙의 범용 TDAS 정책을 구성할 수 있으며, 모든 EMS들은 범용 TDAS 정책을 사용하여 에너지 스케쥴링 결정을 분산 방식으로 결정할 수 있다. 이하에서 FRL에 대한 라운드(round) 인덱스를 r로 나타낸다. 그리고,각 라운드의 지속 시간(duration)은 타임 슬롯 T_{RFR}으로 나타낸다. 라운드 세트는 R = {1, 2, ...}로 정의된다. 범용 TDAS 정책에 대한 Q-네트워크의 중앙 가중치 파라미터는 Θ_C로 나타낸다. 또한 n번째 EMS의 메인 Q-네트워크의 로컬 가중치 파라미터는 Θ_n으로 나타낸다. 각 라운드 동안 각 EMS는 DRL을 기반으로 로컬 가중치 파라미터를 훈련시키고, 각 라운드가 끝날 때 FRL 코디네이터는 로컬에서 훈련된 가중치 파라미터를 집계할 수 있다. 이 경우, FRL에 대한 최적화 문제는 다음의 수학식 21과 같이 정의될 수 있다.

수학식 21

[0115] $$\underset{\boldsymbol{\theta}_C}{\text{minimize}} \ \frac{1}{K} \sum_{n \in N} \sum_{k=1}^{K_n} f_n(\boldsymbol{\theta}_C, k)$$

[0116] 여기서, K_n은 n번째 EMS의 경험 수(number of experiences)를 나타내고, $f_n(\Theta, k)$는 Θ 정책을 사용하는 n번째 EMS의 k번째 경험 기반 손실 함수를 나타낸다.

[0117] 수학식 21의 문제를 해결하기 위해 각 라운드마다 모든 EMS는 로컬 경험을 사용하여 로컬 가중치를 학습시키고, 이를 FRL 코디네이터로 업로드할 수 있다. FRL 코디네이터는 로컬 가중치를 다음의 수학식 22와 같이 집계하여 중앙 가중치를 업데이트할 수 있다.

수학식 22

[0118] $$\boldsymbol{\theta}_C^{r+1} = \boldsymbol{\theta}_C^r - \sum_{n \in N} c_n \nabla g_n^r$$

[0119] 여기서 ∇g_n^r는 라운드 r에서 n번째 EMS의 로컬 가중치의 변화량(gradient)을 나타내고, c_n는 n번째 EMS의 로컬 가중치의 변화량에 대한 중앙 가중치를 나타낸다.

[0120] 중앙 가중치에 대한 n번째 EMS의 기여도에 따라 c_n에 대한 학습 가중치는 다음의 수학식 23과 같이 결정될 수 있다.

수학식 23

[0121] $$c_n = \frac{K_n}{\sum_{n' \in \mathcal{N}} K_{n'}}$$

[0122] FRL 코디네이터는 수학식 22를 기반으로 중앙 가중치 파라미터를 업데이트한 후 업데이트된 중앙 가중치 파라미터를 모든 EMS에 재분배할 수 있다.

[0123] FRL 코디네이터는 일 예로 다음의 표 2와 같은 FRL 기반 다수 가구에 대한 TDAS 알고리즘을 사용할 수 있다.

표 2

[0124]

Algorithm 2 Cooperative Multi-Residential TDAS Algorithm with FRL

The FRL coordinator initializes the central weights, $\boldsymbol{\theta}_C$
Each EMS n initializes its replay memory $\mathcal{D}_n$
Each EMS n initializes its local weights $\boldsymbol{\theta}_n, \boldsymbol{\theta}_n^-$; $\boldsymbol{\theta}_n, \boldsymbol{\theta}_n^- \leftarrow \boldsymbol{\theta}_C$
Each EMS n observes initial state s_n^0
for timeslot $t \in \mathcal{T} = \{0, 1, \ldots\}$ **do**
  for EMS $n \in \mathcal{N}$ **do** ▷ at EMS n in parallel
    Run DQL($t, \boldsymbol{\theta}_n, \boldsymbol{\theta}_n^-, \mathcal{D}_n$) in Algorithm 1
  end for
  if mod(t, T_{FRL}) = 0 **then**
    Each EMS n calculates the local gradients ∇g_n using $\boldsymbol{\theta}_C$ and $\boldsymbol{\theta}_n$
    Each EMS n reports ∇g_n to the FRL coordinator
    The FRL coordinator updates $\boldsymbol{\theta}_C$ as in (22)
    $\boldsymbol{\theta}_n, \boldsymbol{\theta}_n^- \leftarrow \boldsymbol{\theta}_C$, $\forall n \in \mathcal{N}$
  end if
end for

[0125] 표 2를 참조하면, FRL 코디네이터는 범용 정책 모델을 저장할 중앙 Q-네트워크를 생성하고, 범용 TDAS 정책에 대한 Q-네트워크의 중앙 가중치 파라미터 Θ_C를 초기화한다. 각 EMS는 재생 메모리 Dn을 초기화하고, 메인 Q-네트워크의 가중치 파라미터 Θ_n 및 타켓 Q-네트워크의 가중치 파라미터 $\bar{\Theta}_n$를 중앙 가중치 파라미터 Θ_C로 각각 초기화한다.

[0126] 타임 슬롯 t에서, 각 EMS는 표 1의 DRL 기반 TDAS 알고리즘을 병렬적 또는 개별적으로 수행하고 로컬 경험을 사용하여 로컬 가중치를 훈련시킨다. 그리고, 각 EMS는 중앙 가중치 파라미터 Θ_C와 로컬 가중치 파라미터 Θ_n를 사용하여 로컬 가중치 파라미터의 변화량 ∇g_n^r 를 계산한다. 그리고, 각 EMS는 로컬 모델 집계를 위해 FRL 코디네이터로 로컬 가중치 파라미터의 변화량 정보를 전송한다. 여기서, 로컬 가중치 파라미터의 변화량 ∇g_n^r는 r번째 라운드에서 중앙 가중치 파라미터 Θ_C^r와 로컬 가중치 파라미터 Θ_n^r 간의 차이를 나타낸다.

[0127] FRL 코디네이터는 수학식 22와 같이 모든 EMS의 로컬 가중치 파라미터의 변화량을 집계하여 중앙 가중치 파라미터 Θ_C를 계산할 수 있다. 이후 FRL 코디네이터는 중앙 가중치 파라미터를 모든 EMS들에게 브로드캐스트할 수 있으며, 각 EMS는 로컬 가중치 파라미터를 FRL 코디네이터로부터 수신한 중앙 가중치 파라미터로 대체할 수 있다. 따라서, 각 EMS는 표 2와 같은 알고리즘을 통해 범용 TDAS 정책을 분산 방식으로 학습할 수 있다. 이는 EMS의 관점에서 다른 EMS의 경험을 활용하여 TDAS 정책을 학습하는 효과를 제공하므로, EMS가 효율적으로 낮은 전체 비용을 달성할 수 있도록 한다. 또한 표 2와 같은 알고리즘을 사용하는 경우, 각 EMS가 에너지 사용량 데

이터를 코디네이터로 전송하지 않기 때문에 개인 정보 보호 및 보안 문제에서 자유롭다.

[0128] 도 5는 본 발명의 일 실시예에 따른 에너지 관리 방법을 나타내는 도면이다.

[0129] 이하, 도 5를 참조하여 각 가구에 구비된 EMS가 해당 가구의 에너지를 관리하는 방법에 대해 설명한다.

[0130] 도 5를 참조하면, 각 EMS 내 에너지 관리 장치는 제1 타임 슬롯에서의 해당 EMS에 대한 상태 정보를 획득할 수 있다(S500). 여기서, 상기 상태 정보는 해당 타임 슬롯에서 해당 EMS의 에너지 수요량 L^t, 해당 타임 슬롯에서 해당 EMS 내 재생 에너지원의 에너지 생산량 G^t, 해당 타임 슬롯에서 해당 EMS 내 에너지 저장소의 에너지 잔여량 B^t, 해당 타임 슬롯에서 상기 에너지 공급원의 TOU 요금(단위 요금) 정보 P^t_T, 해당 EMS에 대한 DC 요금(단위 요금) 정보 P_D, DC 값 C^t_d, 시간 정보 h^t 등을 포함할 수 있다. 여기서, 상기 DC 값은 현재 주기에서 현재 시간까지 에너지 공급원으로부터 공급받은 최대 에너지 공급량을 기반으로 도출될 수 있다.

[0131] 상기 에너지 관리 장치는 상기 상태 정보를 기초로 상기 제1 타임 슬롯에서 에너지 공급원으로부터 공급받을 에너지 공급량 C^t을 선택할 수 있다(S510). 일 예로, 상기 에너지 관리 장치는 입실론 그리디 정책을 이용하여 각 타임 슬롯에서의 행동(에너지 공급원으로부터 공급받을 에너지 공급량)을 선택할 수 있다.

[0132] 상기 에너지 관리 장치는 제1 타임 슬롯에서 선택된 에너지 공급량과 제1 타임 슬롯에서의 해당 EMS에 대한 상태 정보를 기반으로 에너지 스케쥴링 정책 모델을 이용하여 제1 타임 슬롯에서의 에너지 요금에 대한 정보 및 제2 타임 슬롯에서의 상기 EMS에 대한 상태 정보를 획득할 수 있다(S520). 그리고, 상기 에너지 관리 장치는 제1 타임 슬롯에서의 해당 EMS에 대한 상태 정보 s^t, 제1 타임 슬롯에서 해당 에너지 공급원으로부터 공급받을 에너지 공급량 a^t, 제1 타임 슬롯에서 공급받는 에너지에 대한 요금 정보 c^t 및 제2 타임 슬롯에서의 해당 EMS에 대한 상태 정보 s^{t+1}를 기반으로 경험 정보를 구성하고 이를 이용하여 에너지 스케쥴링 정책 모델을 업데이트할 수 있다(S530). 여기서, 상기 에너지 스케쥴링 정책 모델은 에너지 스케쥴링 정책의 심층강화학습을 위한 메인 심층신경망(main Q-network) 및 목표 에너지 요금의 도출을 위한 타켓 심층신경망(target Q-network)을 포함할 수 있다. 상기 타켓 심층신경망에 대한 가중치 파라미터 Θ^-는 상기 메인 심층신경망에 대한 가중치 파라미터 Θ를 기반으로 업데이트될 수 있다. 예를 들어, 상기 에너지 관리 장치는 메모리에 저장된 경험 정보를 배치 사이즈만큼 샘플링하고 샘플링된 정보와 목표 Q-값을 기반으로 수학식 18 및 수학식 19를 이용하여 에너지 스케쥴링 정책에 대한 학습을 진행할 수 있다.

[0133] 이후, 상기 에너지 관리 장치는 일정 주기마다 메인 심층신경망의 가중치 파라미터를 타겟 심층신경망의 가중치 파라미터에 복사할 수 있으며, 에너지 스케쥴링 정책 모델에 대한 로컬 가중치 파라미터(Θ 및/또는 Θ^-)의 변화량 정보를 에너지 스케쥴링 장치로 전송하고, 상기 에너지 스케쥴링 장치로부터 상기 에너지 스케쥴링 장치의 중앙 에너지 스케쥴링 정책 모델에 대한 중앙 가중치 파라미터 Θ_C를 수신할 수 있다. 여기서, 상기 에너지 스케쥴링 정책 모델에 대한 로컬 가중치 파라미터(Θ 및/또는 Θ^-)는 상기 중앙 가중치 파라미터 Θ_C를 기반으로 업데이트될 수 있다.

[0134] 도 6는 본 발명의 일 실시예에 따른 에너지 스케쥴링 방법을 나타내는 도면이다.

[0135] 에너지 스케쥴링 장치(FRL 코디네이터)는 범용 에너지 스케쥴링 정책 모델을 저장할 중앙 심층신경망(central Q-network)를 생성하고 상기 중앙 심층신경망에 대한 가중치 파라미터를 초기화할 수 있다. 그리고, 에너지 스케쥴링 장치는 일정 주기마다(예를 들어, mode(t, T_{FRL}) = 0이 되는 주기) 복수개의 EMS들 각각으로부터 로컬 에너지 스케쥴링 정책 모델에 대한 로컬 가중치 파라미터의 변화량 정보를 수신할 수 있다(S600).

[0136] 에너지 스케쥴링 장치는 각 EMS로부터 수신한 로컬 가중치 파라미터의 변화량 정보를 기반으로 수학식 22를 기반으로 중앙 가중치 파라미터 Θ_C와 로컬 가중치 파라미터 Θ_n의 변화량을 계산하여 중앙 가중치 파라미터 Θ_C를 업데이트하고(S610), 업데이트된 중앙 가중치 파라미터를 복수개의 EMS들 각각으로 송신할 수 있다(S620).

[0137] 한편, 전술한 실시예에 따른 EMS의 에너지 관리 장치가 수행하는 에너지 관리 방법 및 에너지 스케쥴링 장치(또는 코디네이터)가 수행하는 에너지 스케쥴링 방법에 포함된 각각의 단계는, 이러한 단계를 프로세서가 수행하도

록 하기 위한 명령어를 포함하는 컴퓨터 프로그램으로 구현될 수 있다.

[0138] 또한, 전술한 실시예에 따른 에너지 관리 장치가 수행하는 에너지 관리 방법 및 에너지 스케쥴링 장치가 수행하는 에너지 스케쥴링 방법에 포함된 각각의 단계는, 이러한 단계를 프로세서가 수행하도록 하기 위한 명령어를 포함하는 컴퓨터 프로그램이 기록된 컴퓨터 판독 가능한 기록매체에서 구현될 수 있다.

[0139] 본 발명에 첨부된 흐름도의 각 단계의 조합들은 컴퓨터 프로그램 인스트럭션들에 의해 수행될 수도 있다. 이들 컴퓨터 프로그램 인스트럭션들은 범용 컴퓨터, 특수용 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비의 프로세서에 탑재될 수 있으므로, 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비의 프로세서를 통해 수행되는 그 인스트럭션들이 흐름도의 각 단계에서 설명된 기능들을 수행하는 수단을 생성하게 된다. 이들 컴퓨터 프로그램 인스트럭션들은 특정 방식으로 기능을 구현하기 위해 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비를 지향할 수 있는 컴퓨터 이용 가능 또는 컴퓨터 판독 가능 기록매체에 저장되는 것도 가능하므로, 그 컴퓨터 이용가능 또는 컴퓨터 판독 가능 기록매체에 저장된 인스트럭션들은 흐름도의 각 단계에서 설명된 기능을 수행하는 인스트럭션 수단을 내포하는 제조 품목을 생산하는 것도 가능하다. 컴퓨터 프로그램 인스트럭션들은 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비 상에 탑재되는 것도 가능하므로, 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비 상에서 일련의 동작 단계들이 수행되어 컴퓨터로 실행되는 프로세스를 생성해서 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비를 수행하는 인스트럭션들은 흐름도의 각 단계에서 설명된 기능들을 실행하기 위한 단계들을 제공하는 것도 가능하다.

[0140] 또한, 각 단계는 특정된 논리적 기능(들)을 실행하기 위한 하나 이상의 실행 가능한 인스트럭션들을 포함하는 모듈, 세그먼트 또는 코드의 일부를 나타낼 수 있다. 또한, 몇 가지 대체 실시예들에서는 단계들에서 언급된 기능들이 순서를 벗어나서 발생하는 것도 가능하다. 예컨대, 잇달아 도시되어 있는 두 개의 단계들은 실질적으로 동시에 수행되거나 또는 그 단계들이 때때로 해당하는 기능에 따라 역순으로 수행되는 것도 가능하다.

[0141] 이상의 설명은 본 발명의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 발명의 본질적인 품질에서 벗어나지 않는 범위에서 다양한 수정 및 변형이 가능할 것이다. 따라서, 본 발명에 개시된 실시예들은 본 발명의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예에 의하여 본 발명의 기술 사상의 범위가 한정되는 것은 아니다. 본 발명의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 균등한 범위 내에 있는 모든 기술사상은 본 발명의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

부호의 설명

[0142] 110: 제1 EMS

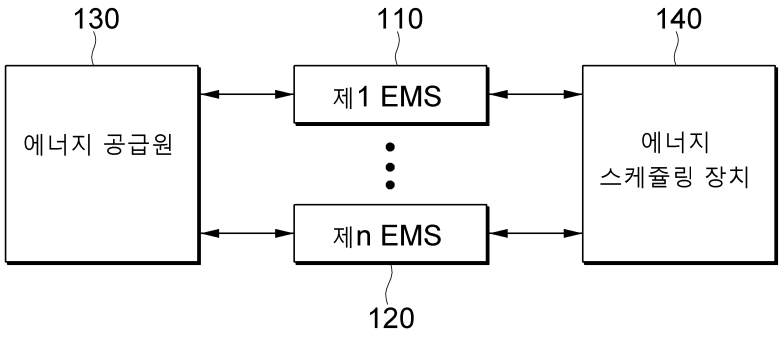
120: 제n EMS

130: 에너지 공급원

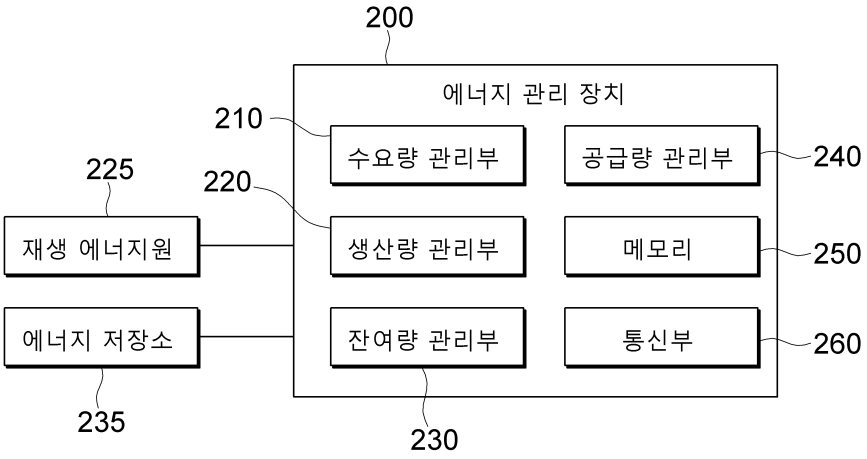
140: 에너지 스케쥴링 장치

도면

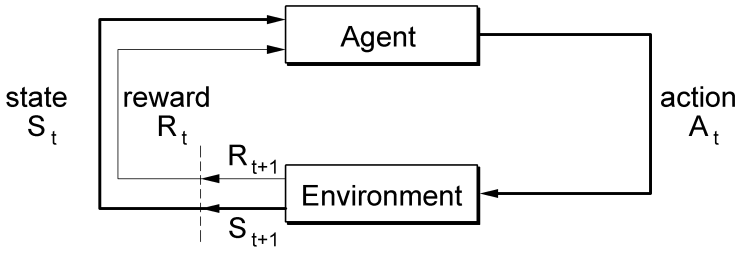
도면1



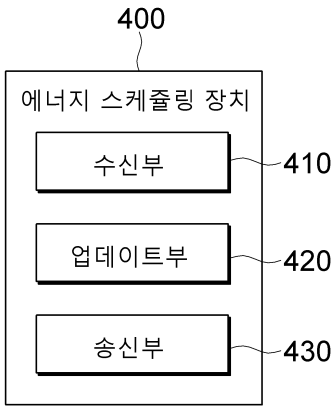
도면2


200
에너지 관리 장치
210
수요량 관리부
공급량 관리부
240
225
220
재생 에너지원
생산량 관리부
메모리
250
에너지 저장소
잔여량 관리부
통신부
260
235
230

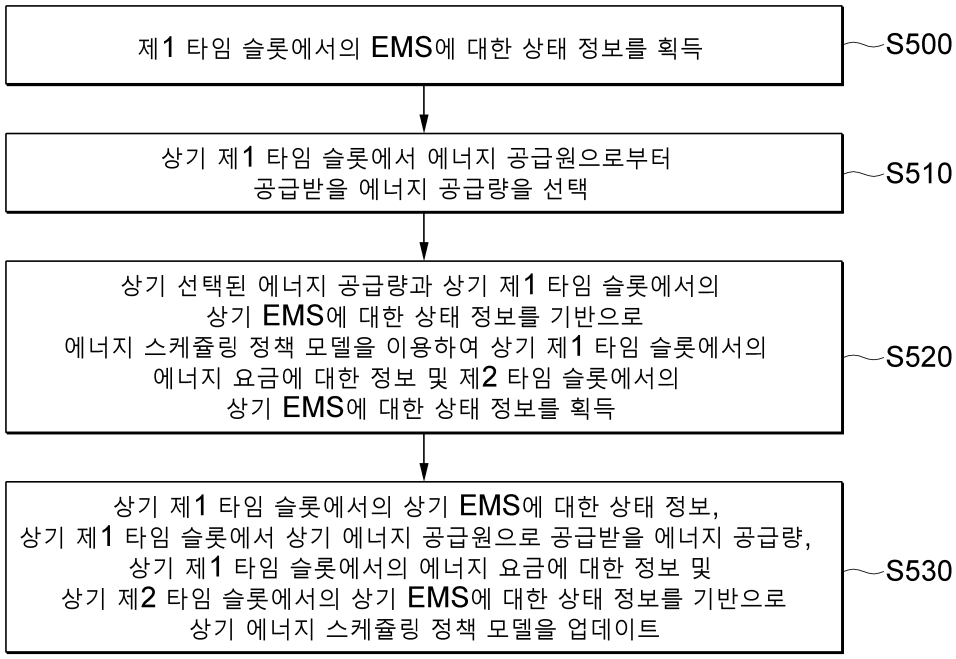
도면3


Agent
state
S_t
reward
R_t
action
A_t
R_{t+1}
Environment
S_{t+1}

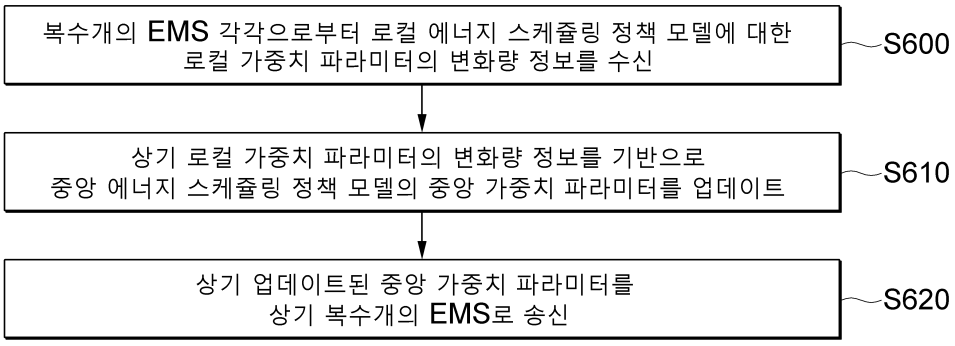
도면4


400
에너지 스케줄링 장치
수신부
410
업데이트부
420
송신부
430

도면5


제1 타임 슬롯에서의 EMS에 대한 상태 정보를 획득
S500
상기 제1 타임 슬롯에서 에너지 공급원으로부터
공급받을 에너지 공급량을 선택
S510
상기 선택된 에너지 공급량과 상기 제1 타임 슬롯에서의
상기 EMS에 대한 상태 정보를 기반으로
에너지 스케쥴링 정책 모델을 이용하여 상기 제1 타임 슬롯에서의
에너지 요금에 대한 정보 및 제2 타임 슬롯에서의
상기 EMS에 대한 상태 정보를 획득
S520
상기 제1 타임 슬롯에서의 상기 EMS에 대한 상태 정보,
상기 제1 타임 슬롯에서 상기 에너지 공급원으로 공급받을 에너지 공급량,
상기 제1 타임 슬롯에서의 에너지 요금에 대한 정보 및
상기 제2 타임 슬롯에서의 상기 EMS에 대한 상태 정보를 기반으로
상기 에너지 스케쥴링 정책 모델을 업데이트
S530

도면6


복수개의 EMS 각각으로부터 로컬 에너지 스케쥴링 정책 모델에 대한
로컬 가중치 파라미터의 변화량 정보를 수신
S600
상기 로컬 가중치 파라미터의 변화량 정보를 기반으로
중앙 에너지 스케쥴링 정책 모델의 중앙 가중치 파라미터를 업데이트
S610
상기 업데이트된 중앙 가중치 파라미터를
상기 복수개의 EMS로 송신
S620