(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0155840
(43) 공개일자 2023년11월13일

(51) 국제특허분류(Int. Cl.)
G06Q 50/06 (2012.01) *G06N 20/00* (2019.01)
G06N 3/02 (2023.01) *G06Q 10/04* (2023.01)
G06Q 10/06 (2012.01) *G06Q 30/02* (2023.01)
G06Q 30/08 (2012.01)
(52) CPC특허분류
G06Q 50/06 (2013.01)
G06N 20/00 (2021.08)
(21) 출원번호 10-2022-0055564
(22) 출원일자 2022년05월04일
심사청구일자 2022년05월04일

(71) 출원인
서강대학교산학협력단
서울특별시 마포구 백범로 35 (신수동, 서강대학교)
(72) 발명자
김홍석
서울특별시 마포구 서강대길 40, 103동 202호 (대흥동, 마포자이2차아파트)
정재익
서울특별시 마포구 신촌로24길 26-12, 2층 (노고산동)
(74) 대리인
팬코리아특허법인

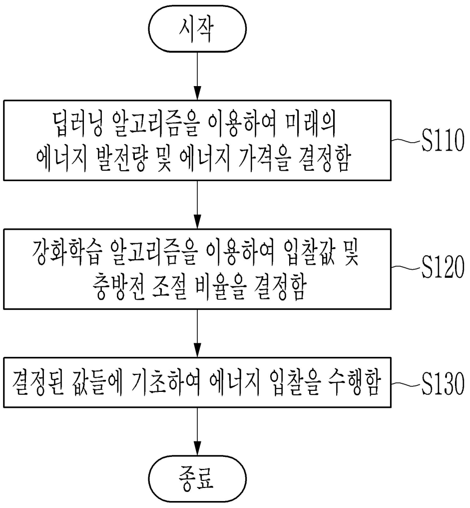
전체 청구항 수 : 총 8 항

(54) 발명의 명칭 **배터리를 활용한 수익 최대화를 위한 에너지 입찰 방법**

(57) 요 약

배터리를 활용한 수익 최대화를 위한 에너지 입찰 방법이 제공된다. 에너지 입찰 방법은 딥러닝 알고리즘을 이용하여 미래의 에너지 발전량 x_t 및 에너지 가격 λ_t 을 결정하는 단계; 강화학습 알고리즘을 이용하여 입찰값 b_t 및 충방전 조절 비율 ρ_t 을 결정하는 단계; 및 상기 결정된 값들에 기초하여 에너지 입찰을 수행하는 단계를 포함할 수 있다.

대 표 도 - 도1



(52) CPC특허분류
G06N 3/02 (2023.01)
G06Q 10/04 (2023.01)
G06Q 10/063 (2023.01)
G06Q 30/0206 (2013.01)
G06Q 30/08 (2023.01)
Y04S 10/12 (2020.08)
Y04S 50/10 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1415175444
과제번호	20192010107290
부처명	산업통상자원부
과제관리(전문)기관명	한국에너지기술평가원
연구사업명	에너지수요관리핵심기술개발(에특)(R&D)
연구과제명	가정 에너지 사용량 실시간 진단 및 지능형 자율제어/관리 시스템 원천기술
기 여 율	1/1
과제수행기관명	한국전자통신연구원
연구기간	2021.05.01~2022.04.30

명 세 서

청구범위

청구항 1

딥러닝 알고리즘을 이용하여 미래의 에너지 발전량 x_t 및 에너지 가격 λ_t을 결정하는 단계;

강화학습 알고리즘을 이용하여 입찰값 b_t 및 충방전 조절 비율 ρ_t 을 결정하는 단계; 및

상기 결정된 값들에 기초하여 에너지 입찰을 수행하는 단계를 포함하는

에너지 입찰 방법.

청구항 2

제1항에 있어서,

입찰 오차 보정 방식을 위해 과소 예측일 경우 $(b_t < x_t)$ $x_t - b_t$ 만큼 충전하고 과대 예측일 경우 $(b_t > x_t)$ $b_t - x_t$ 만큼 방전하는 단계를 더 포함하는 에너지 입찰 방법.

청구항 3

제2항에 있어서,

차익거래 방식을 위해 여기서 결정된 충방전량을 ρ_t 의 비율로 추가로 조절하는 단계를 더 포함하는 에너지 입찰 방법.

청구항 4

제1항에 있어서,

상기 강화학습 알고리즘은 마르코프 결정 과정(Markov Decision Process; MDP)에 기반한 강화학습 알고리즘을 포함하는, 에너지 입찰 방법.

청구항 5

제4항에 있어서,

상기 MDP는 상태(state, s_t), 행동(action, a_t), 보상(reward, r_t)을 포함하고, 시계열의 불확실한 상황에서 s_t 는 관측(observation, o_t)의 시계열 데이터들을 포함하는, 에너지 입찰 방법.

청구항 6

제1항에 있어서,

시계열 데이터를 학습하기 위해 딥러닝 모델 중 LSTM(Long Short-Term Memory)이 사용되는, 에너지 입찰 방법.

청구항 7

제1항에 있어서,

모델 업데이트를 위해 추정된 가치함수 $V_\theta(s_t)$가 사용되는, 에너지 입찰 방법.

청구항 8

제1항에 있어서,

배터리 충방전으로 차익거래 및 입찰 오차 보정 효과를 얻은 후 최종 수익을 이용하여 보상 r_t를 결정하는 단계; 및

보상과 추정된 가치함수로 모델을 업데이트하고 배터리 충방전 후 값 및 다음 시간대에 관측된 에너지 발전량 x_t와 에너지 가격 λ_t로 다음 상태를 결정하는 단계를 더 포함하는 에너지 입찰 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 배터리를 활용한 수익 최대화를 위한 에너지 입찰 방법에 관한 것이다.

배 경 기 술

[0002] 발전된 에너지를 전력시장에 입찰하기 위해 배터리를 활용하여 최종 수익을 극대화하기 위한 방안이 연구되고 있다. 최종 수익을 최대화하기 위한 기술로, 차익거래를 활용하는 기술과 입찰 오차 보정을 활용하는 기술을 들 수 있다. 여기서, 차익거래는 에너지 가격이 낮을 때 충전하고 에너지 가격이 높을 때 방전하여 이익을 증대하는 방식이고, 입찰 오차 보정은 발전량보다 큰 입찰값(과대 입찰)만큼 방전하고, 발전량보다 작은 입찰값(과소 입찰)만큼 충전하여 오차 패널티를 감축하는 방식이다.

[0003] 차익거래를 활용하는 기술은 에너지 가격 불확실성만 고려하고 발전량의 불확실성을 고려하지 않은 경우가 대다수이며, 발전량의 불확실성을 고려했다 하더라도 입찰 정확도가 높다는 가정 하에 오차 보정을 고려하지 않았다. 즉, 기존 입찰값에 방전량만큼 더하거나 충전량만큼 뺀 값을 새로운 입찰값으로 만들어, 기존 오차(발전량 - 기존 입찰값)와 동일한 오차(발전량 + 방전량 - 충전량 - 새로운 입찰값)를 만드는 방식이다. 결과적으로 차익거래로 인해 이익은 증대하지만, 오차로 인한 패널티는 기존과 똑같은 양만큼 부여받게 된다.

[0004] 한편, 입찰 오차 보정을 활용하는 기술은 발전량의 불확실성만 고려하고 에너지 가격의 불확실성을 고려하지 않는다. 재생에너지나 천연가스와 같이 미래 에너지 발전량의 불확실성이 있는 경우 발전량의 예측 기술이 사용되는데, 일반적으로 예측된 발전량을 입찰값으로 삼고 오차가 생긴 만큼 배터리로 보정(충방전)하는 전략을 사용한다. 또한 오차를 줄이는 것이 아닌 보정 가능한 오차를 만드는 전략으로 보정 후 오차를 획기적으로 감축하여 오차 패널티를 크게 줄이는 기술도 있다. 그러나 에너지 가격의 변동에 따른 차익거래는 고려되지 않았고, 배터리는 오차를 보정하기 위해서만 사용되고 있다.

발명의 내용

해결하려는 과제

[0005] 본 발명이 해결하고자 하는 과제는, 발전된 에너지를 전력시장에 입찰할 때, 배터리를 활용하여 차익거래를 통한 이익 증대와 입찰 오차 보정을 통한 패널티 감축을 동시에 달성하여, 궁극적으로 최종 입찰 수익을 최대화할 수 있는 에너지 입찰 방법을 제공하는 것이다.

과제의 해결 수단

[0006] 본 발명의 일 실시 예에 따른 에너지 입찰 방법은, 에너지 입찰 방법은 딥러닝 알고리즘을 이용하여 미래의 에너지 발전량 x_t 및 에너지 가격 λ_t을 결정하는 단계; 강화학습 알고리즘을 이용하여 입찰값 b_t 및 충방전 조절 비율 ρ_t을 결정하는 단계; 및 상기 결정된 값들에 기초하여 에너지 입찰을 수행하는 단계를 포함할 수 있다.

[0007] 일 실시 예에서, 상기 방법은, 입찰 오차 보정 방식을 위해 과소 예측일 경우 $(b_t < x_t)$ $x_t - b_t$만큼 충전하고 과대 예측일 경우 $(b_t > x_t)$ $b_t - x_t$만큼 방전하는 단계를 더 포함할 수 있다.

[0008] 일 실시 예에서, 상기 방법은, 차익거래 방식을 위해 여기서 결정된 충방전량을 ρ_t의 비율로 추가로 조절하는 단계를 더 포함할 수 있다.

[0009] 일 실시 예에서, 상기 강화학습 알고리즘은 마르코프 결정 과정(Markov Decision Process; MDP)에 기반한 강화

학습 알고리즘을 포함할 수 있다.

[0010] 일 실시 예에서, 상기 MDP는 상태(state, s_t), 행동(action, a_t), 보상(reward, r_t)을 포함하고, 시계열의 불확실한 상황에서 s_t는 관측(observation, o_t)의 시계열 데이터들을 포함할 수 있다.

[0011] 일 실시 예에서, 시계열 데이터를 학습하기 위해 딥러닝 모델 중 LSTM(Long Short-Term Memory)이 사용될 수 있다.

[0012] 일 실시 예에서, 모델 업데이트를 위해 추정된 가치함수 $V_\theta(s_t)$가 사용될 수 있다.

[0013] 일 실시 예에서, 상기 방법은, 배터리 충방전으로 차익거래 및 입찰 오차 보정 효과를 얻은 후 최종 수익을 이용하여 보상 r_t를 결정하는 단계; 및 보상과 추정된 가치함수로 모델을 업데이트하고 배터리 충방전 후 값 및 다음 시간대에 관측된 에너지 발전량 x_t와 에너지 가격 λ_t로 다음 상태를 결정하는 단계를 더 포함할 수 있다.

발명의 효과

[0014] 본 발명의 실시 예들에 따르면, 오차 패널티 감축을 위한 입찰값과 차익거래로 이익 증대를 위한 충방전량을 동시에 결정하는 입찰 전략을 제공하며, 이는 기존에 차익거래 또는 입찰 오차 보정 중 하나에만 초점이 맞춰진 방식과 달리 두 방식의 장점을 모두 활용할 수 있어 궁극적으로 최종 입찰 수익을 극대화할 수 있다.

[0015] 또한, 발전량 및 에너지 가격 예측을 위한 딥러닝 구조에 입찰 수익 극대화를 위한 입찰값 및 충방전량을 결정하는 강화학습 알고리즘을 적용한 심층강화학습을 활용하여 발전량 및 에너지 가격 불확실성이 있는 환경(재생에너지, 천연가스 등)에서도 입찰 수익 극대화를 실현할 수 있다.

도면의 간단한 설명

[0016] 도 1은 본 발명의 일 실시 예에 따른 에너지 입찰 방법을 설명하기 위한 도면이다.

도 2는 발전량의 불확실성이 있는 재생에너지를 에너지 시장에 입찰하는 예시적인 방법을 설명하기 위한 도면이다.

도 3은 본 발명의 일 실시 예에 따른 심층강화학습 기반 입찰 알고리즘을 설명하기 위한 도면이다.

도 4는 본 발명의 실시 예들에 기초하여 수행된 입찰 모의 실험 결과를 설명하기 위한 도면이다.

도 5는 본 발명의 실시 예들에 따른 에너지 입찰 방법을 구현하기 위한 컴퓨팅 장치를 설명하기 위한 블록도이다.

발명을 실시하기 위한 구체적인 내용

[0017] 아래에서는 첨부한 도면을 참조하여 본 발명의 실시 예에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시 예에 한정되지 않는다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.

[0018] 명세서 및 청구범위 전체에서, 어떤 부분이 어떤 구성 요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성 요소를 제외하는 것이 아니라 다른 구성 요소를 더 포함할 수 있는 것을 의미한다. 또한, 명세서에 기재된 "...부", "...기", "모듈" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다. 또한, 아래에서 설명되는 실시 예들에 따른 배터리를 활용한 재생에너지 오차 보정 가능한 예측 방법은 프로그램 또는 소프트웨어로 구현될 수 있고, 프로그램 또는 소프트웨어는 컴퓨터로 판독 가능한 매체에 저장될 수 있다.

[0019] 도 1은 본 발명의 일 실시 예에 따른 에너지 입찰 방법을 설명하기 위한 도면이다.

[0020] 도 1을 참조하면, 본 발명의 일 실시 예에 따른 에너지 입찰 방법은, 딥러닝 알고리즘을 이용하여 미래의 에너지 발전량 및 에너지 가격을 결정하는 단계(S110), 강화학습 알고리즘을 이용하여 입찰값 및 충방전 조절 비율을 결정하는 단계(S120) 및 결정된 값들에 기초하여 에너지 입찰을 수행하는 단계(S130)를 포함할 수 있다.

[0021] 불확실한 에너지 발전량이 있는 환경에서 차익거래와 입찰 오차 보정은 모두 수익을 극대화할 수 있는 요소들이다. 2 가지 방식을 단순히 결합한다면 배터리 용량을 약 2배로 늘려야 하는데, 이로 인해 배터리 투자비용이 크게 증가할 수 있다. 실시 예들은, 차익거래와 입찰 오차 보정을 모두 고려하는 방법을 제공한다.

[0022] 이를 위해, 딥러닝 구조를 활용하여 발전량과 에너지 가격의 불확실성을 모두 고려하였으며, 강화학습 알고리즘을 적용한 심층강화학습으로 확장하여 수익 최대화를 위한 입찰값과 충방전량를 결정하게 하였고, 이 때 입찰값은 오차 보정이 가능하도록 결정되며 차익거래를 위한 충방전량이 추가되도록 구성되었다. 이에 따라 동일한 배터리 용량만으로 차익거래와 입찰 오차 보정을 모두 고려하여 추가적인 투자비용 없이 최종 입찰 수익을 극대화할 수 있다.

[0023] <u>에너지 입찰 모델</u>

[0024] 실시 예들에 따른 에너지 입찰 모델 및 이와 관련된 파라미터들은 다음과 같다.

표 1

t	Discrete time slot index
x_t	Real power generation value
b_t	Bidding value
λ_t	Energy price
SoC_t	State-of-charge of the battery
$E_{\max}$	Maximum Battery Capacity
$SoC_{\min}$ / $SoC_{\max}$	Minimum/Maximum state-of-charge of the battery
P_t^c / P_t^d	Charging/Discharging power of the battery
ρ_t	Scaling factor for boosting up the charging/discharging
$\overline{P}_t^c$ / $\overline{P}_t^d$	Charging/Discharging power limitation
$P_{\max}^c$ / $P_{\max}^d$	Maximum charging/discharging power of the battery
η_c / η_d	Charging/Discharging efficiency
β	Battery degradation cost
γ	Discount factor
x_t^D	Dispatched power to the grid
ρ_{pen}	Deviation penalty factor
f_t	The final profit

[0025]

[0026] 그리고 전체적으로 해결하고자 하는 문제는 다음과 같이 정의할 수 있다.

표 2

$$\underset{\{b_t,\rho_t\}_{t=1}^{\infty}}{\text{maximize}} \quad \mathbb{E}_{\{x_t,\lambda_t\}_{t=1}^{\infty}}\left[\sum_{t=1}^{\infty}\gamma^{t-1}f_t\right] \qquad (1)$$

subject to

$$\overline{P}_t^c = \min\left(P_{\max}^c,\ \frac{1}{\eta_c}\times\frac{E_{\max}\times(SoC_{\max}-SoC_t)}{\Delta t}\right) \qquad (2a)$$

$$\overline{P}_t^d = \min\left(P_{\max}^d,\ \eta_d\times\frac{E_{\max}\times(SoC_t-SoC_{\min})}{\Delta t}\right) \qquad (2b)$$

$$P_t^c = \min\left(\max(\rho_t(x_t-b_t),0),\ \overline{P}_t^c\right) \qquad (3a)$$

$$P_t^d = \min\left(\max(\rho_t(b_t-x_t),0),\ \overline{P}_t^d\right) \qquad (3b)$$

$$SoC_{t+1} = SoC_t + \eta_c\frac{P_t^c}{E_{\max}}\Delta t - \frac{1}{\eta_d}\frac{P_t^d}{E_{\max}}\Delta t \qquad (4)$$

$$x_t^D = x_t - P_t^c + P_t^d \qquad (5)$$

$$f_t = \left(\lambda_t\left(x_t^D-\rho_{pen}\left|b_t-x_t^D\right|\right)-\beta\left(P_t^c+P_t^d\right)\right)\Delta t \qquad (6)$$

[0027]

[0028] 식 (2)는 시간대 t 에서 배터리의 충전상태(SoC; State-of-Charge)가 SoC_t 일 때 다음 시간대 $t+1$에서 배터리에 최대로 충전 또는 방전할 수 있는 전력을 나타낸다. 배터리는 SoC가 매우 높거나 매우 낮을 때 열화 현상이 심하다. 따라서 SoC가 배터리 열화가 심하지 않은 범위($SoC_{\min} \sim SoC_{\max}$) 내에 있도록 제한해야 하므로 SoC_t 는 $SoC_{\min} \le SoC_t \le SoC_{\max}$ 을 만족해야 한다. 이 때 충방전 효율($0 \le \eta_c \le 1,\ 0 \le \eta_d \le 1$)과 최대 충방전 전력($P_{\max}^c, P_{\max}^d$)도 고려해야 한다. 배터리를 충전 또는 방전할 때 에너지 손실이 생기기 때문에, 충전할 때는 배터리에 에너지가 덜 쌓이게 되어 충전 가능한 양이 $1/\eta_c$ 의 비율로 늘어나게 되고, 방전할 때는 에너지 손실로 인해 방전된 양이 η_d 의 비율로 줄어들게 된다. 또한 전력변환 시스템의 조건에 의해 용량이 남아있어도 $P_{\max}^c, P_{\max}^d$ 이상으로 충전 또는 방전할 수는 없다. 이에 따라 식 (2)와 같이 시간대 t 의 SoC에 따라 다음 시간대 $t+1$ 에서 최대로 충전 또는 방전할 수 있는 양이 결정된다.

[0029] 식 (3)은 배터리에 충전 또는 방전되는 전력을 나타낸다. x_t 를 시간대 t 에서의 재생에너지 발전량 실측값으로, b_t 를 시간대 t 에서의 재생에너지 발전량 입찰값으로 정의하며, 오차는 $b_t - x_t$가 된다. 입찰 오차 보정 방식을 위해 과소 예측일 경우 $(b_t < x_t)$ $x_t - b_t$ 만큼 충전하고 과대 예측일 경우 $(b_t > x_t)$ $b_t - x_t$ 만큼 방전한다. 차익거래 방식을 위해 여기서 결정된 충방전량을 ρ_t 의 비율로 추가로 조절한다. 즉, 에이전트가 결정해야 되는 값은 b_t 와 ρ_t 2개로 나타난다. 충방전량은 각각 식 (2)에서 계산했던 $\overline{P}_t^c$와 $\overline{P}_t^d$로 제한된다.

[0030] 식 (4)는 배터리 SoC의 변화를 나타낸다. 에너지 손실로 인해 충전할 때는 충전한 전력 P_t^c에서 η_c의 비율로 덜 쌓이게 되고, 방전할 때는 $1/\eta_d$의 비율로 더 방전해야 방전해야 할 전력 P_t^d를 맞출 수 있다. 다른 식과 달리 SoC_{t+1}이 SoC_t에 의해 결정되는 시간 결합성(time-coupling)이 있는 식이다.

[0031] 식 (5)는 전력망에 급전(Dispatch)된 전력량 x_t^D를 나타낸다. 실측값에서 배터리에 충전한 양만큼 빼고 방전된 양만큼 더한 것이다. 급전된 전력량은 에너지 시장으로 보내진 전력량을 의미하고, 이에 따라 에너지 사업자의 수익이 결정된다. 수익은 시장 규칙에 따라 다르게 나타나지만, 발전된 전력을 최대한 에너지 가격이 높을 때 급전할수록, 입찰된 값과 급전된 값의 차이인 입찰 오차가 작을수록 높은 수익을 얻게 된다.

[0032] 식 (6)은 시장 규칙의 한 예시를 보여주며, 미국 PJM 실시간 에너지 시장에서 사용하는 시장 규칙이 적용된 식이다. 먼저 급전된 에너지량 $x_t^D \Delta t$에 단위 에너지당 가격 λ_t을 곱한 만큼 수익이 부여된다. 이때 입찰 오차 $\left|b_t - x_t^D\right| \Delta t$에 오차당 페널티 $\lambda_t \rho_{pen}$를 곱한만큼 페널티를 부과해야 한다. ρ_{pen}은 수익 대비 페널티를 얼마만큼의 비율로 줄 것인지를 나타내는 상수이다. 뿐만 아니라 배터리를 사용하기 때문에 배터리 사용에 대한 요금도 부과된다. SoC가 배터리 열화가 심하지 않은 범위($SoC_{\min} \sim SoC_{\max}$)에 있을 때 배터리 열화는 충방전 전력에 거의 비례한다고 알려져 있으며, 충전 에너지 $P_t^c \Delta t$ 및 방전 에너지 $P_t^d \Delta t$에 비례하는 만큼 부과된다. β는 충방전 전력당 배터리 열화에 대한 비용을 나타내는 상수이다. 따라서 최종 수익 f_t는 급전된 에너지에 대한 수익에서 오차에 대한 페널티 비용과 배터리 열화 비용을 뺀 값이 된다. 다른 시장 규칙에 대해서도 식 (6)과 같이 입찰값에 따른 수익이 어떻게 나타내는지 표현하면 적용이 가능하다. 에너지 시장은 공통적으로 에너지 가격의 변동이 있고, 입찰 오차에 따른 페널티 또는 입찰 정확도에 따른 인센티브가 부여되므로 어떠한 시장 규칙에서도 차익거래와 입찰 오차 보정 효과를 노리도록 b_t와 ρ_t를 결정하게 할 수 있다.

[0033] 궁극적으로 목적 함수는 식 (1)과 같이 나타난다. 미래의 에너지 발전량 $\{x_t\}_{t=1}^{\infty}$ 및 에너지 가격 $\{\lambda_t\}_{t=1}^{\infty}$의 분포에 따라 수익 f_t의 합을 최대화할 수 있는 입찰값 $\{b_t\}_{t=1}^{\infty}$와 충방전 조절 비율 $\{\rho_t\}_{t=1}^{\infty}$을 결정하는 것이다. 이때 감가 상각($0 < \gamma < 1$)의 의미는 먼 미래에 발생한 수익일수록 그 중요도를 낮춘 것이다. 미래의 수익까지 모두 고려해야 하는 이유는 식 (4)로 인해 시간 결합성(time-coupling)이 생기기 때문이다.

[0034] <u>심층강화학습 적용 방법</u>

[0035] 미래의 에너지 발전량 $\{x_t\}_{t=1}^{\infty}$과 에너지 가격 $\{\lambda_t\}_{t=1}^{\infty}$은 현재 알려지지 않은 불확실한 요소들이므로 과거 데이터에 기반한 기계학습 알고리즘을 필요로 한다. 이때 불확실한 요소들을 예측하기 위해 로 데이터(Raw data)에서도 의미있는 예측을 수행할 수 있는 딥러닝 알고리즘이 널리 사용되며, 본 발명의 실시 예들 역시 딥러닝에 기반할 수 있다. 또한 식 (4)로 인해 시간 결합성(time-coupling)이 있다는 것은 해결하고자 하는 문제가 입찰값 b_t와 충방전 조절 비율 ρ_t을 순차적으로 결정해야 하는 순차적 의사 결정임을 의미한다. 이는 마르코프 결정 과정(Markov Decision Process; MDP)에 기반한 강화학습 알고리즘을 사용하여야 한다. 따라서 본 발명의 실시 예들의 경우, 딥러닝과 강화학습에 기반한 심층강화학습이 적용될 수 있다.

[0036] MDP는 상태(state, s_t), 행동(action, a_t), 보상(reward, r_t)으로 구성된다. 시계열의 불확실한 상황에서 s_t는 관측(observation, o_t)의 시계열 데이터들로 구성된다. 본 발명에서 o_t, s_t, a_t, r_t는 아래와 같이 정의된다.

[0037]

$$o_t = (x_{t-1}, \lambda_{t-1}, SoC_t) \quad (7)$$

$$s_t = (o_0, o_1, \cdots, o_t) \quad (8)$$

$$a_t = (b_t, \rho_t) \quad (9)$$

$$r_t = f_t - \lambda_t \times x_t \times \Delta t \quad (10)$$

[0038] 시간대 t 에서의 에너지 발전량 x_t 와 에너지 가격 λ_t 는 미리 관측될 수 있는 사항들이 아닌 불확실한 요소들이다. 따라서 관측 o_t 의 요소에는 기존에 관측되었던 x_{t-1} 과 λ_{t-1}가 대신 들어가게 된다. 상태 s_t는 이들의 시계열 데이터로 구성되었으며, 이를 기반으로 행동 a_t 를 결정하게 된다. 에이전트가 결정해야 되는 값은 b_t와 ρ_t 이므로 행동 a_t 는 이 둘로 구성된다. 보상 r_t 는 에이전트의 목적에 의해 결정되기에 f_t로 구성할 수 있다. 그러나 f_t 는 x_t 와 λ_t 에 의해서도 결정되는데, 이는 학습 효율을 떨어뜨릴 수 있다. 예를 들어 x_t 또는 λ_t가 낮은 상황에서는 수익 f_t가 어쩔 수 없이 낮게 되는데 학습 알고리즘은 잘못된 학습으로 f_t가 낮아진 것으로 오판할 수 있다. 반대로 x_t 또는 λ_t 가 높으면 잘못된 결정을 내려도 수익 f_t가 높을 수 있는데 학습 알고리즘이 맞는 결정을 내렸다고 오판할 수 있다. 따라서 학습에 사용될 보상함수는 f_t와 기본적인 수익 $\lambda_t \times x_t \times \Delta t$ 의 차이로 재정의할 수 있다. 이를 통해 랜덤하게 높아지거나 낮아지는 보상 f_t가 아닌 에이전트의 결정에 의해 높아지거나 낮아지는 보상 r_t 를 가질 수 있게 된다.

[0039] 행동 a_t 의 구성 요소인 b_t 와 ρ_t 는 연속적인 값이므로 정책 기반 강화학습이 사용된다. 정책이란 어떤 상태 s_t 에서 어떤 행동 a_t 를 취해야 하는지 알려주는 것이고, 정책 π 는 s_t 에 따라 a_t가 취해질 확률 $\pi(a_t|s_t)$을 나타내 준다. 정책 기반 강화학습은 정책을 파라미터 θ 로 모델링하여 (이때의 정책을 π_θ라 함) $\sum_{t=1}^{\infty} \gamma^{t-1} f_t$ 을 최대화할 수 있는 파라미터 θ 를 찾는다. 이때 θ 는 딥러닝의 파라미터를 의미한다.

[0040] 딥러닝 모델의 입력은 상태 s_t 이고, 출력은 여기서의 정책 $\pi_\theta(\cdot|s_t)$이다. 일반적으로 가우시안 분포로 모델링되며 가우시안 분포의 평균 벡터를 출력한다. (여기서는 구성요소가 b_t와 ρ_t이므로 2차원 벡터이다.) 가우시안 분포의 공분산 행렬은 하이퍼파라미터로 설정되어 모델 트레이닝 과정 중에 조절된다. 구성요소가 2개이므로 2X2 단위 행렬에 설정된 분산 하이퍼파라미터 σ^2 를 곱한 값이 공분산 행렬로 사용된다. 행동 a_t는 정책 $\pi_\theta(\cdot|s_t)$에서 추출된다.

[0041] 해결해야 하는 문제인 식 (1)을 강화학습 형태로 바꾸면 아래와 같다.

[0042]

$$\underset{\theta}{\text{maximize}} \quad \mathbb{E}_{\{x_t, \lambda_t\}_{t=0}^{\infty}}\left[\sum_{t=1}^{\infty} \gamma^{t-1} r_t\right] \text{ where } a_t \sim \pi_\theta(\cdot|s_t) \quad (11)$$

[0043] 식 (1)과 달라진 점은 목적함수가 f_t 대신 r_t가 사용되었다는 점과, 모든 시간대에서의 $a_t = (b_t, \rho_t)$를 결정하는 대신 각 상태 s_t에서 어떤 a_t를 취할지 알려주는 정책 $\pi_\theta(\cdot|s_t)$을 결정한다는 점이다. f_t 대신 r_t를 사용한 것은 학습 효율을 높이기 위함이고, $\pi_\theta(\cdot|s_t)$를 결정하는 것은 모든 시간대에서의 $a_t = (b_t, \rho_t)$를 결정하는 것과 같기 때문에 식 (1)에서 목적하는 바를 이룰 수 있다. 모든 연속값 제어가 가능한 정책 기반 강화학습 알고리즘이 식 (11)을 해결하는 데 사용될 수 있다.

[0044] 도 2는 발전량의 불확실성이 있는 재생에너지를 에너지 시장에 입찰하는 예시적인 방법을 설명하기 위한 도면이

다.

[0045] 도 2를 참조하면, 발전량의 불확실성으로 유명한 재생에너지를 에너지 시장에 입찰할 때 프로세스의 예시(미국 PJM 시장)를 나타낸다. 재생에너지 사업자가 시간대 t 에서의 입찰값 b_t 와 충방전 조절 비율 ρ_t를 사전에 결정하고, 이후 태양광 또는 풍력과 같은 재생에너지 발전소에서 발전량 x_t 가 정해진다. x_t, b_t, ρ_t에 의해서 충방전 전력 P_t^c 와 P_t^d 가 정해지면 급전되는 전력량 x_t^D 와 배터리 사용 비용 $\beta\times\left(P_t^c+P_t^d\right)\times\Delta t$ 이 정해진다. 급전되는 전력량에 따라 이익 $\lambda_t\times x_t^D\times\Delta t$ 이 부여되며 x_t^D 와 b_t 의 차이로 인해 생기는 입찰 오차로 인해 페널티 $\lambda_t\times\rho_{pen}\times\left|b_t-x_t^D\right|\times\Delta t$ 도 부과된다.

[0046] 도 3은 본 발명의 일 실시 예에 따른 심층강화학습 기반 입찰 알고리즘을 설명하기 위한 도면이다.

[0047] 도 3을 참조하면, 본 발명의 일 실시 예에 따른 심층강화학습 기반 입찰 알고리즘을 도식화한 것이다. 먼저 시계열 데이터를 학습하기 위해 딥러닝 모델 중 LSTM(Long Short-Term Memory)이 사용된다. LSTM은 시간순으로 두 벡터 c_t 와 h_t 를 다음 LSTM 셀(cell)에 전달하므로, 시간대 t 에 모든 관측값 $o_0, o_1, \cdots, o_t$ 이 사용되어 결과적으로 상태 s_t 가 입력된 형태임을 알 수 있다. LSTM의 파라미터를 θ 라 하면 정책 π_θ의 평균이 출력된다. 이는 하이퍼파라미터로 설정된 정책의 공분산 행렬과 같이 가우시안 분포를 만들고 행동 a_t는 이 분포로부터 추출된다. 모델의 또 다른 출력 $V_\theta(s_t)$는 추정된 가치함수로 모델 업데이트에 사용된다. 배터리 충방전으로 차익거래 및 입찰 오차 보정 효과를 얻은 후 최종 수익을 이용하여 보상 r_t 가 결정된다. 보상과 추정된 가치함수로 모델을 업데이트하고 배터리 충방전 후 값 및 다음 시간대에 관측된 에너지 발전량 x_t와 에너지 가격 λ_t로 다음 상태를 결정한다.

[0048] 도 4는 본 발명의 실시 예들에 기초하여 수행된 입찰 모의 실험 결과를 설명하기 위한 도면이다.

[0049] 도 4를 참조하면, 본 발명의 실시 예들로 입찰 모의 실험을 했을 때 재생에너지(RES; Renewable Energy Sources) 발전량, 에너지 가격, 배터리에 저장된 에너지의 변화를 나타낸다. 재생에너지는 발전량의 불확실성을 갖는 태양광과 풍력을 대상으로 모의 실험하였다. 전체적으로 에너지 가격이 높고 배터리에 저장된 에너지가 많을 때 과대 입찰을 하고, 에너지 가격이 낮고 배터리에 저장된 에너지가 적을 때 과소 입찰을 하려는 경향을 보임을 알 수 있다.

[0050] 도 5는 본 발명의 실시 예들에 따른 에너지 입찰 방법을 구현하기 위한 컴퓨팅 장치를 설명하기 위한 블록도이다.

[0051] 도 5를 참조하면, 본 발명의 실시 예들에 따른 에너지 입찰 방법은 컴퓨팅 장치(500)를 이용하여 구현될 수 있다.

[0052] 컴퓨팅 장치(500)는 버스(520)를 통해 통신하는 프로세서(510), 메모리(530), 사용자 인터페이스 입력 장치(540), 사용자 인터페이스 출력 장치(550) 및 저장 장치(560) 중 적어도 하나를 포함할 수 있다. 컴퓨팅 장치(500)는 또한 네트워크(40), 예컨대 무선 네트워크에 전기적으로 접속되는 네트워크 인터페이스(570)를 포함할 수 있다. 네트워크 인터페이스(570)는 네트워크(40)를 통해 다른 개체와 신호를 송신 또는 수신할 수 있다.

[0053] 프로세서(510)는 AP(Application Processor), CPU(Central Processing Unit), GPU(Graphic Processing Unit) 등과 같은 다양한 종류들로 구현될 수 있으며, 메모리(530) 또는 저장 장치(560)에 저장된 명령을 실행하는 임의의 반도체 장치일 수 있다. 프로세서(510)는 도 1 내지 도 4에서 설명한 기능 및 방법들을 구현하도록 구성될 수 있다.

[0054] 메모리(530) 및 저장 장치(560)는 다양한 형태의 휘발성 또는 비 휘발성 저장 매체를 포함할 수 있다. 예를 들어, 메모리는 ROM(read-only memory)(531) 및 RAM(random access memory)(532)를 포함할 수 있다. 본 발명의 실시 예에서 메모리(530)는 프로세서(510)의 내부 또는 외부에 위치할 수 있고, 메모리(530)는 이미 알려진 다양한 수단을 통해 프로세서(510)와 연결될 수 있다.

[0055] 또한, 본 발명의 실시 예들에 따른 에너지 입찰 방법 중 적어도 일부는 컴퓨팅 장치(500)에서 실행되는 프로그

램 또는 소프트웨어로 구현될 수 있고, 프로그램 또는 소프트웨어는 컴퓨터로 판독 가능한 매체에 저장될 수 있다.

[0056] 또한, 본 발명의 실시 예들에 따른 에너지 입찰 방법 중 적어도 일부는 컴퓨팅 장치(500)와 전기적으로 접속될 수 있는 하드웨어로 구현될 수도 있다.

[0057] 이제까지 설명한 본 발명의 실시 예들에 따르면, 오차 패널티 감축을 위한 입찰값과 차익거래로 이익 증대를 위한 충방전량을 동시에 결정하는 입찰 전략을 제공하며, 이는 기존에 차익거래 또는 입찰 오차 보정 중 하나에만 초점이 맞춰진 방식과 달리 두 방식의 장점을 모두 활용할 수 있어 궁극적으로 최종 입찰 수익을 극대화할 수 있다.

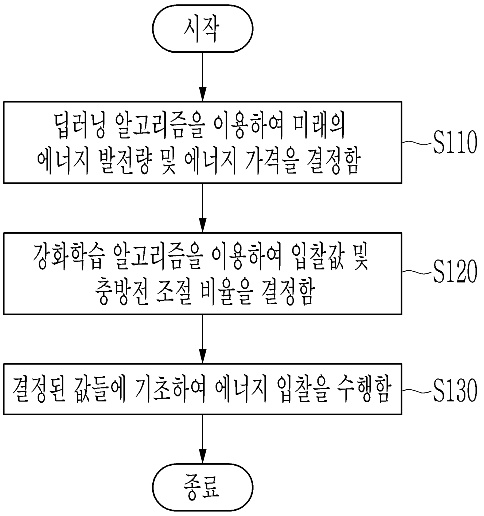
[0058] 또한, 발전량 및 에너지 가격 예측을 위한 딥러닝 구조에 입찰 수익 극대화를 위한 입찰값 및 충방전량을 결정하는 강화학습 알고리즘을 적용한 심층강화학습을 활용하여 발전량 및 에너지 가격 불확실성이 있는 환경(재생에너지, 천연가스 등)에서도 입찰 수익 극대화를 실현할 수 있다.

[0059] 이상에서 본 발명의 실시 예에 대하여 상세하게 설명하였지만 본 발명의 권리 범위는 이에 한정되는 것은 아니고, 다음의 청구범위에서 정의하고 있는 본 발명의 기본 개념을 이용한, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자의 여러 변형 및 개량 형태 또한 본 발명의 권리 범위에 속한다.

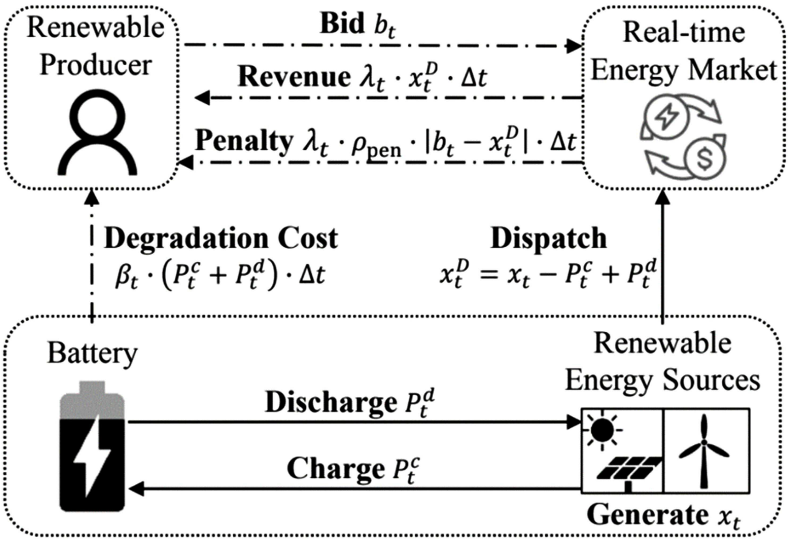
[0060] 이상에서 본 발명의 실시 예에 대하여 상세하게 설명하였지만 본 발명의 권리 범위는 이에 한정되는 것은 아니고, 다음의 청구범위에서 정의하고 있는 본 발명의 기본 개념을 이용한, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자의 여러 변형 및 개량 형태 또한 본 발명의 권리 범위에 속한다.

도면

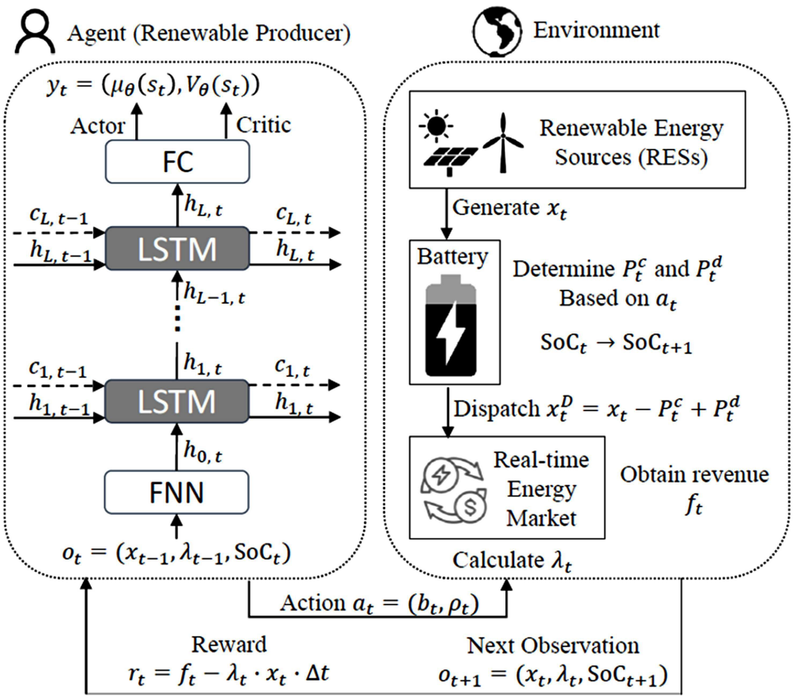
도면1



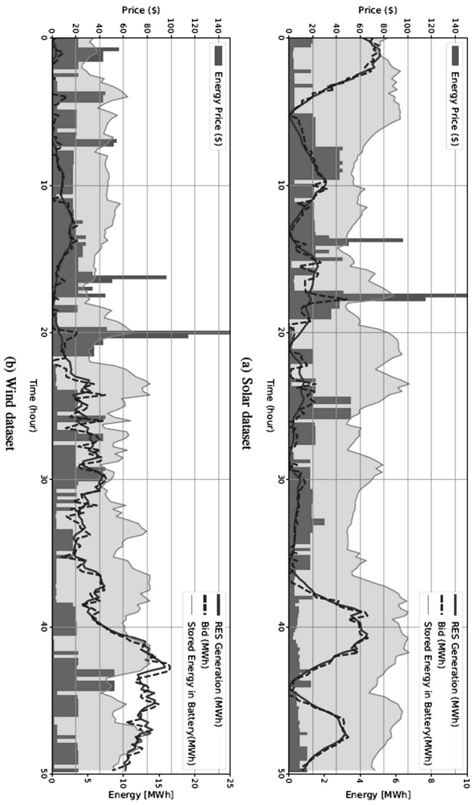
도면2


Renewable Producer
Bid b_t
Real-time Energy Market
Revenue $\lambda_t \cdot x_t^D \cdot \Delta t$
Penalty $\lambda_t \cdot \rho_{\text{pen}} \cdot |b_t - x_t^D| \cdot \Delta t$
Degradation Cost
$\beta_t \cdot (P_t^c + P_t^d) \cdot \Delta t$
Dispatch
$x_t^D = x_t - P_t^c + P_t^d$
Battery
Renewable Energy Sources
Discharge P_t^d
Charge P_t^c
Generate x_t

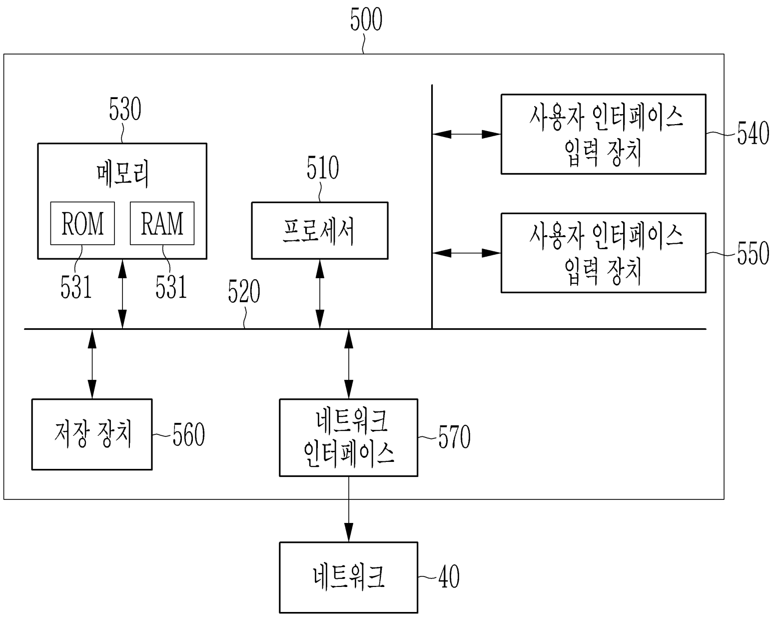
도면3


Agent (Renewable Producer)
Environment
$y_t = (\mu_\theta(s_t), V_\theta(s_t))$
Actor
Critic
FC
$h_{L,t}$
$c_{L,t-1}$
$h_{L,t-1}$
LSTM
$c_{L,t}$
$h_{L,t}$
$h_{L-1,t}$
$h_{1,t}$
$c_{1,t-1}$
$h_{1,t-1}$
LSTM
$c_{1,t}$
$h_{1,t}$
$h_{0,t}$
FNN
$o_t = (x_{t-1}, \lambda_{t-1}, \text{SoC}_t)$
Renewable Energy Sources (RESs)
Generate x_t
Battery
Determine P_t^c and P_t^d Based on a_t
$\text{SoC}_t \rightarrow \text{SoC}_{t+1}$
Dispatch $x_t^D = x_t - P_t^c + P_t^d$
Real-time Energy Market
Obtain revenue f_t
Calculate λ_t
Action $a_t = (b_t, \rho_t)$
Reward
$r_t = f_t - \lambda_t \cdot x_t \cdot \Delta t$
Next Observation
$o_{t+1} = (x_t, \lambda_t, \text{SoC}_{t+1})$

도면4


Price ($)
Energy Price ($)
RES Generation (MWh)
Bid (MWh)
Stored Energy in Battery(MWh)
Time (hour)
Energy [MWh]
(a) Solar dataset
(b) Wind dataset

도면5


500
530
메모리
ROM
RAM
531
531
510
프로세서
520
사용자 인터페이스
입력 장치
540
사용자 인터페이스
입력 장치
550
저장 장치
560
네트워크
인터페이스
570
네트워크
40