

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)(11) 공개번호 10-2025-0101169
(43) 공개일자 2025년07월04일

(51) 국제특허분류(Int. Cl.)
H02J 3/38 (2022.01) *G06N 20/00* (2019.01)
G06N 3/08 (2023.01) *G06Q 10/04* (2023.01)
G06Q 50/06 (2024.01) *H02J 3/32* (2006.01)

(52) CPC특허분류
H02J 3/381 (2023.08)
G06N 20/00 (2021.08)

(21) 출원번호 10-2023-0192534
 (22) 출원일자 2023년12월27일
 심사청구일자 2023년12월27일

(71) 출원인
 고려대학교 산학협력단
 서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)

(72) 발명자
 신경철
 서울특별시 성북구 고려대로 7가길 57, 102호
 주성관
 서울특별시 노원구 노원로18길 19, 211동 404호
 (뒷면에 계속)

(74) 대리인
 김등용

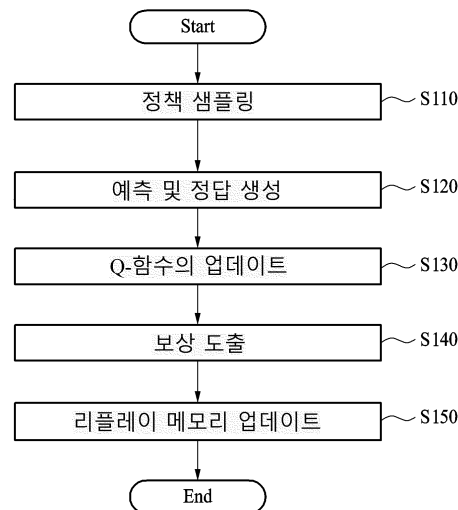
전체 청구항 수 : 총 8 항

(54) 발명의 명칭 강화학습을 이용한 전력계통 운용을 위한 ESS와 P2X 적정용량 산정 장치 및 방법

(57) 요약

강화학습을 이용한 전력계통 운용을 위한 ESS(Energy Storage System)와 P2X(Power To X) 적정용량 산정 장치 및 방법이 개시된다. 상기 적정용량 산정 방법은 프로세서(Processor)를 포함하는 컴퓨팅 장치에 의해 수행되고, 리플레이 메모리(Replay memory)로부터 정책을 샘플링하는 단계, 샘플링된 정책의 행동(Action)에 대한 Q-네트워크의 예측을 생성하고, 상기 샘플링된 정책의 행동(Action)에 대한 타겟 네트워크의 출력인 정답을 생성하는 단계, 미리 정의된 오류를 최소화하는 방향으로 상기 Q-네트워크의 가중치를 업데이트하는 단계, 상기 Q-네트워크를 통해 예측된 행동(Action)에 대한 보상을 도출하는 단계, 및 상기 리플레이 메모리를 업데이트하는 단계를 포함한다.

대표도 - 도1



(52) CPC특허분류

G06N 3/08 (2023.01)
 G06Q 10/04 (2023.01)
 G06Q 50/06 (2024.01)
 H02J 3/32 (2023.08)
 H02J 2203/20 (2023.08)
 H02J 2300/20 (2023.08)
 Y04S 10/123 (2020.08)
 Y04S 10/14 (2020.08)
 Y04S 40/20 (2020.08)

(72) 발명자

이득영

경기도 의왕시 내손순환로 7, 310동 503호

정현철

서울특별시 성북구 안암로 93, 601호(동안빌딩)

이 발명을 지원한 국가연구개발사업

과제고유번호	1415186969
과제번호	20204010600220
부처명	산업통상자원부
과제관리(전문)기관명	한국에너지기술평가원
연구사업명	에너지인력양성사업(예특)
연구과제명	에너지거래 플랫폼 융합인력양성
기 여 율	1/1
과제수행기관명	고려대학교산학협력단
연구기간	2023.01.01 ~ 2023.12.31

명세서

청구범위

청구항 1

프로세서(Processor)를 포함하는 컴퓨팅 장치에 의해 수행되는, 강화학습을 이용한 전력계통 운용을 위한 ESS(Energy Storage System)와 P2X(Power To X) 적정용량 산정 방법에 있어서,

리플레이 메모리(Replay memory)로부터 정책을 샘플링하는 단계;

샘플링된 정책의 행동(Action)에 대한 Q-네트워크의 예측을 생성하고, 상기 샘플링된 정책의 행동(Action)에 대한 타겟 네트워크의 출력인 정답을 생성하는 단계;

미리 정의된 오류를 최소화하는 방향으로 상기 Q-네트워크의 가중치를 업데이트하는 단계;

상기 Q-네트워크를 통해 예측된 행동(Action)에 대한 보상을 도출하는 단계; 및

상기 리플레이 메모리를 업데이트하는 단계를 포함하는,

ESS와 P2X 적정용량 산정 방법.

청구항 2

제1항에 있어서,

상기 리플레이 메모리는 현재 상태, 행동, 보상, 및 다음 상태를 포함하는 정책이 저장되어 있는,

ESS와 P2X 적정용량 산정 방법

청구항 3

제2항에 있어서,

상기 행동은 ESS와 P2X의 충·방전 용량인,

ESS와 P2X 적정용량 산정 방법

청구항 4

제3항에 있어서,

상기 리플레이 메모리를 업데이트하는 단계는,

상기 샘플링된 정책에 포함된 현재 상태, 행동, 행동에 대한 보상, 및 행동으로 인한 다음 상태를 상기 리플레이 메모리에 저장하는,

ESS와 P2X 적정용량 산정 방법

청구항 5

제4항에 있어서,

상기 보상을 도출하는 단계는,

상기 샘플링된 정책에 포함된 현재 상태와 행동을 기초로, 전력계통의 전력수요에 대한 평가 결과를 수치화하는,

ESS와 P2X 적정용량 산정 방법

청구항 6

제4항에 있어서,

상기 보상을 도출하는 단계는,

상기 샘플링된 정책에 포함된 현재 상태와 행동을 기초로, 전력계통의 전압 안정도에 대한 평가 결과를 수치화 하는,

ESS와 P2X 적정용량 산정 방법

청구항 7

제4항에 있어서,

상기 보상을 도출하는 단계는,

상기 샘플링된 정책에 포함된 현재 상태와 행동을 기초로, 상기 ESS 및 상기 P2X의 설치 비용 및 운영 비용에 대한 평가 결과를 수치화하는,

ESS와 P2X 적정용량 산정 방법

청구항 8

제4항에 있어서,

상기 샘플링하는 단계 내지 상기 리플레이 메모리를 업데이트하는 단계는 미리 정해진 주기 동안 반복적으로 수행되고,

상기 미리 정해진 주기가 경과한 경우, 상기 타겟 네트워크의 가중치를 상기 Q-네트워크의 가중치로 업데이트 한 후에 상기 샘플링하는 단계 내지 상기 리플레이 메모리를 업데이트하는 단계를 다음 주기 동안 반복적으로 수행하는,

ESS와 P2X 적정용량 산정 방법

발명의 설명

기술 분야

[0001] 본 발명은 전력계통에 관한 것으로, 특히 강화학습을 이용해 전력계통을 운용하기 위한 ESS와 P2X의 적정용량을 산정하는 장치 및 방법에 관한 것이다.

배경 기술

[0002] 강화학습(Reinforcement Learning)이란 기계 학습의 한 영역으로, 어떤 환경 안에서 정의된 에이전트가 현재의 상태를 인식하여 선택 가능한 행동들 중 보상을 최대화하는 행동 혹은 행동 순서를 선택하는 방법이다. 최근에는 심층 학습(Deep Learning)을 활용한 DQN(Deep Q-Network)과 A2C(Advantage Actor Critic)와 같은 기술이 대표적이다.

[0003] ESS(Energy Storage System)란 에너지저장장치로, 저장에 어렵고 사용 후 없어져 버리는 전기에너지를 효율적으로 사용할 수 있도록 충·방전이 가능하다. 최근 신재생에너지와 연계하여 변동성이 심한 신재생에너지를 기반으로도 안정적으로 전력계통을 운용할 수 있다. ESS의 종류로는 초단주기·단주기·장주기 저장장치로 구분된다. 초단주기·단주기용 ESS는 대표적으로 리튬이온 배터리가 있으며, 빠른 응답속도로 전력계통의 주파수를 조정할 수 있다. 장주기용 ESS는 열에너지를 이용한 카르노배터리 등이 있으며, 계통·수용가의 피크를 저감하여 전기요금의 절약 및 수요반응자원으로 활용이 가능하다.

[0004] P2X(Power To X, 전기에너지를 기타 합성연료 형태로 저장하는 방식)란 잉여 전기에너지를 다른 형태의 연료·에너지 등으로 저장하는 기술이다. 수소·메탄 등과 같은 기체 연료, 메탄올·경유 등과 같은 액체 연료, 또는 열에너지로의 전환 등이 존재한다. 대표적으로 P2G(Power To Gas)는 태양광이나 풍력으로 얻은 전기를 이용하여 수소 또는 메탄 등의 연료를 생성하고 저장하며 이산화탄소를 천연가스로 전환한다는 측면에서 친환경적인 기술이다. P2X는 ESS와는 다르게 변환된 연료·에너지 등을 장기간 저장할 수 있어 계절용으로 사용이 가능하다.

[0005] 제10차 전력수급기본계획에 따르면, 미래의 전력계통은 신재생에너지의 확정설비 용량이 2023년 32.8GW에서 2036년 108.3GW로 확충된다. 날씨와 일사량의 영향을 통해 변동성이 발생하는 신재생에너지의 증가는 전력계통의 안정성에 큰 저해를 끼친다. 이러한 상황에서 전력을 저장하고 필요 시 방전을 할 수 있는 장치인 ESS와 P2X

기술이 요구되며 전력계통을 운용하기 위해서는 ESS와 P2X의 적정용량 산출이 필요하다.

[0006] 우크라이나와 러시아 간 전쟁 등으로 인해 천연가스 공급이 어려워져 천연가스 가격은 약 5배 이상에 달하기도 하였다. 국내는 해외의 천연가스를 사용하여 발전기를 발전시켜 전력 계통을 운용하는 데, 이로 인하여 전기요금의 인상과 막대한 부채를 끼안아야 했다. 이러한 상황에서 신재생에너지와 ESS·P2X를 활용하여 외국에 의존하지 않고 자체적으로 전력을 생산하고 운용하며 가격 안정화를 위하여 저비용의 ESS·자원의 설치가 필요하다.

[0007] 이상 기후로 인해 세계적으로 탄소 중립을 중요시하는 경향이 나타나고 있다. 국내에서는 국가 온실가스 배출량의 많은 부분을 차지하고 있는 에너지 부문은 2050년 탄소중립 달성을 위하여 전력으로의 에너지 전환이 증가하고 있다. 대표적으로는 전기차 보급량 증가와 히트펌프 등의 화석연료를 대체할 것으로 보이며 매해 전력수요의 증가가 예상된다. 이러한 상황에서 탄소 중립을 완전히 달성하기 위해서는 화석연료 기반의 석탄·가스 발전기의 사용을 줄여야 할 필요가 있다. 따라서 친환경적 에너지인 ESS·의 활용이 필수적이다.

선행기술문헌

특허문헌

[0008] (특허문헌 0001) 대한민국 공개특허 제2022-0101875호 (2022.07.19. 공개)

(특허문헌 0002) 대한민국 공개특허 제2021-0001172호 (2021.01.06. 공개)

(특허문헌 0003) 대한민국 등록특허 제2190845호 (2020.12.14. 공고)

(특허문헌 0004) 대한민국 등록특허 제2497271호 (2023.02.08. 공고)

발명의 내용

해결하려는 과제

[0009] 본 발명이 이루고자 하는 기술적인 과제는 강화학습을 이용한 전력계통 운용을 위한 ESS와 P2X 적정용량 산정 장치 및 방법을 제공하는 것이다.

과제의 해결 수단

[0010] 본 발명의 일 실시예에 따른 ESS(Energy Storage System)와 P2X(Power To X) 적정용량 산정 방법은 프로세서(Processor)를 포함하는 컴퓨팅 장치에 의해 수행되는 강화학습을 이용한 전력계통 운용을 위한 ESS와 P2X 적정용량 산정 방법으로, 리플레이 메모리(Replay memory)로부터 정책을 샘플링하는 단계, 샘플링된 정책의 행동(Action)에 대한 Q-네트워크의 예측을 생성하고, 상기 샘플링된 정책의 행동(Action)에 대한 타겟 네트워크의 출력인 정답을 생성하는 단계, 미리 정의된 오류를 최소화하는 방향으로 상기 Q-네트워크의 가중치를 업데이트하는 단계, 상기 Q-네트워크를 통해 예측된 행동(Action)에 대한 보상을 도출하는 단계, 및 상기 리플레이 메모리를 업데이트하는 단계를 포함한다.

발명의 효과

[0011] 본 발명의 실시예에 따른, 강화학습을 이용한 전력계통 운용을 위한 ESS와 P2X 적정용량 산정 장치 및 방법에 의할 경우, 복잡한 전력계통에서도 ESS와 P2X의 적정용량을 산출하여 안정적이고 경제적으로 계통을 운용할 수 있는 효과가 있다.

도면의 간단한 설명

[0012] 본 발명의 상세한 설명에서 인용되는 도면을 보다 충분히 이해하기 위하여 각 도면의 상세한 설명이 제공된다.

도 1은 본 발명의 일 실시예에 따른, 강화학습을 이용한 전력계통 운용을 위한 ESS와 P2X 적정용량 산정 방법을 설명하기 위한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0013] 본 명세서에 개시되어 있는 본 발명의 개념에 따른 실시예들에 대해서 특정한 구조적 또는 기능적 설명들은 단

지 본 발명의 개념에 따른 실시예들을 설명하기 위한 목적으로 예시된 것으로서, 본 발명의 개념에 따른 실시예들은 다양한 형태들로 실시될 수 있으며 본 명세서에 설명된 실시예들에 한정되지 않는다.

[0014] 본 발명의 개념에 따른 실시예들은 다양한 변경들을 가할 수 있고 여러 가지 형태들을 가질 수 있으므로 실시예들을 도면에 예시하고 본 명세서에서 상세하게 설명하고자 한다. 그러나, 이는 본 발명의 개념에 따른 실시예들을 특정한 개시 형태들에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물, 또는 대체물을 포함한다.

[0015] 제1 또는 제2 등의 용어는 다양한 구성 요소들을 설명하는데 사용될 수 있지만, 상기 구성 요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성 요소를 다른 구성 요소로부터 구별하는 목적으로만, 예컨대 본 발명의 개념에 따른 권리 범위로부터 벗어나지 않은 채, 제1 구성 요소는 제2 구성 요소로 명명될 수 있고 유사하게 제2 구성 요소는 제1 구성 요소로도 명명될 수 있다.

[0016] 어떤 구성 요소가 다른 구성 요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성 요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성 요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성 요소가 다른 구성 요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는 중간에 다른 구성 요소가 존재하지 않는 것으로 이해되어야 할 것이다. 구성 요소들 간의 관계를 설명하는 다른 표현들, 즉 "~사이에"와 "바로 ~사이에" 또는 "~에 이웃하는"과 "~에 직접 이웃하는" 등도 마찬가지로 해석되어야 한다.

[0017] 본 명세서에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로서, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 본 명세서에 기재된 특징, 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0018] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 갖는 것으로 해석되어야 하며, 본 명세서에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.

[0019] 이하, 본 명세서에 첨부된 도면들을 참조하여 본 발명의 실시예들을 상세히 설명한다. 그러나, 특허출원의 범위가 이러한 실시예들에 의해 제한되거나 한정되는 것은 아니다. 각 도면에 제시된 동일한 참조 부호는 동일한 부재를 나타낸다.

[0020] 변동성이 심한 신재생에너지, 예컨대, 풍력 발전, 태양광 발전, 조력 발전 등이 추가적으로 보급되고, 송전선의 건설이 지연되는 등 전력계통에는 다양한 상황이 발생하고 있다. 이러한 상황에서 강화학습을 통해 복잡한 전력계통에서도 ESS(Energy Storage System, 에너지 저장 장치)와 P2X(Power To X)의 적정용량을 산출하여 안정적이고 경제적으로 계통을 운용할 수 있다.

[0021] 본 발명에서는 강화학습을 위해, DQN(Deep Q-Network)을 활용한다. DQN은 리플레이 메모리(Replay Memory)를 통하여 학습을 진행하며, 학습의 계산은 심층 학습(Deep-Learning)으로 진행될 수 있다.

[0022] 리플레이 메모리는 에이전트가 환경에서 탐험하여 얻은 샘플(S, A, R, S')을 메모리에 저장하는 방법이다. 즉, 에이전트는 학습할 때 리플레이 메모리에서 샘플을 무작위로 추출하여 그 샘플에 대해 인공신경망(Q-네트워크를 의미할 수 있음)을 업데이트한다. 그리고 이 과정을 매 타임스텝마다 반복한다. 여기서, 샘플은 상태(S), 액션(A), 보상(R), 및 다음 상태(S')를 포함할 수 있다.

[0023] 리플레이 메모리는 미리 정해진 크기를 갖기 때문에, 리플레이 메모리가 가득 차게 되면 가장 먼저 들어온 샘플부터 삭제된다. 이에 따라, 에이전트가 학습을 통해 점점 더 높은 점수를 받게 되면 더 좋은 샘플들이 저장될 것이다.

[0024] 심층 학습은 모델의 성능을 위하여, 활성화 함수(Activation Function), 손실 함수(Loss Function), 옵티마이저(Optimizer), 렙스(Depth), 입력(Input), 출력(Output) 등을 설정한다. DQN에서는 정책(Policy)을 평가하기 위해 정책이 저장되어 있는 리플레이 메모리에서 데이터를 임의로 추출해서 학습에 사용하고, 이때 큐함수(Q-

Function) 또는 Q-네트워크의 업데이트를 통해 모델 학습을 진행한다. 해당 과정에 대한 수학적 1은 다음과 같다.

[0025] [수학적 1]

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha(R_{t+1} + \gamma \max_{a'} Q(S_{t+1}, a') - Q(S_t, A_t))$$

[0027] 수학적 1에서, Q는 행동가치함수를, S_t 는 t 시간에서의 상태(State)를, S_{t+1} 은 t+1 시간에서의 상태를, A_t 는 t 시간에서의 행동(Action)을, α 는 스텝 사이즈(step size)를, R_{t+1} 은 t+1 시간에서의 보상(Reward)을, γ 는 할인율을 나타낸다.

[0028] 그리고, 오류함수는 MSE(Mean Squared Error, 평균제곱오차)를 활용할 수 있다. 수학적 2로 표현된 오류함수를 최소화하는 방향으로 DQN이 업데이트되며 학습을 진행한다.

[0029] [수학적 2]

$$MSE = (\text{정답} - \text{예측})^2$$

[0031] 리플레이 메모리를 사용하는 에이전트는 매 타임스텝마다 리플레이 메모리에서 샘플을 미니배치만큼 추출하여 학습에 사용한다. 학습을 수행하는 것은 인공신경망, 즉 Q-네트워크의 가중치 및/또는 편향을 업데이트 하는 것을 의미하며, 이는 오류함수의 오류(또는 오차)를 감소시키는 방향으로 업데이트된다. 여기서, 문제는 업데이트의 목표가 되는 정답이 계속 변한다는 것이다.

$$\text{Error} = (R_{t+1} + \gamma \max_{a'} Q(s', a', \Theta) - Q(s, a, \Theta))^2$$

[0033] 상기 수식에서 $R_{t+1} + \gamma \max_{a'} Q(s', a', \Theta)$ 은 정답 값인데, 인공신경망이 계속 업데이트 되면서 가중치와 편향값이 바뀐다면 정답값이 영향을 주어 학습이 잘 이루어지지 않게 된다. 이를 위해, Q-네트워크와 동일한 구조를 갖는 타겟 네트워크가 활용된다. 구체적으로, 기존의 인공신경망(Q-네트워크)을 계속 학습시켜 나가는 동안에, 타겟 네트워크의 파라미터는 일정 시간 동안 바꾸지 않고 유지하면서, 일정 스텝 주기로 기존 인공신경망의 가중치와 편향값을 타겟 네트워크에 업데이트 시켜주는 것이다.

[0034] 오류함수에서 이를 구분하기 위해 타겟 네트워크는 Θ^- 를 매개변수로 표현하고 Q-네트워크는 Θ 를 매개변수로 표현한다. 이와 같이, 타겟 네트워크를 따로 두고 학습을 진행하면 정답 값이 계속 바뀌지 않기 때문에 오류를 줄여가며 학습을 효과적으로 수행할 수 있다.

$$\text{Error} = (R_{t+1} + \gamma \max_{a'} Q(s_{t+1}, a', \Theta^-) - Q(S_t, A_t, \Theta))^2$$

[0036] 다시 말하면, 동일한 구조를 갖는 Q-네트워크와 타겟 네트워크를 활용하여, Q-네트워크를 이용하여 예측값을 생성하고, 타겟 네트워크를 이용하여 정답을 생성할 수 있다.

[0037] **환경(Environment)**

[0038] 환경은 강화학습을 사용하여 풀고자 하는 대상이나 문제를 의미한다. 안정적인 전력계통 운용을 위한 ESS와 P2X의 적정용량을 산정하기 위한 환경을 구축하여야 한다. 환경에서는 ESS와 P2X의 적정용량을 산정하기 위한 전력계통의 기본적인 상태가 필요하다. 예컨대, 발전기, 변압기, 선로, ESS 설비, P2X 설비 등과 같은 전기 설비들을 모델링하여 환경에 추가하여야 한다. 또한, 필요한 부하와 신재생 발전기에서 발전되는 전력량 등을 설정하여 ESS와 P2X에 충·방전에 필요한 양을 산출하여야 한다. 또한, ESS와 P2X의 충·방전량 및 횟수에도 제한을 걸어 전력계통 운용에 필요하게 환경을 모델링할 수 있다. 제약으로 표현이 가능한 수학적식은 아래 수학적 3 및 수학적 4와 같다.

[0039] [수학적 3]

$$\sum_{t=1}^{24} \max[u(t) - u(t-1), 0] \leq 2$$

[0040] [수학적 4]

- [0042] $u(t) \in \{0, 1\}$
- [0043] 상기 수학적식에서 $u(t)$ 는 ESS와 P2X의 온(ON)/오프(OFF) 여부를 나타내는 함수로, 예컨대 0의 값을 가지는 경우 온(또는 오프) 상태이고 1의 값을 가지는 경우 오프(또는 온) 상태일 수 있다.
- [0044] **행동(Action) 및 정책(Policy)**
- [0045] 행동에는 전력계통을 운용하기 위하여 대표적으로는 부하와 발전량을 일치시키기 위한 ESS와 P2X의 충·방전량으로 설정할 수 있다. 필요한 충·방전량을 통해 ESS 및 P2X의 용량 산출이 가능하다. 그리고, 강화학습의 성능을 향상시키기 위하여 정책, 즉 행동에 따른 Q-벨류(Q-Value, 행동가치)를 리플레이 메모리에 저장할 수 있다. 학습을 진행할 경우에는 리플레이 메모리에 저장되어 있는 데이터를 무작위로 추출하여 학습을 진행할 수 있다.
- [0046] **학습 과정**
- [0047] 보상은 행동 및 정책을 통해 심층 학습으로 산정된 ESS와 P2X의 충·방전 용량을 통하여 전력 계통 평가를 진행한다. 전력계통에 대한 평가는 설계자의 판단에 따라 우선순위를 설정할 수 있다. 대표적으로, 전력 수요(MW)와 전압(V)의 안정도 등에 대한 평가가 존재한다. 전력 수요에 대한 평가는 전력계통에 존재하는 발전량과 ESS·P2X의 방전량의 총 합과 부하를 사용하는 수요량과 ESS·P2X의 충전량의 총 합이 같을 때를 보상값으로 모델링하여 전력계통에 대한 수요 관리를 최적으로 이행시켜 주는 ESS·P2X 충·방전량을 산정할 수 있다. 또한, 전압의 안정도에 대해서는 안정적 계통 전압을 유지하기 위한 모델링을 통하여 그 결과에 따른 보상을 통해 강화학습에 대한 최적화를 진행할 수 있게 된다. 그리고, ESS·P2X의 설치 비용 및 운용 비용에 관한 환경 데이터를 활용하여 최저비용의 설비에 대한 보상을 모델링하여 다양한 ESS·P2X 설비 중 가장 경제적인 설비의 충·방전량에 대한 산출도 진행할 수 있게 된다.
- [0048] 페널티(Penalty)는 강화학습을 진행하는 중 보상이 상승하는 행동도 있지만, 임의로 행동을 학습하기 때문에 보상이 하락하는 경우도 존재한다. 해당 경우에는 전력계통 안정화가 강화학습의 목표이기 때문에 전력계통의 붕괴가 발생하는 상황, 즉 큰 오차를 발생하며 전력 수요를 일치시키지 못할 경우나 전압을 유지하지 못할 경우, 설치 및 운용 비용이 예산을 초과하는 경우에는 보상에서 큰 감점인 페널티를 부여하게 된다.
- [0049] 도 1은 본 발명의 일 실시예에 따른, 강화학습을 이용한 전력계통 운용을 ESS와 P2X 적정용량 산정 방법을 설명하기 위한 흐름도이다.
- [0050] 도 1을 참조하면, ESS와 P2X 적정용량 산정 모델 생성 방법 등으로 명명될 수도 있는 ESS와 P2X 적정용량 산정 방법은 적어도 프로세서(Processor) 및/또는 메모리(Memory)를 포함하는 컴퓨팅 장치에 의해 수행될 수 있다. 다시 말하면, ESS와 P2X 적정용량 산정 방법을 이루는 단계들 중 적어도 일부는 컴퓨팅 장치에 포함되는 프로세서의 동작으로 이해될 수도 있고, 이 경우 컴퓨팅 장치는 ESS와 P2X 적정용량 산정 장치로 명명될 수도 있다. 컴퓨팅 장치는 PC(Personal Computer), 서버(server), 랩탑 컴퓨터, 태블릿 PC 등을 포함할 수 있다. 실시예에 따라, 컴퓨팅 장치는 물리적으로 구분된 복수의 장치에 의해 구현되는 것도 가능하다. 이하에서, ESS와 P2X 적정용량 산정 방법을 설명함에 있어, 앞선 기재와 중복되는 내용에 관하여는 그 구체적인 기재를 생략하기로 한다.
- [0051] 우선, DQN의 리플레이 메모리(Replay Memory)에 저장되어 있는 정책을 샘플링하여 적어도 하나의 정책을 추출한다(S110). 리플레이 메모리에는 다수의 정책들이 저장되어 있을 수 있으며, 학습을 진행하는 동안 새로운 정책이 생성되어 저장될 수 있다. 이때, 리플레이 메모리의 미리 정해진 저장 공간이 모두 점유되면, 가장 먼저 저장된 정책이 리플레이 메모리에서 삭제될 수 있다. 또한, 샘플링은 임의의 방식으로 수행될 수 있다.
- [0052] DQN의 Q-네트워크와 타겟 네트워크를 이용하여 예측 및 정답을 생성한다(S120). 이때, Q-네트워크와 타겟 네트워크는 동일한 구조를 갖는 인공 신경망(Artificial Neural Network, ANN) 모델로써, 상태(state)를 입력으로 수신하여 행동(action)을 출력할 수 있다.
- [0053] Q-함수에 대한 업데이트가 수행된다(S130). 다시 말하면, 미리 정의된 오류를 최소화하는 방향으로 Q-네트워크의 가중치 등에 대한 업데이트가 수행될 수 있다. 구체적으로, 타겟 네트워크의 가중치 등은 Q-네트워크와 동일하게 설정된 후, 미리 정해진 주기 동안 타겟 네트워크의 출력을 정답으로 이용하여 Q-네트워크에 대한 업데이트가 수행된다. 미리 정해진 주기가 경과하면, Q-네트워크의 업데이트 결과가 타겟 네트워크에 반영되고, 다시 Q-네트워크에 대한 업데이트가 수행될 수 있다.
- [0054] 도출된 행동(Q-네트워크의 출력을 의미할 수 있음)에 대한 보상이 도출된다(S140). 보상은 도출된 행동(action)에 따른 1) 전력수요 평가 결과, 2) 전압 안정도, 및 3) ESS/P2X의 설치 비용 및 운영 비용 중 적어도 하나의

정도를 수치화한 값을 의미할 수 있다. 실시예에 따라, 복수개의 평가 결과에 가중치를 부여하여 보상을 산출할 수도 있다.

- [0055] 예컨대, 전력수요에 대한 평가 결과로 보상을 산출하는 경우, 전력계통에 존재하는 발전량과 ESS/P2X의 방전량의 총 합과 부하를 사용하는 수요량과 ESS/P2X의 충전량의 총 합의 차이의 정도를 수치화하여 보상을 산출할 수 있다. 이를 위해, 각 행동에 대한 시뮬레이션이 수행될 수 있고, 시뮬레이션 결과를 기초로 보상이 산출될 수 있다. 따라서, 컴퓨팅 장치에는 전력계통을 모의할 수 있는 시뮬레이션 프로그램(또는 시뮬레이션 수단)이 포함될 수 있다.
- [0056] 다른 예로, 전압 안정도에 대한 평가 결과로 보상을 산출하는 경우, 전압 유지의 안정성의 정도를 수치화하여 보상을 산출할 수 있다.
- [0057] 또 다른 예로, ESS/P2X의 설치 비용 및 운영 비용에 대한 평가 결과로 보상을 산출하는 경우, 소요되는 비용의 정도를 수치화하여 보상을 산출할 수 있다. 이 경우, 필요한 비용이 적을수록 큰 보상이 산출될 수 있다.
- [0058] 다음으로, 리플레이 메모리가 업데이트된다(S150). 즉, 상태, 행동, 다음 상태, 및 보상이 새로이 리플레이 메모리에 저장될 수 있다.
- [0059] 상술한 S110 단계 내지 S150 단계는 미리 정해진 주기 동안 반복적으로 수행될 수 있다. 미리 정해진 주기가 경과되는 경우, Q-네트워크의 가중치는 타겟 네트워크에 반영되고, 다음 주기 동안 다시 S110 단계 내지 S150 단계가 반복적으로, 다음 주기 동안, 수행될 수 있다.
- [0060] 상술한 과정을 통해, ESS와 P2X 적정용량 산정 모델이 생성될 수 있다. 실시예에 따르면, 추론된 총·방전량 중 가장 큰 값이 적정용량으로 산정될 수 있다.
- [0061] 이상에서 설명된 장치는 하드웨어 구성 요소, 소프트웨어 구성 요소, 및/또는 하드웨어 구성 요소 및 소프트웨어 구성 요소의 집합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성 요소는, 예를 들어, 프로세서, 컨트롤러, ALU(Arithmetic Logic Unit), 디지털 신호 프로세서(Digital Signal Processor), 마이크로컴퓨터, FPA(Field Programmable array), PLU(Programmable Logic Unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(Operation System, OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술 분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(Processing Element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 컨트롤러를 포함할 수 있다. 또한, 병렬 프로세서(Parallel Processor)와 같은, 다른 처리 구성(Processing Configuration)도 가능하다.
- [0062] 소프트웨어는 컴퓨터 프로그램(Computer Program), 코드(Code), 명령(Instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(Collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성 요소(Component), 물리적 장치, 가상 장치(Virtual Equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호 파(Signal Wave)에 영구적으로, 또는 일시적으로 구체화(Embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.
- [0063] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(Magnetic Media), CD-ROM, DVD와 같은 광기록 매체(Optical Media), 플롭티컬 디스크(Floptical Disk)와 같은 자기-광 매체(Magneto-optical Media), 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상

기된 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[0064]

본 발명은 도면에 도시된 실시예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성 요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성 요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다. 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 등록청구범위의 기술적 사상에 의해 정해져야 할 것이다.

도면

도면1

