

블록체인과 강화학습을 이용한 P2P 전력 거래

최정용, 황동엽, 김기형*

아주대학교, 아주대학교, *아주대학교

comrry@ajou.ac.kr, bc8c@naver.com, *kkim86@ajou.ac.kr

P2P electricity trading using blockchain and reinforcement learning

Jungyong Choi, DongYeop Hwang, Ki-Hyung Kim*

Ajou Univ., Ajou Univ., *Ajou Univ.

요 약

프로슈머의 등장과 IT 기술의 발전으로, 전력회사를 통하지 않는 P2P 전력거래가 늘어날 것으로 예상된다. 본 논문에서는 블록체인을 활용한 P2P 에너지 거래 커뮤니티 모델을 제안하여 거래 투명성 향상, 데이터 무결성 보장, 중개 수수료 감소 효과를 얻으며 이러한 블록체인 기반 P2P 전력 거래 커뮤니티 모델 환경에서 에너지저장장치를 사용하는 프로슈머의 최대 보상을 위해 강화학습을 이용하는 방법을 제안한다.

I. 서론

전기의 생산, 소비, 운반등을 IT 기술에 접목한 스마트 그리드 기술이 발전함과 동시에, 프로슈머라는 개념이 탄생하였다. 프로슈머는 스스로 태양광, 풍력 같은 재생 가능 에너지 자원을 이용해 스스로 에너지를 생산하며 소비하는 사람이다[1]. 이러한 과정에서 잉여전력이 발생하게 되고, 잉여전력을 효율적으로 관리하기 위해 다양한 에너지저장장치(Energy Storage System, ESS)기술이 발전하고 있다[2].

이러한 스마트 그리드 기술과 다양한 IT 기술들을 통해, 기존 한국전력공사 같은 중앙기관을 통해 전기를 공급받아 소비하였던 흐름에서, P2P(Peer To Peer) 전력을 거래할 수 있는 환경이 마련되고 있다. 이를 위해, P2P 로 거래할 상대방을 찾아 안전하게 거래할 수 있는 방법이 필요하다. 이 뿐만 아니라, 에너지저장장치를 사용하는 프로슈머가 에너지를 효율적으로 관리하기 위한 방법이 필요하다.

P2P 거래가 가능하기 위해 블록체인 기술이 활용될 수 있다. 블록체인은 사토시 나카모토가 제안한 최초 암호화폐인 비트코인의 기반 기술이며[3], 블록체인은 데이터를 누구든 임의로 수정할 수 없고 모두에게 데이터가 공개되는 P2P 기반의 분산 원장 기술이다. 이와 함께, 에너지저장장치를 사용하는 프로슈머의 에너지 구매와 판매시기를 조절하여 최적의 에너지 관리 효율을 통해 최대 보상을 얻기 위한 방법으로 강화학습 기술을 사용할 수 있다. 강화학습 기술은 기계학습의 한 분야로, 어떤 환경 안에 정의된 에이전트가 현재의 상태를 인식하여, 환경과 상호작용 하며 최대의 보상을 얻는 방법으로 여러 가능한 행동 중 하나의 행동을 고르는 의사결정 문제(Decision-making)를 해결할 수 있는 방법 중 하나이다[4].

본 논문에서는 블록체인과 강화학습을 활용해 P2P 전력거래를 하는 방법을 제안한다. 이렇게 블록체인을 통해 공유된 데이터로, 에너지저장장치를 사용하는 프로슈머들은 상태를 인지할 수 있고 전력 저장, 구매, 판매를 강화학습을 활용하여 결정하여 최대 보상을 낼 수 있도록 하는 방법을 제안한다.

II. 관련연구

2.1 전력 거래 모델

전력 거래를 위한 동일 배전망 단위의 Pool 기반 전력소매시장 모델이 제안되었다[5]. 소규모 Pool 을 이용해 공동주택단지, 주상복합건물 같은 동일 배전망 내에서 거래가 가능하도록 설계가 되어 구매자들은 Pool 소유자의 중개를 통해서 거래를 하게 될 수 있게 되었다. 이러한 과정은 복잡한 매칭 과정 없이 참여자들의 편익 확보를 가능하게 하여 소규모 수용가들의 참여확대가 용이함을 확인하였으며 추후 최적의 잉여전력 분배방안에 관한 연구가 필요함을 언급하였다.

2.2 Q-learning

강화학습 알고리즘 중, Q-learning 은 모델 없이 학습하는 시간차 학습 (Temporal difference) 알고리즘으로, 유한 마르코프 결정 과정(FDMP, Finite Markov Decision Process)에서 에이전트가 특정 행동을 통해 미래 보상을 최대화하는 최적의 정책을 찾기 위해 사용할 수 있다[6]. Q 값은 최적의 상태-행동가치를 추정할 수 있는 값으로 상태 s 에서 행동 a 를 선택했을 때, 기대할 수 있는 할인된 미래 보상의 합이다. Q 값을 구하는 식은 (1)과 같고 Q 값을 학습시키기 위한 식은 (2)과 같다.

$$Q_{\pi}(s, a) = E_{\pi}[R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \dots | S_t = s, A_t = a] \quad (1)$$

$$Q(s_t, a_t) \leftarrow (1 - \alpha)Q(s_t, a_t) + \alpha(r_t + \gamma \max_a Q(s_{t+1}, a)) \quad (2)$$

위 식에서 γ 는 미래의 보상을 계산하기 위한 0~1 사이의 Discount Factor 이며 α 는 학습속도 조절을 위한 0~1 사이의 learning rate 이다. Q-learning 이 시작되기전 Q 값은 임의의 값을 가지며 각시간 t 에서 에이전트는 행동을 취해 새로운 상태인 s_{t+1} 로 전이한다. 이때 보상 r_t 를 얻으며 Q 값이 갱신되고, 도달한 s_{t+1} 이 종결 상태인 경우 에피소드 하나가 끝나며 이러한 과정을 반복하여 Q 값을 학습한다.

III. 본론

3.1 블록체인의 기반 P2P 전력 거래 커뮤니티 모델

[5]에서 제안한 pool 기반 거래 모델을 블록체인을 통해 구성한다면, 거래 투명성을 향상시킬 수 있으며 데이터 조작 같은 보안 위협에 대한 위험이 낮아져 무결성이 보장되며, 중앙화된 시스템에 비해 수수료를 절감할 수 있다. 이에 따라 다음 [그림 1]과 같은 블록체인을 활용한 P2P 전력 거래 커뮤니티 모델을 제안한다.

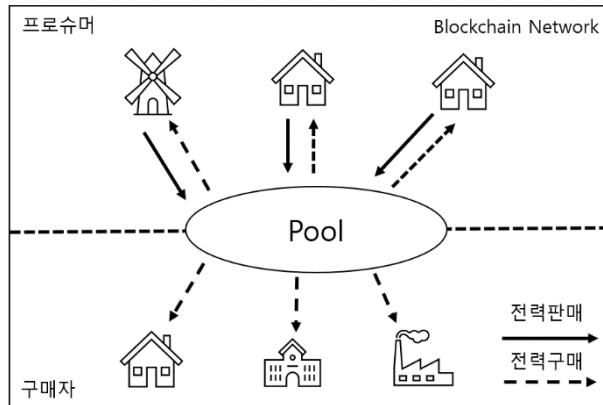


그림 1 블록체인 기반 P2P 전력 거래 커뮤니티 모델

[그림 1]에서, 구매자들은 프로슈머가 판매하고자 하는 전력량, 가격 같은 정보를 블록체인을 통해 알 수 있으며 구매자 및 프로슈머들은 Pool 소유자에게 약간의 수수료를 지불하여 P2P 간 거래가 가능하도록 할 수 있다. 에너지저장장치를 사용하는 각 프로슈머들은 생산한 전력을 소비하며, 잉여 전력은 저장하거나 판매할 수 있다. 이 모델에서는 프로슈머가 에너지저장장치를 사용한다고 가정하며, 프로슈머도 전력을 구매할 수 있다는 점을 추가하였다.

3.2 프로슈머의 보상 최대화

제안한 모델 환경에서 자신의 잉여전력을 어느 시기에 판매할지, 혹은 구매해야 하는지 최적의 판매 시기, 구매시기를 결정하여 최대의 보상을 얻을 수 있도록 강화학습을 활용할 수 있다. 현재의 시장 전력 가격, 잉여 전력량같은 현재 상태를 파악한 후, 자신의 잉여전력을 팔아야 할지 말아야 할지를 결정한다. 이를 위해 강화학습 기법 중 하나인 Q-learning을 사용하는 방법을 제안한다.

본 논문에서는 상태가 유한하다고 가정하여, 프로슈머의 의사결정을 Q-Learning을 통해 학습한 Q 값을 이용하는 방법을 제안한다. 상태 s 는 프로슈머의 전력 생산량, 에너지저장장치에 남아있는 잉여 전력, 현재 pool의 전기 가격, 인근 전력 수요등으로 구성될 수 있다. 행동 a 는 pool 시장에 전력 판매 및 구매, 전력 저장 등으로 구성될 수 있다. 보상 r 은 전력 거래를 통해 얻은 이득을 보상으로 할 수 있다.

이렇게 Q-learning을 통해 프로슈머의 최적의 의사결정을 도울 수 있지만, 본 논문에서 제시한 상황은 한정적인 상황이다. Q-learning은 많은 상태나 행동을 가진 대규모의 MDP에 적용하기 힘들며, 실제 환경에서는 상태가 무한히 커질 수 있기 때문에 단순한 Q-learning을 실제 환경에 적용하기 힘들 것이다. 문제가 복잡해지고 커지는 상황에서, 최적의 Q 값을 구하기 위해 approximate-Q-learning 기법을 적용하여 상태-행동쌍의 Q 값을 근사하여 해결할 수 있다. 특히

심층 신경망을 사용해 복잡한 문제에서 좋은 결과를 낸 Deep-Q-learning을 적용해볼 수 있을 것이다[7].

추후 연구는 위에서 제안한 P2P 전력 거래 커뮤니티 모델 환경을 만들어 Deep-Q-learning을 적용을 통해 에너지저장장치를 사용하는 프로슈머의 의사결정을 도와줄 수 있는 프레임 워크를 제안하여 성능을 평가해 볼 것이다.

IV. 결론

본논문에서는 거래 투명성을 향상, 데이터의 무결성을 보장, 중개 수수료 감소의 효과를 볼 수 있는 블록체인을 활용한 P2P 전력 거래 커뮤니티 모델을 제안하였다. 이와 함께 블록체인을 통해 얻은 데이터를 활용하여 에너지저장장치를 사용하는 프로슈머의 의사결정 문제를 도와줄 수 있도록 강화학습 기법중 하나인 Q-learning을 적용하는 방법을 제안하였다. 하지만 이 논문에서 제안한 Q-learning은 한정적인 상황에서 사용이 가능하여 상태가 무한히 커질 수 있는 실제 환경에서 적용하기 위해 Deep-Q-learning을 적용하는 추후 연구가 필요하다. 이에 추후 연구에서는 본 논문에서 제안한 P2P 전력 거래 커뮤니티 모델 환경을 만들어 Deep-Q-learning을 적용하여 에너지저장장치를 사용하는 프로슈머의 의사결정을 도울 수 있는 프레임 워크를 제안하고, 해당 프레임워크의 성능을 평가해 볼 것이다.

ACKNOWLEDGMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학 ICT 연구센터지원사업과 2018년도 정부(교육과학기술부)의 재원으로 한국연구재단의 기초연구사업의 지원과 2020년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (IITP-2020-2018-0-01396, NRF-2018R1D1A1B07048697, P0008703, 2020년 산업전문인력역량강화사업)

참고 문헌

- [1] Tuğrul Atasoy, Hülya Erdener Akınç, Özden Erçin, "An analysis on smart grid applications and grid integration of renewable energy systems in smart cities", International Conference on Renewable Energy Research and Applications (ICRERA). 2015
- [2] 이정인, 이일우, "스마트그리드를 위한 에너지저장장치 기술", 정보과학회지 31(3), pp. 33-42, 2013.3
- [3] Satoshi Nakamoto, Bitcoin: A Peer-to-Peer Electronic Cash System, 2008
- [4] 위키백과 강화학습, https://ko.wikipedia.org/wiki/%EA%B0%95%ED%99%94_%ED%95%99%EC%8A%B5
- [5] 손은태, 곽형근, 김성열, 김동민, "소규모 에너지 프로슈머를 위한 Pool 거래 기반 전력소매시장에 관한 연구", 2018
- [6] 위키백과 Q 러닝, https://ko.wikipedia.org/wiki/Q_%EB%9F%AC%EB%8B%9D
- [7] Volodymyr, Mnih Koray, Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, Martin Riedmiller, "Playing Atari with Deep Reinforcement Learning", 2013