Blockchain and Artificial Intelligence Enabled Peer-to-Peer Energy Trading in Smart Grids

by

Md. Moniruzzaman

A thesis submitted to the Lakehead University in partial fulfillment of the requirement for the degree of Doctor of Philosophy (Ph.D.)
in
Electrical and Computer Engineering

Lakehead University, Thunder Bay, Ontario, Canada

Examining Committee Membership

The following served on the Examining Committee for this thesis.

External Examiner:	Dr. Irfan S. Al-Anbagi Professor, Dept. of Engineering and Applied Science University of Regina, Canada
Internal Member (s):	Dr. Amir Ameli Professor, Dept. of Electrical and Computer Engineering Lakehead University, Canada
	Dr. Shafiqul Hai Professor, Dept. of Electrical and Computer Engineering Lakehead University, Canada
Supervisor:	Dr. Abdulsalam Yassine Professor, Dept. of Software Engineering Lakehead University, Canada
Co-supervisor:	Dr. Rachid Benlamri Professor, Dept. of Software Engineering Lakehead University, Canada

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Peer-to-peer (P2P) energy trading allows smart grid-connected parties to trade renewable energy with each other. It is widely considered a scheme to mitigate the supply-demand imbalances during peak-hour. In a P2P energy trading system, users (e.g., prosumers, Electric Vehicles (EV)) increase their utility by trading energy securely with each other at a lower price than that of the main grid. However, three challenges hinder the development of secured P2P energy trading systems. First, there is a lack of implicit trust and transparency between trading participants because they do not know each other. Second, P2P energy trading systems cannot offer an intelligent trading strategy that could maximize users' (agents') utility. This is because the agents may lack previous trading experience data that enable them to select an optimal trading strategy. Third, the current energy trading platforms are mainly centralized, which makes them vulnerable to malicious attacks and Single point of failure (SPOF). This may interrupt the transaction validation mechanism when the system is compromised, and the central database is unavailable.

To address the aforementioned challenges, we investigate the security, transparency, and economic benefit of users in a P2P trading system. In particular, we comprehensively take into account the security-based preferences of users for choosing a trading partner, unit energy trading price, and the decision to trade electricity with other users or with the main grid. To this end, firstly, this thesis proposes a novel system that combines cooperative game theory and blockchain technology to stimulate users to maximize their profits and trade energy securely. In our model, users can store renewable energy credits as assets in the blockchain and trade them with others. These assets could be converted into money by third-party exchange or cryptocurrency by the underlying blockchain platforms, which determine the conversion time. We also develop Proof of Energy Generation (PoEG) by including energy balance and distribution line loss as a consensus mechanism among blockchain energy trading members that assist coalition formation among small-scale prosumers (individual users who consume and produce energy). Secondly, this thesis proposes a novel Federated Reinforcement Learning (FRL) system combined with blockchain technology to maximize EV users' utility while preserving the security and privacy of trading transactions using the Vehicle to Everything (V2X) scheme. Thirdly, the thesis proposes the concept of Proof of State of Charge (PoSOC) as a consensus protocol to determine the winning EVs and reward them as block miners without explicitly knowing the mechanism of the blockchain system. The block miners control the system decentrally, which can avoid SPOF. Lastly, this thesis conducts comprehensive theoretical analysis and simulation experiments with respect to users' security and economic properties to demonstrate its real-world feasibility.

Acknowledgements

The journey to a Ph.D. is filled with ups and downs, adventures and challenges, stresses and breakthroughs, and failures and rewards. Personally, I find immense pleasure in having the opportunity to live the life of a Ph.D. candidate. It enables me to develop critical thinking skills, refine my writing abilities, delve into scientific problems, expand the boundaries of human knowledge, and, above all, embark on a journey of self-discovery and the quest for life's meaning. I would like to take this opportunity to express my heartfelt appreciation and deepest gratitude to a remarkable group of individuals. Without their invaluable contributions, this endeavor would not have reached such a happy ending.

First and foremost, I am incredibly fortunate to have had an exceptionally supportive and helpful supervisor, Dr. Abdulsalam Yassine. His unwavering support, dedication, insightful advice, prompt feedback, and constructive criticisms have played a crucial role in my academic journey. Honestly, I do not have words to express my gratitude to him for choosing me as his student, even though I had minimal research experience.

I would like to express my appreciation and gratitude to my co-supervisor Dr. Rachid Benlamri, for his profound supervision, guidance, support, and encouragement. I am incredibly fortunate to have had the privilege of being supervised by him.

I would also like to express my sincere gratitude to all my labmates with whom I shared the room for several years. Specifically, I want to thank Andrew Tittaferrante for his unconditional support and encouragement.

Finally, I would like to thank Lakehead University, the Department of Engineering, and NSERC for providing financial support during my Ph.D. studies.

Dedication

This thesis is dedicated to my family, specifically my parents, who have raised and supported me through every step of my life. I would like to thank them with all my heart for their continuous support and encouragement to get me where I am today. I would like to express my sincere gratitude and thanks to my lovely wife, Tahsina Faruq, for all her encouragement, sacrifices, patience, and motivation. I would like to thank my son Zohan and my daughter Zaafira for their sacrifice throughout the entire journey.

Table of Contents

List of Figures

List of Tables

List of Algorithms

List of Symbols

$\mathcal{I}$ – Set of cooperative prosumers.

$\mathcal{P}$ – Set of cooperative prosumers energy production.

$\mathcal{C}$ – Set of cooperative prosumers energy consumption.

$2^{\mathcal{I}}$ – Power set of $\mathcal{I}$.

$\mathcal{S}$ – Coalition of prosumers.

$\mathcal{S}_{seller}$ – Subset of prosumers (Seller) in the coalition $\mathcal{S}$.

$\mathcal{S}_{buyer}$ – Subset of prosumers (Buyer) in the coalition $\mathcal{S}$.

$E^{loss}_{\mathcal{S}_i}$ – Energy loss of prosumer i in the coalition $\mathcal{S}$.

$E^{loss}_{\mathcal{D}_i}$ – Energy loss of prosumer i assuming acting alone.

$\mathcal{S}_{stable}$ – Stable coalition after the algorithm converges.

$\mathcal{S}_{win}$ – Wining coalition.

$\mathcal{PS}$ – Proof score.

$\mathcal{PS}_i$ – Proof score of prosumer i when acting alone.

$\mathcal{PS}_{\mathcal{S}}$ – Proof score of coalition $\mathcal{S}$.

$\mathcal{PS}_{\mathcal{S}_i}$ – Proof score of prosumer i in the coalition $\mathcal{S}$.

E_{id} – Amount of energy transfer from prosumer i to DC.

E_{di} – Amount of energy transfer from DC to prosumer i.

E_{ij} – Amount of energy transfer from prosumer i to prosumer j in coalition, $\mathcal{S}$.

E^{lo}_{id} – Amount of energy loss from prosumer i to DC.

E_{di}^{lo} – Amount of energy loss from DC to prosumer i.

E_{ij}^{lo} – Amount of energy loss from prosumer i to prosumer j in coalition, $\mathcal{S}$.

E_i^{bal} – Energy balance $(\mathcal{P}_i - \mathcal{C}_i)$ of prosumer i.

$E_{\mathcal{S}}^{save}$ – Amount of energy savings of coalition $\mathcal{S}$.

$E_{\mathcal{S}_i}^{save}$ – Amount of energy savings of prosumer i in oalition $\mathcal{S}$.

d – DC as a prosumer.

$U_{\mathcal{S}}$ – Utility of coalition S.

$\mathcal{M}_r$ – Mining reward.

$\mathcal{P}_i(\mathcal{M}_r)$ – Part of mining reward of prosumer i.

$p_s(t)$ – Unit price of energy at time t.

X_{ij} – Amount of energy traded between prosumer i and prosumer j.

$dist_{di}$ – Distance between DC and prosumer i.

$dist_{ij}$ – Distance between prosumer i and prosumer j.

Π – Partition structure.

Π_{stable} – Stable partition structure.

$OET_{\mathcal{S}}$ – Optimized energy transactions of coalition $\mathcal{S}$.

$Block_t$ – New block of transactions at time t.

$Blockchain_{t-1}$ – Existing blockchain without $Block_t$ at time t.

$Blockchain_t$ – Updated blockchain with $Block_t$ at time t.

$\mathcal{I}$ – Set of EVs.

$\mathcal{D}$ – Set of amount of energy that the battery discharge.

$\mathcal{C}$ – Set of amount of energy required to charge the battery.

$\mathcal{PS}_i$ – Proof score of EV i.

SOC_i – State of charge of EV i.

δ_x – Scaling factor to determine score, $x \in \{on, mid, off\}$.

λ_x – Energy price at peak hour, here $x \in \{on, mid, off\}$.

s_t	– State of an EV agent at time t.
r_t	– Reward of an EV agent by taking action a_t at time t.
$\mathcal{M}_i$	– Computing machine of agent i, EV i.
p_t^{v2v}	– Price of per unit electricity using V2V method.
η	– Range anxiety coefficient.
w_t^i	– Local model of machine M_i at time t.
w_t^G	– Global model at time t.
$E_{max,i}$	– Maximum battery charging capacity of EV i.
$\mathcal{A}_t^{EV_i}$	– Action taken by the EV i at time t.
$E_{t,i}^{ch/dis}$	– Amount of electricity charge/discharge of an EV i.
$\mathcal{N}_t^{Ag}$	– Aggregator/miner at time t.
$\mathcal{N}_{t+\mathcal{H}}^{Ag}$	– Aggregator/miner at time $t + H$.
β	– The portion of the profit allocated for mining reward.
R^{EV_i}	– Reward of EV i.
p_t^{Gprice}	– Grid electricity buying/selling price at time t.
Ω_t	– Cost parameter based on the V2V charging mode at time t.

List of Abbreviations

P2P	– Peer-to-peer.
RES	– Renewable energy sources.
PoEG	– Proof of Energy Generation
PoEC	– Proof of Energy Consumption
LEM	– Local energy market
PoW	– proof of work
PoS	– proof of stake
SHA – 256	– Secure Hash Algorithm
PBFT	– Practical Byzantine Fault Tolerance
DPoS	– Delegated Proof of Stake
PoAu	– Proof of Authority
EVM	– Ethereum Virtual Machine
PoET	– Proof of Elapsed Time
AVAX	– Avalanche Blockchain
ML	– Machine Learning
CNNs	– Convolutional Neural Networks
RNNs	– Recurrent Neural Networks
LSTM	– Long Short-Term Memory Networks
Stacked – LSTM	– Stacked Long Short-Term Memory Networks Stacked
FPP	– Federated power plant
PHEVs	– Plug-in hybrid electric vehicles
DBFT	– Delegated Byzantine fault tolerance
PoE	– Proof of energy
PBFT – CB	– Practical Byzantine Fault Tolerance based-Consortium Blockchain
CF – Module	– Coalition Formation Module
IoT	– Internet of Things
ICT	– Information and Communication Technology
EV	– Electric Vehicle

MDP – Markov Decision Process
V2X – Vehicle-to-Everything
FRL – Federated reinforcement learning
RL – Reinforcement learning
PoSOC – Proof of State of Charge
V2V – Vehicle-to-Vehicle
V2B – Vehicle-to-Building
V2G – Vehicle-to-Grid
PoS – Proof of Stake
SOC – State of charge
LHEMSs – Local home energy management systems
CS – Charging station
LAG – Local aggregator
PoW – Proof of Work
PS – Proof score

Chapter 1

Introduction

For decades, energy consumers have relied on large-scale power grids to meet their domestic and commercial energy demands. However, the highly centralized nature of these power grids has been a significant concern for system reliability and resiliency and is not user-centric, leading to the emergence of Peer-to-peer (P2P) electricity trading. Technically, P2P electricity trading is a business model that is widely considered as the "Uber" or "Airbnb" of energy [1]. This is because it is a decentralized platform that enables users to leverage their own resources to achieve some economic benefit. Specifically, it facilitates users with local Distributed Energy Resources (DER) to sell their electricity (independently from the grid) to consumers. Currently, it has attracted significant attention from both industry and academia because it not only empowers individuals to take control of their own energy and costs but also promotes renewable energy usage and reduces carbon footprint.

The P2P energy market is a complex system consisting of many distributed participants with distinct energy resources and demand/supply. To this end, two state-of-the-art schemes are leading the P2P market; firstly, the Local Energy Market (LEM), and secondly, Vehicle to Everything (V2X) energy trading systems. In LEMs, prosumers, who both produce and consume energy, can trade their locally produced renewable energy (e.g., from roof-top PV panels) with each other. Alternatively, in V2X systems, an Electric Vehicle (EV) may exchange energy with the surrounding infrastructure (e.g., main grid, building, and other EVs) using cables or wireless to achieve a greater profit by leveraging their energy storage capabilities. However, these schemes come with their challenges. Specifically, for a sustainable P2P energy trading scheme, it is essential to provide users with transaction security and transparency, as well as a mechanism to efficiently adapt trading strategies to maximize their economic benefit. The key to a sustainable P2P energy trading scheme lies in its ability to adopt cutting-edge technologies.

There has been a great technological leap in the P2P trading paradigm in recent years. This leap is largely due to a combination of advances in decentralization and artificial intelligence. One real-life example is the Brooklyn MicroGrid [2], which enhanced the traditional energy grid with a 2-part mechanism, firstly, users can exchange their locally produced/stored energy with other users, and secondly, the energy transaction records are secured by blockchain, which is essentially a decentralized ledger technology. In that example, blockchain enables decentralization and democratization of the energy market and facilitates secure, efficient, and transparent P2P transactions. Specifically, due to its complex transaction-validation process, blockchain can provide a secure and transparent way to verify and record transactions while ensuring legitimate energy trading with a reliable process. On the other hand, AI techniques can maximize the utility of P2P trading participants by analyzing strategic interactions between groups of participants. For example, developing cooperative game theory mechanisms to form coalitions between prosumers with common goals can maximize the utility of individuals during trading negotiations. Another example would be to use Reinforcement learning (RL), which can be used to determine optimal trading strategies to fit the user's objectives.

Currently, there is a plethora of work on P2P energy trading as discussed later in Chapter 2. The utilization of blockchain technology in prior studies such as [3], [4], [5], and [6] is limited to maintaining immutable records in a distributed database. Furthermore, consensus protocols employed in these studies are computationally expensive, therefore, cannot support fast transaction authentication and verification process for a P2P trading system. In terms of economic benefit for prosumers to participate in P2P systems, prosumers need mechanisms that support the collaborative nature of their communities. Indeed, effective self-maintained communities require frameworks where not only individual prosumers can prioritize their own interests, but also enhances the social welfare reward of all participants. Studies such as in [3] and [5] address these issues, however, our proposed cooperative model shows that the benefits to the communities outperform those presented in studies such as [7]. Also, unlike existing work where a central entity takes decisions on behalf of participants [8, 9] , we model users' decisions, especially EV owners, as intelligent learning agents where the learning process of agents is distributed and autonomous which ensures scalability and immunity to single point of failure (SPOF).

1.1 Challenges and Motivations

Despite the opportunities brought by P2P energy trading models, there are some key challenges that hinder their wide adoption. In the following, we include some of the technical challenges and the motivations for solving them:

1. The trading participants in a P2P system do not know each other, and there is no implicit trust between them. The challenge here is to design a secure trading system that allows them to interact without necessarily trusting each other. To address this challenge, blockchain technology could be a probable technology to consider because, firstly, it is a Distributed Ledger Technology (DLT), where a database called a ledger is shared across all the nodes in the network. Secondly, the transactions can be validated by all the computing nodes in the network. In particular, the miners' node collects transactions into the block, verifies them, and starts a consensus protocol to add to the blockchain, often called mining. Thirdly, the block of data is connected through cryptographic hash functions (e.g., SHA-256), which makes it very difficult to tamper with its data. Tampering blocks implies changing all previous blocks across the majority of the nodes in the network. Finally, the transactions can be verified without the need for a trusted intermediary (e.g., a broker or bank).

2. The current trend is to develop a self-sustained prosumer community that often forms a microgrid, where they prefer to make their decision independently. However, this is difficult in a distributed system without having a mechanism that enables them to do that on their own [10]. This problem is rather significant because RES is mainly operated by local communities where users cooperate to tap into locally produced energy resources. In particular, cooperative decisions of prosumers enable trading energy with neighboring parties by avoiding costly remote utility grids and thus render green energy more affordable. To this end, self-organization using cooperative game theory can be a suitable approach to enable users' cooperative strategy, minimize distribution line loss (by trading energy with users in close proximity), and alleviate the load on the main grid. Here, self-organization represents a system that is independent and not controlled by a centralized party. A self-organized system coordinates all the demand and supply of the participants so that they are able to maximize their trading benefits through P2P transactions

3. The P2P trading system requires an intelligent mechanism that supports the users' (agents) decision-making strategy to maximize their utility. The trading strategy must take into consideration factors such as diverse locations, energy quantity, trading price, and trading time preferences. This is a challenging prospect because the agents may lack training data on previous trades that help them select an optimal trading strategy. To overcome this difficulty, we can consider the use of Federated Reinforcement Learning (FRL) for two main reasons. Firstly, it is a trial and error

process that can navigate users to adopt the best action to take with minimum or no prior experience. Secondly, the model can learn collaboratively and evolve with the dynamics of the environment without necessarily sharing users' data with others.

4. Training a continuously evolving intelligent system is expensive in terms of computing and communication overhead. The existing model-based approaches (e.g., Linear regression) use deductive processes where a general understanding of the system is required to derive an optimal output. However, these approaches can not alleviate the overhead because data needs to be transmitted from all distributed sources to a central server for processing [11]. Therefore, it is challenging to develop a training mechanism that enables distributed learning with minimal data transactions among computing nodes. To address this issue, FRL would be a suitable technology because the model is trained in a distributed fashion without explicitly transferring the data from all the local nodes to a central server. Instead, the local model updates called gradients are shared with a central aggregator to aggregate them to build an updated global model.

5. Most of the transaction platforms are controlled by a centralized third party, which is vulnerable to malicious attacks and Single point of failure (SPOF) [12]. This is a core reliability issue that may interrupt transaction authentication and payment services when the system is compromised, or the database is unavailable. To overcome such a challenge, we consider blockchain technology because it eliminates the centralized server, the cryptographic hash function prevents tampering, and data are replicated to all the computing nodes to ensure high availability. Specifically, with a single node in the network, the database can be retrieved and restored.

6. Transaction validation in a decentralized system requires substantial computational resources that may hinder the wide adoption of a P2P energy trading system because most trading systems require real-time verification without significant delay. The challenge here is to develop a mechanism that can authenticate and validate transactions with little or no delay and with minimal computing resources without sacrificing security. The consensus/mining protocol in the blockchain is a probable solution for a decentralized system. Technically, the mining node collects transactions, adds them to the block, and executes a consensus protocol to insert them into the blockchain. However, the conventional mining mechanisms (e.g., Proof of Work (PoW)) need extensive resources, which affect the transaction throughput significantly. Therefore, we need to develop a computationally inexpensive consensus protocol to resolve this issue.

1.2 Research Goal and Objectives

The main goal of this thesis is to develop an architecture that enables users (e.g., household prosumers and EVs) to participate in an intelligent P2P energy trading system that enhances their transaction profit, security, and privacy. The following objectives are deduced from the above goal:

Objective 1: To design a P2P trading architecture that enables users to trade energy with each other while ensuring their transaction security and transparency.

Objective 2: To develop a mechanism that allows users to participate in a P2P system without sharing their private and sensitive energy usage information with others.

Objective 3: To design an intelligent mechanism that will help users determine an optimal trading strategy that leverages their energy resources to maximize profit in a P2P energy trading network.

Objective 4: To develop a learning mechanism that allows P2P energy trading systems to evolve over time considering various factors such as location and energy cost with minimal computational complexity.

1.3 Research Approach and Methodology

The methodology to achieve the above objectives are as follows,

- To achieve the first objective, we comprehensively investigate the key factors (e.g., authentication, encryption, and access control) that influence users' security and transparency in a P2P trading system. To eliminate the external security threat, we use a type of blockchain called "private blockchain," which restricts the blockchain network to a specific group of participants. Then, inside this private network, we mathematically model the energy behavior of prosumers to develop a novel consensus protocol called Proof of Energy Generation (PoEG). This protocol includes a mechanism where users are penalized if they act in a way that disrupts the system. The results show that the above method enhances security and transparency, and we showed real-world feasibility by demonstrating that the scheme can be adapted to commercial blockchain platforms.
- To achieve the second objective, we thoroughly studied the users' anonymity in the blockchain system which may not be ensured due to transaction linkability and recovery issues. This may create a serious privacy risk when users share their sensitive

data with each other. To mitigate this risk, in our P2P system, we design the application layer with two main components, Control Module (CM) and blockchain. The CM is integrated with all nodes, which supervises the execution of the consensus protocol and generates a score for each user. Then, we share this score in the blockchain instead of users' sensitive energy usage information, which eliminates data-sharing risk.

- To achieve the third objective, we explore the dynamics of social cooperation to maximize trading benefits among the cooperative users (e.g., prosumers) in the P2P trading network. This leads us to develop a Coalition formation algorithm using Cooperative game theory where users' interactions based on their distance and energy price are mathematically modeled as an optimization problem. The algorithm minimizes the energy loss for every P2P transaction, which is the baseline to form a coalition because rational users want to maximize their economic benefit. Then, we analytically prove that the algorithm enables the final formed coalition to be Pareto optimal and stable. The algorithm is validated by conducting experiments using a real-world dataset (Ausgrid) where the test results show that users can increase 6% energy savings compared to the baseline model [7]. However, a limitation of this approach is that optimization using game theory is computationally expensive because it requires solving an entire optimization problem in each time step. This limitation leads us to solve this problem using a data-driven approach.

- To achieve the fourth objective, we develop a collaborative machine-learning mechanism by employing FRL in the P2P trading system. In particular, we design an efficient state space and mathematically formulate the reward function by considering the dynamics of the trading environment. Then, we develop a blockchain consensus protocol, Proof of State of Charge (PoSOC), to enable the collaborative learning aggregation to be decentralized. This method significantly reduces computing complexity. Finally, we used a real-world dataset from Ontario Electric Board (OEB) to simulate and validate the system. The result shows that the learning process is faster, and users in this system can improve at least 5% more benefit compared to the baseline models [13].

1.4 Major Contributions

The main contributions of this study are:

- We propose a blockchain-based P2P energy trading mechanism among prosumers and a novel coalition formation algorithm to determine the winning coalitions as block miners. In the coalition, an optimization technique is used to determine the list of optimal energy transactions to improve their utility.

- We develop and improve the Proof of energy generation (PoEG) protocol by introducing distribution line loss with energy balance to provide a higher transaction advantage to participants and reduce the verification latency in the blockchain.

- We design a miner selection algorithm (based on the Weighted Random Selection (WRS) method) and a reward mechanism to motivate prosumers to verify transactions and add blocks to the blockchain. The mining reward is then fairly distributed among coalition members by using the Shapeley value solution concept. In addition, we introduce a technique to punish the malicious prosumer who either refuses to mine the block or verifies invalid transactions in the blockchain.

- We propose a novel blockchain-based V2X energy trading mechanism with FRL to enable EV users to select optimal energy trading strategies for maximizing trading benefits while protecting security, privacy, trust, and transparency.

- We design the Proof of State of Charge (PoSOC) consensus protocol as a function of EVs SOC and the amount of energy trading at peak hours. The mechanism selects a miner for a specific interval which ensures transaction validity and transparency. The significance of this protocol is that EV users in the system are motivated to sell energy more during peak hours to maximize their chances of being selected as miners and receiving rewards.

- We introduce an approach to integrate the conventional Local aggregator (LAG) into the miner called mLAG, which prevents the system from SPOF and malicious attacks. Also, we propose a mining reward mechanism and a technique to deter selfish users by using the smart contract. The main advantage of this approach is that it increases systems reliability and reduces computing complexity.

1.5 Peer-reviewed Publications

Published:

1. **Moniruzzaman**, Md, Abdulsalam Yassine, and Rachid Benlamri. ”Blockchain and Federated Reinforcement Learning for Vehicle-to-Everything Energy Trading in Smart Grids.” IEEE Transactions on Artificial Intelligence (2023).

2. **Moniruzzaman**, Md, Abdulsalam Yassine, and Rachid Benlamri. ”Blockchain and cooperative game theory for peer-to-peer energy trading in smart grids.” International Journal of Electrical Power & Energy Systems 151 (2023): 109111.

3. **M. Moniruzzaman**, A. Yassine and R. Benlamri, ”Blockchain-based Mechanisms for Local Energy Trading in Smart Grids,” 2019 IEEE 16th International Conference on Smart Cities: Improving Quality of Life Using ICT & IoT and AI (HONET-ICT), 2019, pp. 110-114, doi: 10.1109/HONET.2019.8908024.

4. **Moniruzzaman, Md**, Abdulsalam Yassine, and Rachid Benlamri. ”Blockchain and Metaverse For Peer-to-peer Energy Marketplace: Research Trends and Open Challenges.” In 2022 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD), pp. 1-8. IEEE, 2022.

5. **Moniruzzaman, Md**, Seyednima Khezr, Abdulsalam Yassine, and Rachid Benlamri. ”Blockchain for smart homes: Review of current trends and research challenges.” Computers & Journal of Electrical Engineering, Elsevier 83 (2020): 106585.

6. Khezr, Seyednima, **Md Moniruzzaman**, Abdulsalam Yassine, and Rachid Benlamri. ”Blockchain technology in healthcare: A comprehensive review and directions for future research.” Journal of Applied sciences, MDPI 9, no. 9 (2019): 1736.

1.6 Thesis Organization

The remainder of this thesis is organized as follows:

Chapter 2: Background and State of the Art. Presents background information such as an overview of P2P energy transactions, the fundamental concepts of blockchain, consensus protocols, and the reason behind choosing PoS as the foundation of our proposed algorithm. We also introduce several well-known industry-level open-source blockchain platforms, such as Ethereum and Avalance, and show their compatibility to adopt into a small-scale private use case. Finally, we present the literature review to demonstrate the state-of-the-art methods used in this area.

Chapter 3: Blockchain and Cooperative Game Theory for P2P Energy Trading. Presents a proposed blockchain-based P2P energy trading mechanism among prosumers and a novel coalition formation algorithm based on PoEG consensus protocol. Then, we validate the model by using the IEEE 14-bus system and a real-world dataset called Ausgrid. We also present a novel mining reward mechanism as a function of energy savings. Furthermore, through comprehensive analysis, we prove that the system results in a stable coalition, which is beneficial for rational prosumers. Here, we implement the model by creating a sandbox consisting of the Avalanche blockchain platform. We also implement the concepts of native and universal tokens by using the Solidity of Ethereum s' Smart Contract to promote energy as a digital asset.

Chapter 4: Blockchain and Federated Reinforcement Learning based Vehicle-to-Everything Energy Trading Presents a proposed novel Federated Reinforcement Learning (FRL) system combined with blockchain technology to maximize EV users' utility while preserving the security and privacy of energy trading transactions. The proposed system is validated through comprehensive simulation experiments utilizing a real-world dataset. Furthermore, the model is implemented on the Avalanche blockchain platform to demonstrate its real-world feasibility.

Chapter 5: Conclusion and Future Work. Includes a summary of the research work, possible future direction, and timeline for the remaining work.

The organization of the thesis is illustrated in Fig. 1.1. Finally, references are included.

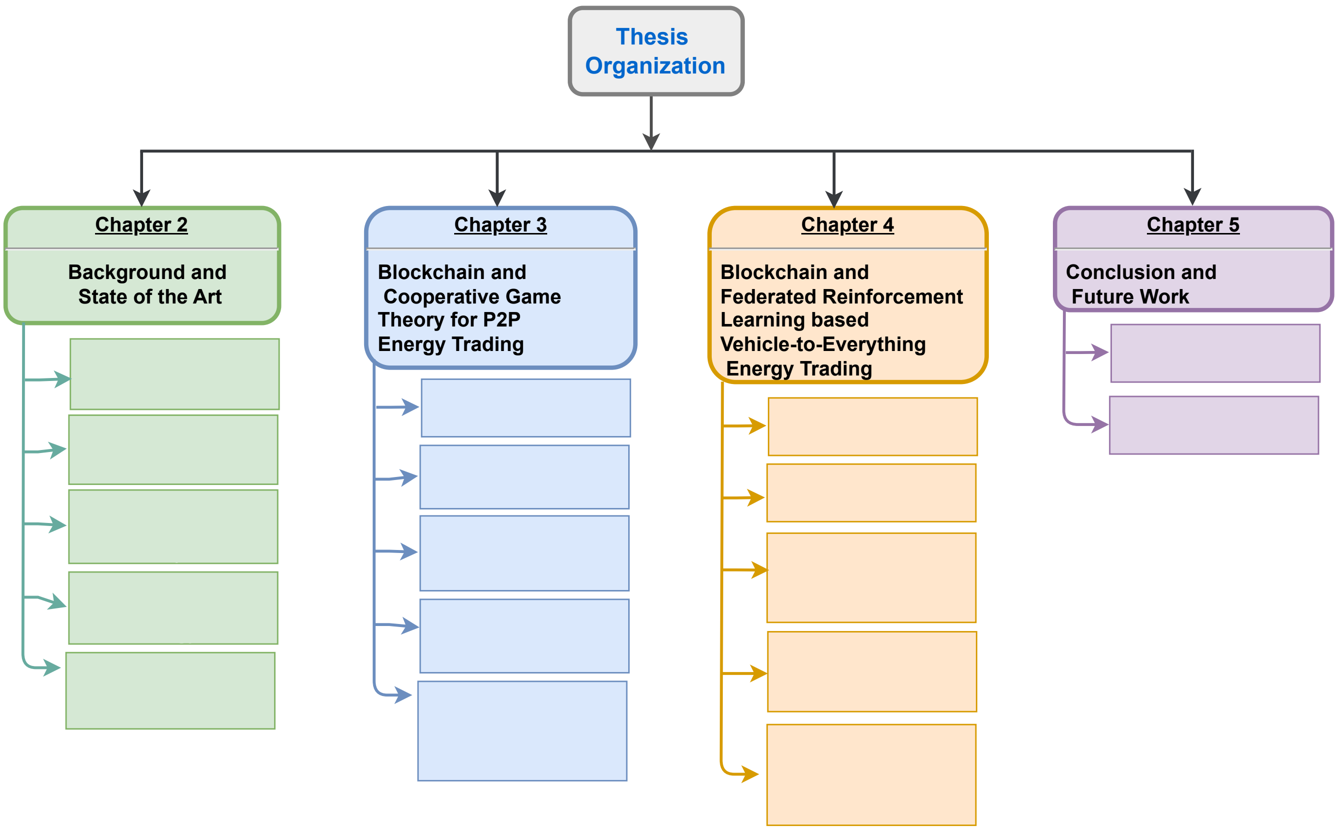


Figure 1.1: Thesis Organization.

Chapter 2

Background and State-of-the-art

In this chapter, we explore the background, state-of-the-art, and current literature surrounding a Peer-to-peer (P2P) trading platform that leverages Blockchain and Artificial Intelligence in smart grids. Specifically, we will first explore the workings of the energy system in section 2.1, before moving on to the fundamental concept of the P2P energy transaction system in section 2.2. Then, in section 2.3, we discuss the details of blockchain technology. Finally, we review the existing studies regarding blockchain, game theory, and Federated reinforcement learning (FRL) based energy trading systems in section 2.5.

2.1 Background

Susan Furnell, an energy consultant based in London, says, "The future is moving toward distributed energy, distributed generation for local businesses and for consumers" [14]. Managing the energy demand and supply during peak hours is a challenging prospect, especially in areas where the energy demand grows rapidly. To alleviate the energy demand/supply management, an increasing number of distributed Renewable Energy Resources (RESs) are being integrated into the main grid. Example of such resources consists of photovoltaic (PV) systems, wind turbine mounted on a tower, energy storage devices, Electric vehicles (EVs), and controllable loads. Rooftop solar panels installed in households contribute significantly to distributed energy resources. The global market for smart-home PV panels is projected to grow by 11% over the next six years, and residential energy storage systems are expected to increase production from 95 megawatts (MWs) in 2016 to over 3,700 MWs by 2025. [14]. Additionally, EVs can exchange energy with surrounding infrastructure to achieve a greater profit by utilizing their energy storage capabilities. According to a report by BloombergNEF, the global market for EV energy trading is expected to grow significantly in the coming years, reaching 1,000 GWh by 2030 [15]. The primary goal for renewable energy and electric vehicles is to transition to a more sustainable and environmentally-friendly energy and transportation system. To achieve such an ambitious goal, it is crucial for small-scale PV owners to ensure active participation in the energy market.

In this context, two prevailing schemes for reimbursing RES have been extensively used to stimulate prosumers to participate in the energy market: 1) Net metering and 2) Feed-in-tariff (FiT). Net metering is an energy policy that enables prosumers to save some of their surplus RES as credit by providing it to the main utility grid [16]. The energy credit can be utilized to procure additional energy at a later time when prosumers require it. This mechanism facilitates small-scale prosumers in a way that they do not require to invest extra money to buy expensive energy storage systems. However, the leading criticism of this scheme is that utility companies have no way to recover money from PV owners that they spend on fixed maintenance and infrastructure costs. On the other hand, FiT is a strategy designed as a catalyst to expedite the improvement of RES production by providing above-the-market price to small-scale prosumers. Therefore, prosumers sell their surplus renewable energy to the grid and buy from it when required. Unfortunately, the economic benefit of FiT is marginal for the participating prosumers, which fails to attract prosumers extensively. As a consequence, in many countries, the program has been suspended, and no novel scheme is offered. For example, in Ontario, the state of Canada, such a program has been eliminated because of failing to attract new users [17].

The aforementioned mechanisms face many challenges, including but not limited to ensuring the profitability of market parties and controlling by a central authority. To overcome these limitations, P2P trading mechanisms have emerged as the next-generation energy management techniques for smart grids. These techniques facilitate the active participation of prosumers in the trading system by allowing them to sell their excess energy to other prosumers or increase their self-energy consumption. P2P energy trading offers significant benefits to prosumers by enabling them to set their own terms, conditions, and prices, resulting in expected substantial gains. Also, the utility grid receives compelling benefits from this trading system in terms of curtailing peak demand [18], reducing operating costs [19], and improving seamless power supply [20]. Nevertheless, modeling a P2P energy trading system is also challenging. This is because a P2P energy trading system allows prosumers greater freedom than traditional systems by eliminating the need for centralized control. However, this lack of authority creates a trustless platform, which can make it difficult to convince prosumers to actively participate. Thus, to ensure a seamless power flow, novel mechanisms must be developed that prioritize network security and protect the privacy of rational prosumers [21]. The mechanism must ensure a consistent power flow throughout the energy network with minimal network loss. To address these challenges, one solution is to utilize Distributed ledger technology (DLT), such as blockchain. Elecbay is an example of a mechanism that utilizes this technology [18].

Blockchain is considered a disruptive technology that impacts various industries. The decentralized nature of the technology, coupled with features such as security, transparency, and privacy, makes it more attractive than centralized systems. One of the main industries that are expected to adopt blockchain extensively is the energy sector which has already been transformed from an aging sector to a modern digitized one [21]. Blockchain technology has the ability to evade existing monopolistic energy markets, offering small and large prosumers a digital platform to directly trade energy among each other in a P2P fashion. Another main advantage of blockchain is its ability to execute smart contracts, which facilitate a legally binding agreement between two parties. Several market players (e.g., WePower, Electron, etc.) are emerging with an ambitious blockchain approach to transform the energy industry and create a balanced electricity market. These new players are benefiting from the fundamentals of blockchain as a transactional platform that can trace energy production and consumption and make price adjustments. Despite its ability to provide security and transparency, blockchain technology alone cannot optimize energy trading strategies to fully maximize economic benefits for rational users. To this end, one effective solution to ensure enhanced profits for users in the P2P energy market is through the use of AI models. AI techniques can enhance the utility of P2P trading participants by analyzing strategic interactions among groups of participants. For instance, Cooperative game theory can be utilized to establish coalitions among prosumers with shared goals. Another approach involves using Reinforcement Learning (RL), which has demonstrated success in real-world scenarios. Applying RL to P2P trading enables the selection of optimal trading strategies aligned with user objectives. Next, we discuss the basics of P2P energy transactions and their interactions with some leading users (e.g., small-scale prosumers and EVs).

2.2 Peer-to-peer Energy Transaction

P2P Energy Transactions considers a next-generation energy trading mechanism that benefits users (small-scale prosumers and EVs) who are actively engaged in the energy market. In a P2P energy trading system, it is expected that the users will trade their energy with one another, unlike the conventional trading system, where a centralized third-party utility company manages the entire energy management system. Traditionally, the utility company trivially considers the economic benefits of small-scale prosumers and EV users. Hence, this decade-old system proves obsolete, leading to the emergence of a system that enables users to provide freedom to choose their buyer or seller, price, and regulation.

2.2.1 Energy Trading among Small-Scale Prosumer

The main source of energy that we consume today is generated by large fossil fuel power plants or nuclear technology. Over the past decades, researchers have been relentlessly looking for alternative energy sources. As a result, the market pattern of energy production and distribution is changing, leading to the emergence of P2P energy trading. In this evolving energy marketplace, small-scale prosumers (e.g., smart homes) that are capable of generating their energy from Renewable Energy Sources (RES), such as rooftop solar panels, can trade it with others (e.g., neighbors, factories, main grids, etc.). Examples of this scheme started in Vanderbron, Netherlands, in 2014 when it enabled an online market platform to facilitate individual consumers to buy required electricity from producers [22]. Currently, P2P electricity marketplace trading is at its early stage and is expected to grow significantly in the future. Another enabling component of the P2P marketplace stems from the generation of a large volume of usage data from prosumers' smart meters. Harnessing energy usage patterns from smart meter data by means of machine learning models could lead to the discovery of valuable information about supply and demand. Also, energy usage patterns may provide knowledge for resource optimization, demand response, and energy pricing. Such knowledge allows a party involved in P2P energy trading to plan its purchase from a peer offering energy in the forward market.

2.2.2 Energy Trading using Electric Vehicles (EVs)

In this section, we briefly discuss how we can use Electric Vehicles (EVs) to engage them in the P2P energy trading systems. First, we discuss the general overview of EVs, and then we discuss state-of-the-art P2P energy trading techniques using EVs.

EVs

The growing concerns about greenhouse gas emissions (GHG) lead to the emergence of widespread adoption of Electric Vehicles (EVs). The reason is that EVs can use clean energy and eliminate fossil fuel dependency, which offers zero CHG. Many countries in the world have made ambitious plans to transform fossil fuel-based internal combustion engine (ICE) vehicles into EVs to reduce their fuel dependency and amount of CO2 emissions. For example, by the year 2040, countries like the U.S., U.K., and Canada adopt policies to replace all their ICE-based vehicles with EVs. Additionally, Canada has set a goal that the transportation industry must reduce its greenhouse gas emissions by 12 megatonnes by the year 2030. Such an ambitious goal may impact the entire power grid, which may lead to severe energy management issues, specifically during peak hours. Electrifying transport vehicles naturally boost electricity demand, and it is estimated that there will be more than 250 million EVS on the road by 2030, resulting in an energy demand of 1.1 PWh [23]. To meet this added load to the existing power grid, innovative and intelligent energy management systems and new charging stations (CSs) are required. The existing charging infrastructure across the countries is not enough to meet the changing demands of the growing number of EVs in every day. Hence, alternative ways of charging systems need to be explored, and one way to meet the demand is using the vehicle-to-Everything (V2X) mechanism.

Vehicle to Everything (V2X) Energy Scheme

Vehicle-to-Anything (V2X) refers to the usage of EV batteries to deliver energy services and create added value from the battery asset while not in use. V2X is a paradigm where an EV can trade energy with grids (V2G), with buildings (V2B), and with other EVs (V2V) [24]. The main objective of V2X service is to create profits from the storage/battery asset of EVs through bi-directional charging control in order to benefit the electric grid, flatten the peak energy demand, or supply backup power to consumers. In this context, the term energy services can be demonstrated as to trade EVs stored charge as a form of cumulative flexible capacity in the wholesale market, which provides flexibility to the system operators and other key parties. In this study, we only focus on the V2V energy trading system and propose a novel trading mechanism that enables EVs to exchange energy intelligently using bidirectional charging technology. When an EV is charged from the grid, the supply voltage can vary depending on the location and infrastructure of the grid. The Society of Automotive Engineers has defined three levels of charging for electric vehicles (EVs). Level 1 charging uses a standard household outlet with 120 V, level 2 charging requires a dedicated Electric Vehicle Supply Equipment (EVSE) with 220-240 V, and level 3 charging, also known as dc fast charging, provides up to 90 kW of charging power at 200/450 V, making it suitable for implementing a V2G architecture in micro-grids [25]. On the other hand, V2V charging involves charging an EV from another EV, where the supply voltage depends on the capabilities of the vehicle. Due to the lower power output, V2V charging may require special cables or adapters, and the supply voltage is determined by the EV with the lower charging capabilities. The typical voltage range for direct V2V charging is from 300V to 450V [26]. For example, a Nissan leaf could be charged with a voltage of 350V [27].

Vehicle-to-Vehicle (V2V) Energy Trading

The V2V paradigm refers to a group of EVs connected locally through an application and participating in trading energy between themselves in a P2P fashion using wired or wireless connection method as discussed in section 4.2.3. This mechanism has two key benefits, firstly, the trading losses between the local EVs and the power grid can be significantly reduced, and secondly, EVs have full control of their power and can trade power at a reduced price with each other. Fundamentally, an application enables EVs to participate in such an energy market, where they need to share their trading information. An aggregator, in this case, is responsible for coordinating the control of the energy of each EV and automating the entire trading system [28]. Generally, V2V involves multiple EVs and uses smart homes and parking lots for power exchange. However, nowadays, V2V can occur anywhere, anytime, even when the EVs are on the move by using roadside charging wireless charging pads [29].

2.2.3 Overview of a Blockchain Enabled P2P Energy Trading Framework

The conventional way of trading energy in a P2P fashion for small-scale prosumers using blockchain can be described by a general framework as shown in Fig. 2.1. The diagram presents a high-level architecture of a classical blockchain-based P2P energy transaction framework. In this framework, a private blockchain platform plays the role of the core transaction processing system. Smart contracts in this framework pave the way for prosumers to create digital agreements or rules among trading parties. In fact, smart contracts eliminate the middleman in a traditional centralized marketplace. Additionally, the framework introduces a recommended system that facilitates users' feedback on the services they receive. This feedback mechanism stores user ratings in the blockchain so that future prosumers can analyze and freely choose their trading partners for better services. The main features of the framework are illustrated as follows:

- In this framework, prosumers are closely connected with a community microgrid where the energy transaction takes place after a successful negotiation and monetary

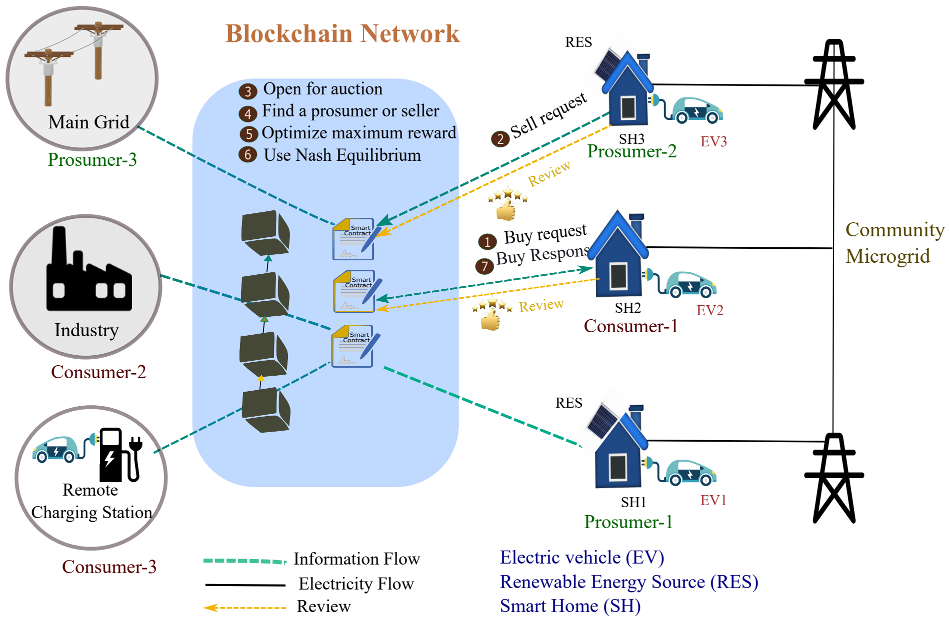


Figure 2.1: An example Framework of a Classical Blockchain-based P2P Energy Network

transaction between active agents (consumers and prosumers). The whole market price is controlled by an auction mechanism. The framework assumes that the auction mechanism is already implemented in the local mining nodes.

- The final agreement between prosumers and consumers is recorded in the smart contract provided by blockchain technology. Instead of using conventional money as a form of payment, this architecture allows for the use of a cryptographic coin or token-based transaction. These coins could be converted into money by third-party exchange (e.g., Binance). The details can be found in the Abstract on page iv. The core blockchain platform could be either a private blockchain, such as Hyperledger, or a private/public blockchain, such as Ethereum.
- When a prosumer intends to sell extra energy resources, it calls for an open auction using an energy marketplace. Buyers who participate in the auction are either consumers or prosumers. These buyer agents are normally community members or industrial consumers. These auction mechanisms may follow an agreed-upon scheme such as First Bid Sealed Auction, Second Price Auction, Double-sided auction, etc. The aim here is to provide the players with various trading vehicles to complete their transactions in a fair fashion.

- EV in this model acts as dynamic energy storage when it is attached to the smart home; it charges its high-capacity battery and stores extra electricity. These EVs could sell the stored energy once the level of storage becomes greater than or equal to a predefined threshold.
- The framework embodies a recommended mechanism to capture the users' experience measured in the form of a rating value. Both sellers and buyers recommend and rate each other on a ranking scale (e.g., 1 to 5), which represents the users' experience. This feedback is stored in the blockchain for future consumers' effortless analysis.

2.3 Blockchain Technology

The rise of blockchain technology as a transparent and responsible mechanism to distribute and store data is paving the way for new potentials of solving serious data privacy, security, and integrity issues in the smart home [30]. The ability of blockchain technology to provide an immutable decentralized ledger for securing, sharing, and keeping logs of information with permissions is becoming an attractive solution for future smart home systems [31]. Indeed, as a flagship of cybersecurity technology, blockchain has achieved remarkable performance in different types of smart home applications such as home access control [32], [33], data sharing [34], [35], and P2P energy trading [36], [37], etc. Furthermore, implementing blockchain in smart home networks is justifiable since it is independent of existing heterogeneous protocols such as Zigbee, Thread, Z-Wave, and Bluetooth LE, which are often used in smart homes [38].

2.3.1 Blockchain Basics

Blockchain was invented by Satoshi Nakamoto in 2008 [39]. It is the underlying platform of cryptocurrencies (e.g., Bitcoin) that facilitates a peer-to-peer transaction system to eliminate third-party and double spent problems [40]. It is a decentralized data structure where every block of data is cryptographically connected with the previous block's hash [41]. This hash is generated using SHA-256 (Secure Hash Algorithm), a one-way function that produces 256-bit output by transforming input values [41]. The fundamental structure of a block comprises the block number, the previous block's hash, transaction data, nonce, and timestamp [39]. The timestamp is a continuous variable, and the nonce is a random variable [42]. The static (such as a block of transaction data) and dynamic (timestamp of the transaction and a random number called nonce) data are continuously hashed by the validators or miners (computational nodes) to find a value that starts with a number of consecutive significant leading zeros. This process is widely known as a cryptographic puzzle. The miner who finds the valid hash value first considers the winner, who is given permission to add the block to the blockchain. The methodology of certifying a block, whether it is valid or not, is called Proof-of-Work (PoW) consensus algorithm [39]. The main goal of the PoW consensus protocol is to limit denial of service (DoS) attacks and potential synchronization complexities by delaying the block insertion frequency into the chain. Several consensus algorithms such as Proof-of-Stake (PoS), Practical Byzantine Fault Tolerance (PBFT), Delegated Proof of Stake (DPoS), and Proof of Authority (PoAu) perform a similar job with dissimilar characteristics [41], [43].

Generally, mainstream blockchain can be broadly categorized into two types [44], first, permissionless (or public) blockchain where everyone has permission to register, access, verify, and perform transactions, and second, permissioned (or private) blockchain where every participant require to get permission from the owner of the network to perform access, join, verify, and send and receive transactions. Bitcoin and Ethereum are examples of permissionless blockchains, while Hyperledger is a permissioned blockchain platform. Also, Avalanche is a platform that could be used for both public and private blockchain networks. Every machine in a blockchain network is called a node that stores the blockchain data structure. A node is broadly categorized into two types. Firstly, a full node that stores a full copy of the whole blockchain data structure, and secondly, a lightweight node that needs the help of a full node to participate in the network. Fig. 2.2 illustrates the data structure and the relevant algorithms used in blockchain technology for a single node. The following steps describe the main functionality of classical blockchain technology.

1. Every node (connected IoT devices in case of smart home), including the miners in a blockchain network, comprises a Memory Pool (Mempool), which includes all current transactions that are waiting to be added to the blockchain to create a new block [39].

2. A Merkle tree verifies and summarizes all the transactions.

3. If it is valid, then selected transactions are included in the block, which becomes ready for mining by miners across the smart home network.

4. Miners generates a hash of block by changing nonce and time stamp.

5. The system then compares the generated hash with the target. Once a miner finishes mining the block, it is then added successfully to the chain.

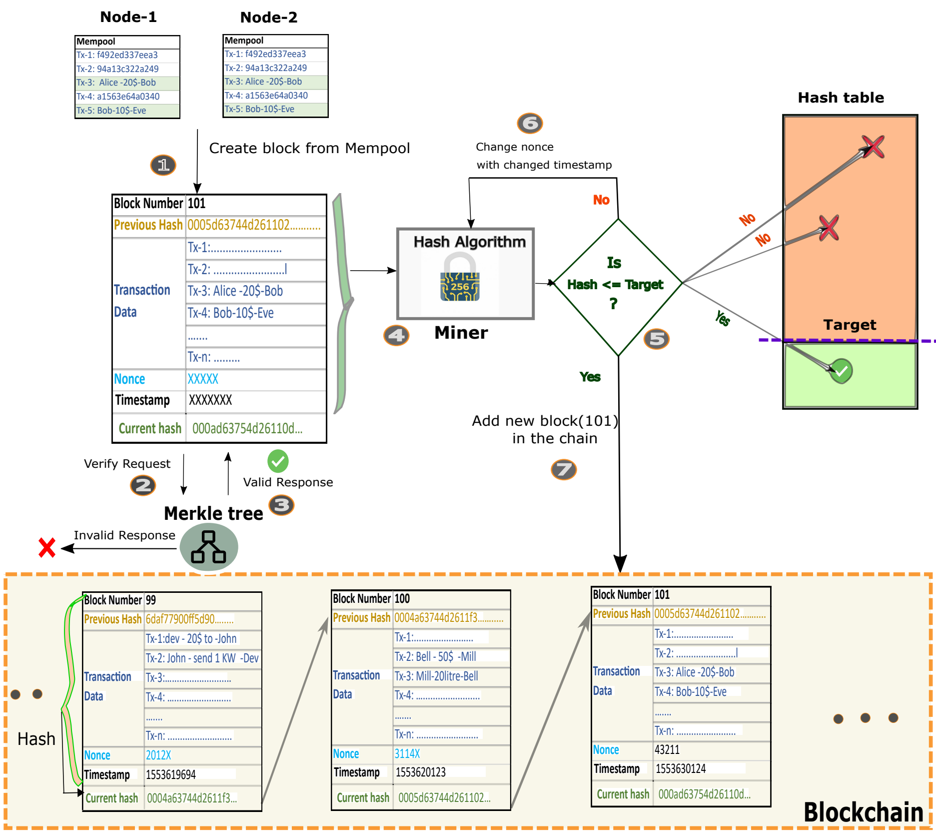


Figure 2.2: A classical blockchain architecture, internal mechanism, and workflow.

6. If the hash is above the target value, then it starts again from step 4. What it means is that the target value starts with several zeroes followed by random numbers (0000****). Therefore, if the generated hash has fewer zeroes at the beginning, it will not be considered and classified as an unsuccessful attempt. Essentially, this mechanism of leading zeroes in the target value defines the hardness of the cryptographic puzzle because the probability of getting a number less than the target becomes low.

7. If the hash is below the target value, then the PoW is verified as a success, and added the block to the blockchain. Consequently, this notification is broadcasted to the whole network to notify every connected node to delete processed transactions from the mempool [42].

2.3.2 Consensus Protocols

The consensus algorithms are the core mechanisms of a blockchain that facilitates an agreement between highly decentralized nodes to validate the correctness of transactions in a block. For example, every node in an Ethereum blockchain network uses the PoW algorithm to reach an agreement by competing to solve a cryptographic puzzle. The researcher proposes many types of blockchain consensus protocols, where each one offers distinctive features, merits, and demerits. The methodology used for reaching a consensus determines by some main performance characteristics such as transaction throughput, scalability, security, and spending resources. The main consensus protocols are the PoW, PoS, DPoS, and PBFT. Next, we will briefly discuss the fundamental concept of those consensus protocols,

Proof of Work (PoW)

The origin of PoW is to develop a mechanism so that it could limit the denial of service attacks on digital resources [45]. In this protocol, miners or validators compete with each other to add a new block to the existing blockchain by solving a complex cryptographic puzzle. Essentially, this cryptographic puzzle is simply generating a hash output that starts with several leading zeroes. A nonce (a randomly generated number) is added to a block in every iteration and thus generates a hash. The goal for each miner is to generate a hash that is lower than the specific target. This technique guarantees that no miners have the ability to influence or predict the outcome of the generated hash value. Therefore, the only feasible action might be to brute force the nonce value and hash the entire block. This trial-and-error mechanism needs high computing power that increases exponentially with the number of leading zeroes of the target hash. Once the correct hash that is lower than the target is found, it is then returned to the network to accept by all other miners if all the transactions in the block are valid and unspent. This protocol ensures high security; however, the transaction throughput is very low. For example, bitcoin only processes seven transactions per second, which is infeasible for many systems that require fast processing.

Proof of Stake (PoS)

The energy-intensive PoW algorithm led to an alternate solution for consensus called Proof of stake (PoS). The PoS protocol potentially eliminates cryptographic puzzles and introduces a random selection technique. Instead of staking high computing power, the validator stakes their wealth or asset, such as surplus renewable energy. The probability of successful mining with this protocol is proportionately associated with the staking amount of the participating validators. For example, a validator with a 200 value of asset/coin will be two times as likely to be selected as another validator with 100 assets/coin. Therefore, this protocol potentially results in faster transactions in the blockchain consuming much lower computing power. Unlike PoW, this approach does not require generating new cryptocurrency to stimulate validation, and the rewards are paid strictly by the transaction fees [46]. In this protocol, incorporating the game-theoretical mechanism may prevent collusions and centralization. The main drawback of this protocol is creating multiple blockchains, known as the fork, and mining on that is not expensive for validators. This is known as the "nothing at stake" problem, where several solutions have been proposed to address this vulnerability. One of the solutions is to lock the stake value of the participants and slash it if they try to double-sign or fork the system. Currently, PoS draws significant attention to implementing highly scaled performance systems. For example, the Avalanche blockchain platform could handle more than 4500 transactions per second because it uses PoS as its consensus protocol. Ethereum is currently using PoW but is planning to move to PoS.

Delegated Proof of Stake (DPoS)

The DPoS protocol employs a distributed voting mechanism to select delegates and witnesses. The stakeholders hold the ability to participate in the validation process. The mechanism is a more efficient and democratic version than that of the PoS because each participant votes to elect several witnesses to generate the block [47]. Every stakeholder gets a number of votes proportionate to the number of coins they own. Alternatively, they could choose to delegate their stake to other stakeholders to compete on their behalf. The delegator is responsible for validating transactions in a block and adding to the existing blockchain. Fundamentally, the delegator takes turns to produce a block every few seconds.

Delegates who misbehave or constantly fail to produce a block will lose their reputation and be quickly expelled from the system. The key difference between PoS and DPoS is that, unlike PoS, the ability to produce a block in DPoS does not depend on participants' resource ownership. The main objective of this protocol is to achieve high throughput and low electricity consumption. However, the main drawback of DPoS is that the low participation of nodes in the election process may risk the distributed system becoming a centralized one.

2.3.3 Smart Contract

A smart contract is a computer program that is embedded in the blockchain to digitally facilitate, verify or enforce the negotiation of a contract. It was created and recognized in 1994 by Nick Szabo, who is a legal scholar and cryptographer [48], [49]. It is a set of rules followed by different parties to govern a relationship to exchange values [50]. The details and permissions written in a smart contract code require an exact sequence of events to succeed and initiate the agreement of the rules written in the contract. Moreover, a smart contract may have a deadline similar to real-life traditional contracts. Traditional contracts are governed by law, while smart contracts are propelled by programming code, and they are legally binding. For example, in the US, enforcement of blockchain smart contracts may fall under the jurisdiction of the E-Sign Act as any other 'electronic contract' [51]. The main idea of the smart contract is based on a simple logic, IF-THEN. Technically there is no limit to using IF-THEN in the smart contract code. A trivial logical example might be,

- IF you transfer some asset to person A, THEN the sum of the exchange value of that asset (cryptocurrency or any mainstream currency) will be transferred to your wallet.
- A real-life example would be if a consumer research institute A agrees to pay X amount for certain product usage data inside a smart home, then the smart home-owner releases relevant information.

A logical example of a smart contract between two parties involved in buying and selling energy might be,

- If person P_1 transfers A amount of electricity to person P_2, then the exchange value of that electricity (cryptocurrency or any mainstream currency) will be transferred to the wallet of P_1.

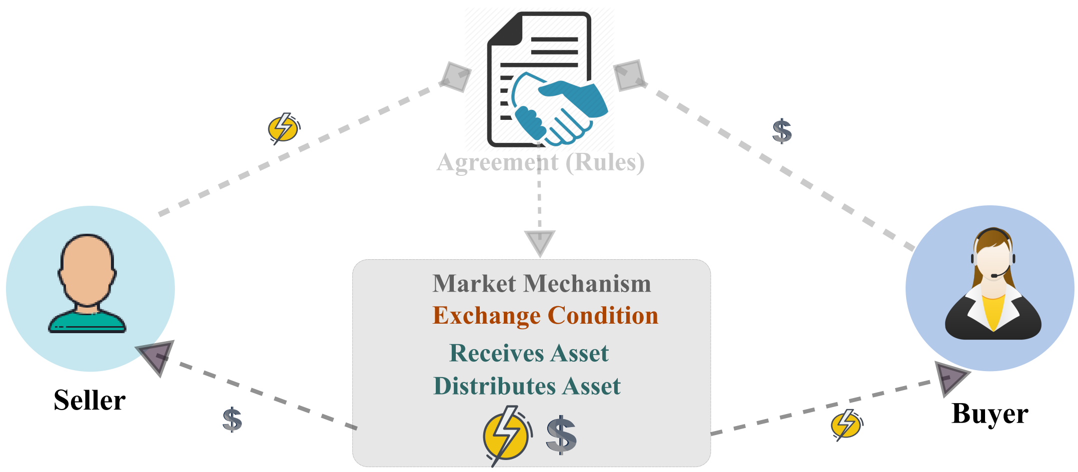


Figure 2.3: Transaction rules in a Smart Contract.

Bitcoin was the first to introduce some kind of smart contract to verify whether the amount of value transferred is actually available in the senders' account or not. Unfortunately, the contract written in Bitcoin is a Turing-incomplete language. Later, Ethereum introduced a smart contract mechanism that is more accurate and powerful since developers can create their own customized contract in a Turing-complete language, Solidity [52]. Similarly, the Hyperledger platform has its own smart contract facility, which is written in another Turing-complete language called chaincode [53]. In general, each node in the blockchain network keeps a copy of three crucial pieces of information, including the history of all smart contracts and transactions as well as the current state of all smart contracts [54]. Fig. 2.3 demonstrates the transaction rules of a smart contract.

2.3.4 Industry Level Blockchain Platforms

In this section, we briefly discuss some of the popular and mainstream industry-level blockchain platforms, including Ethereum, Hyperledger, and Avalanche.

Ethereum

The first-generation blockchain that developed Bitcoin has a limited facility for application programmings such as Turing-complete languages and storage of embedding auxiliary data

[55]. This problem was resolved by a new technology called Ethereum, which was introduced by Vitalik Buterin in 2013 [54]. Ethereum supports a general-purpose programmable platform for not only storing transactions but also executing customized programs in the blockchain [55]. For instance, Solidity, which is a programmable platform to write a smart contract for Ethereum, runs in a virtual machine widely known as Ethereum Virtual Machine (EVM) [56]. Furthermore, Ethereum reduces the mining time of blocks by 40 times less than that of mining time in Bitcoin [54], [55]. The fundamental property of EVM is to encapsulate the whole blockchain into a node which isolates a local machine from the blockchain program [49]. Therefore, the malicious application cannot affect the performance of a local computer or even can not gather any knowledge [49]. Thus, the system is obscure. Ethereum has an embedded programming language that allows the smart home user to develop their applications. Besides, this platform has a smart contract and Decentralized Applications (DApps), which is open source and operate autonomously. Therefore, blockchain-based smart home applications (e.g., P2P energy marketplace) would be smooth to implement and maintain on top of this platform.

Hyperledger

Bitcoin and Ethereum blockchain platforms employ the PoW mechanism as part of their consensus protocol to allow miners to insert new blocks in the blockchain. This technique enhances both platforms to become more secure, immutable, and resilient as a public blockchain; on the contrary, this is an obstacle to becoming a high-performance and scalable platform. Although Ethereum could be used for both public and private blockchain networks, low scalability resists its' extensive adoption. In the case of a small-scale application where the resources are limited, this affects even more.

To address this problem, Hyperledger, which was established by the Linux Foundation as an open-source project in 2016, emerged as a popular private blockchain platform for running smart contracts, with a modular architecture allowing various pluggable functions [57]. The fundamental property that makes Hyperledger different from Bitcoin and Ethereum is that the consensus protocol is pluggable, which means that the user has the freedom to choose a consensus algorithm (e.g., PoS). Currently, the PBFT protocol is widely used as a consensus mechanism. This protocol is one of many that can be employed in a private blockchain where a two-thirds similar response from the miner requires adding a new block. This mechanism empowered Hyperledger to scale up its performance significantly. Unlike Bitcoin and Ethereum, Hyperledger does not have in-built commercial cryptocurrency [57]. However, the platform has the capability that allows users to create customized cryptocurrency or coins. Due to this flexibility of modification, many organizations and individuals contributed to this generic platform.

Avalanche

Avalanche [58] is an open-source blockchain platform often called "Blockchain 3.0" used for launching highly decentralized applications. The key property that makes a difference between Avalanche and other mainstream blockchain platforms is the consensus mechanism. The platform has significantly improved blockchain 1.0 (Bitcoin) and 2.0 (Ethereum). The first and second-generation blockchain platforms rely on PoW. The drawback of utilizing PoW is that the platform can not scale high-performance transaction throughput. However, Avalanche employs the Snowman consensus protocol similar to the PoS, which scales with high performance. Furthermore, the Avalanche (AVAX) facilitates infrastructure for creating a trading platform, decentralized applications, and digital assets. It is considered the Internet of assets for issuing and trading all digital goods. For example, prosumers could create their blockchain network for their extra renewable energy as a digital asset and trade it with other prosumers. The platform allows anyone to create their private blockchain without developing a new blockchain, cryptocurrency, and digital assets from scratch. In this platform, a digital asset issuer can determine the features, such as privacy-enabled transactions, and decide whether it wants to be on a private or public network. Here, a custom smart contract creates the asset to help to design the market functionality.

The Avalanche blockchain consists of X-chain, P-chain, and C-chain. Fundamentally, X-chain is the blueprint and an instance of the Avalanche virtual machine used for exchanging assets. The C-chain is the smart contract chain which is an Ethereum VM; as a result, the platform supports the Ethereum toolkit and contract applications written in Solidity. The P-chain is the administration chain that enables users to stake a certain amount of assets to become validators or delegators and get a reward out of it.

2.4 Artificial Intelligence for P2P Energy Trading

In this section, we will explain some AI techniques that we used in our proposed P2P energy trading systems, including cooperative game theory and Federated reinforcement learning.

2.4.1 Cooperative Game Theory

Cooperative game theory is a branch of game theory that deals with the analysis of strategic interactions among a group of players who can cooperate with each other to achieve a common goal. In contrast to non-cooperative game theory, where players operate independently, cooperative game theory centers on how players unite to form coalitions and work towards maximizing their combined outcomes. The main concept of a cooperative game theory is making coalitions which can be any subset of the players in the game, and its members share a combined payoff that depends on the actions they take collectively. Cooperative game theory investigates various ideas such as the core, the Shapley value, the nucleolus, and the bargaining set. The core of a cooperative game is a collection of stable outcomes that can be deviated from by any coalition. The Shapley value is a technique for allocating the total payoff among the players based on their marginal contributions to each potential coalition. Despite its promising origins, cooperative game theory has been utilized far less frequently than non-cooperative theory as a predictive tool in economics. However, cooperative theory offers a significant advantage in providing insights into the behavior of coalitions and how subgroups of players engage in negotiations over which actions to take. It is widely used, particularly in determining trading strategies. Currently, many research studies use it for maximizing trading profit for small-scale prosumers' benefit [7, 59, 60]. The detailed description of the mathematical formulation is described in section 3.2.3.

2.4.2 Federated Reinforcement learning

Federated Reinforcement Learning (FRL) is a promising ML technique that enables the training of neural network models on local devices in a distributed manner without compromising data privacy. The FRL training process comprises two stages: local model training and global aggregation of updated local models. Let $\mathcal{M} = \{\mathcal{M}_1, \mathcal{M}_2, ..., \mathcal{M}_n\}$ be the set of n local machines. At time t, assume that each local machine $\mathcal{M}_i$ trains the global model w_t^G using its local trading experiences referred to as a local dataset to create a locally updated model w_t^i. Once the local updates are transmitted to the LAG, it estimates a new global model denoted as $w_{t+1}^G = f(w_t^1, w_t^2, ...w_t^n)$. Then, the newly formed model w_{t+1}^G is sent back to all local devices to replace their global reference model to train at time $t+1$. In contrast to the conventional distributed Deep Reinforcement Learning (DRL) model, which requires global access to all local device data, the FRL method does not necessitate local data sharing, preserving data privacy for local devices. Thus, FRL is an attractive option for organizations seeking to ensure data privacy while utilizing ML techniques. Fig. 2.4

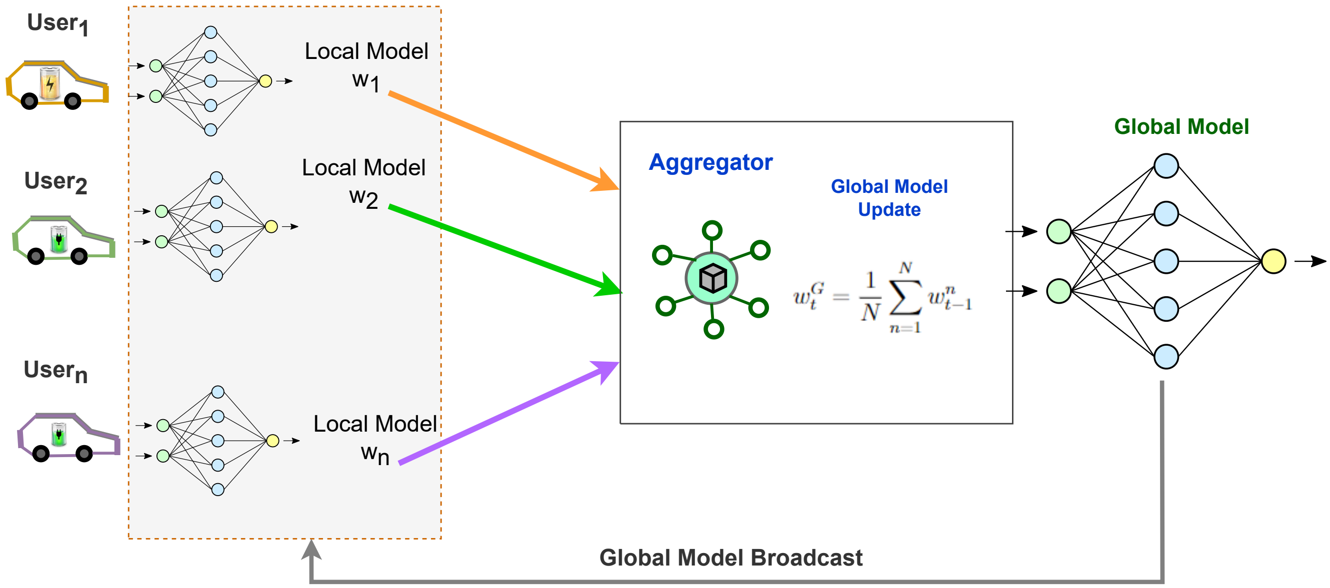


Figure 2.4: The mechanism of training and aggregation for an FRL-based P2P trading system.

shows the mechanism of training and aggregation for an FRL-based P2P trading system. The detailed description of the mathematical formulation is described in section 4.2.4

2.5 Literature Review

The current literature proposed plenty of research work on P2P energy trading systems employing cooperative game theory, Federated Reinforcement Learning, and blockchain. In this section, we present the most relevant studies and show the leading differences compared to our proposed model. For convenience, we divide the entire section into two subsections: P2P Energy Trading in Local Energy Market and V2V Energy Trading in Smart Grid. In the following, we explain the related works for each of them.

2.5.1 P2P Energy Trading in Local Energy Market

The study in [61] proposes a P2P energy trading system using non-cooperative game theory to facilitate trading benefits to rational agents (e.g., seller or buyer) individually. However, the study does not consider the transaction security and privacy of the participants. Similarly, the authors in [62], [63], and [64], analyze and develop models to support rational agents who demonstrate non-cooperative behavior. The main advantage of employing a non-cooperative game in a P2P energy trading system is that rational prosumers try to maximize their own profit regardless of the benefits of other prosumers with limited information. However, the main drawback of such a game is that parties do not explore the dynamics of social cooperation to achieve enhanced benefits. The study in [65] formulates the market prosumers as a generalized aggregative game. The author also uses a distributed market-clearing mechanism that guarantees the convergence of a strategically stable and economically profitable system using a generalized Nash equilibrium. However, the main limitation here is that the users need to rely on a distribution network operator (DNO) to operate the system. Similarly, in [66], the authors focus on multiobjective function (MOF) to optimize cost from DERs that include solar and wind in multi-Microgrid scenarios. However, the research does not consider the security of P2P transactions and social cooperation for a self-sustained community.

To this end, several studies, such as those in [67] and [68], were tailored to facilitating microgrids to form coalitions by utilizing a cooperative game theory. Although these studies propose mechanisms that are new in terms of cooperative strategies, they were primarily focused on minimizing distribution line loss. For example, in [67], the authors proposed a novel method that provides a Microgrid the ability to become autonomous and self-organized by utilizing cooperative games to maximize their utility by reducing distribution line loss. The work in [69], considers several main properties, including geographical location, demand-supply, and pricing mechanism, while creating a coalition between prosumers to improve their social welfare. Even though the study ensures users' economic benefit by using Nash equilibrium, they do not focus on security and reliability issues regarding transaction verification. However, the aforementioned proposed strategies, including [70] and [71], suffer from several limitations. Firstly, those mechanisms can not guarantee a prosumer-centric mechanism because the trading terms and conditions are controlled by a centralized party or the main grid operator. Secondly, the inherent security risk in P2P energy transaction systems (e.g., DDoS attack) is not well addressed as the system needs to be immune from malicious attackers. Finally, the trust between untrusted participants in a P2P transaction mechanism is of utmost importance for a sustainable P2P market that is not considered.

Existing studies such as [60] address the security, transparency, and trust issues of P2P energy trading systems using dynamic pricing using blockchain. The optimization is coordinated by a constrained game that provides an optimal price of exchange. Similarly, the work in [7] proposes a demand-side management model to reduce the peak-to-average ratio where the blockchain guarantees the security of the trading profiles. In [6], the authors propose a Stackelberg game with consortium blockchain to eliminate trusted intermediaries in credit-based P2P energy trading systems. In addition, research in [59] uses the Practical Byzantine Fault Tolerance based-Consortium Blockchain (PBFT-CB) technique, but the proposed solution is limited to the pricing mechanism. Furthermore, the work in [72] and [73] is closely related to the above works but focused on P2P residential energy systems and controllable distributed energy resources (DERs), respectively. The Brooklyn Microgrid [2] is an example of a real-life application where a number of peers in different microgrids trade renewable energy using blockchain. The study in [74] shows another blockchain-based energy trading application inside the campus of Washington State University. Therefore, apart from enhancing security, blockchain usage in previous studies is limited to record-keeping in a distributed database. Several studies have proposed consensus protocol in a P2P transaction framework for distributed energy transactions, such as the work in [75], [76], and [77]. However, these studies do not consider mining reward mechanisms for both non-cooperative and cooperative users in a community.

With regards to the above analysis, we can infer that the existing research has several limitations: 1) security and privacy of individual users are not well addressed, 2) lack of decentralization results in SPOF, 3) the current consensus protocol is computationally expensive to support fast transaction authentication and verification, 4) lack of mechanisms to support cooperative decisions to maximize economic benefit for a self-sustained community, and 5) mining reward mechanism in terms of energy as an asset is not considered. Table 2.1 summarizes the comparison between existing works and our study with several key technical properties that are closely associated with the above limitations. Unlike the existing work, our study includes all those essential components for a sustainable P2P energy trading system.

2.5.2 V2V Energy Trading in Smart Grid

The open literature provides plenty of research work on V2V energy trading mechanisms. This section presents the most relevant work and exhibits significant contrast compared to this study.

The energy trading of EVs' intelligent charging control mechanisms is a relatively new and fast-progressing field. In [79], the authors employ deep RL to control EV charging

Table 2.1: Comparison between our work and existing P2P energy trading studies

Properties	Work[67], [78], [68]	Work[65], [66]	Work[3] ,[4]	Work[7]	Work[5]	Work[6], [2], [69]	**Our Study**
Decentralization	-	-	✓	✓	✓	✓	✓
Trading Cooperation	✓	-	-	✓	✓	✓	✓
Cost Optimization	✓	✓	✓	✓	✓	✓	✓
Security and Privacy	-	-	-	-	-	-	✓
Consensus Protocol	-	-	-	-	✓	-	✓
Mining Mechanism	-	-	-	-	-	-	✓
Self Executing Contract	-	-	✓	✓	-	-	✓
Energy as an Asset	-	-	-	✓	-	-	✓
Demand-side Management	-	-	-	✓	-	-	-

to maximize self-consumption of smart home renewable from rooftop photovoltaic (PV) panels and SOC of EVs at departure. However, the study maximizes EV users' utility for a limited time. To address this issue, the study in [13] proposes a deep reinforcement learning-based approach to training the ML learner by optimal charging control policy employing charging Control Deep Deterministic Policy Gradient (CDDPG). However, only the V2G method is considered in their model, and there is no apparent formulation for parameters while determining participants' rewards. Similarly, researchers have proposed a plethora of approaches as in [80, 81], and [8, 9] to enable single and multi-objective reinforcement learning, respectively, to achieve an optimal charging control strategy. Although these studies empower EV users by taking intelligent strategies, the learning process of agents is coordinated by a central system.

The aforementioned research works raise a critical concern, including but not limited to users' data privacy and security, which may hinder the systems' widespread adoption. A typical way to address the issue in the literature is to train intelligent agents locally with their own devices. One of the early research work in [82] employs FL to reduce communication overhead and enhance data privacy for mobile users on wireless-edge servers. Unlike [83], the work in [84] tailored FRL to predict energy demand for EVs in a connected network. This study aims to reduce the communication overhead significantly between a Charging Station Provider (CSP) and a Charging Station (CS). In this work, EVs send their local model to the CSP without revealing the entire training dataset, which protects users' privacy. More recently, FRL has become widely used in sectors involved with control problems. One of the recent studies in [85] tailored FRL to control home appliances and energy storage devices for multiple smart homes for efficient Local Home Energy Management Systems (LHEMSs), which is closely similar to the work in [86].

The work in [87] is closely similar to our work, as the author optimizes the EV charging/discharging strategy while securing the Vehicle energy network (VEN). They proposed a joint Stackelberg-matching-auction game and consortium blockchain-based mechanism. The work considers dynamic wireless power transfer protocol (DWPT) to enable EVs to charge while they are on the move [87]. However, optimization using game theory is computationally expensive because it requires solving an entire optimization problem in each timestep. The work in [88] proposes a novel Private charging pile (PCP) sharing system where the EVs and PCPs are modeled as a joint coalition-matching game. In this case, the blockchain is efficiently used to preserve the security of the PCP network by minimizing encryption signature size. This reduces the computing burden of consensus protocol. As for the optimization, the work use game theory, which also solves an entire optimization problem in each timestep, similar to work in [87]. To address this problem, we consider an FRL agent for providing an optimal trading strategy that continuously evolves and improves efficiency without sharing their energy trading records with a central server.

To this end, one of the common approaches for FRL-based systems in the current literature is that the Global Server (GS) is controlled by a centralized authority. Those approaches do not consider security and privacy issues. To address the security issue, several works proposed blockchain-based FL mechanisms in areas including Mobile-edge Computing (MEC) as in [12] and EV energy trading systems in [89]. Unlike the existing works, our proposed model decentralizes the GS by randomly selecting an EV agent/user as an aggregator to replace a classical centralized GS. In addition, the same agent will become the miner and add transaction data blocks to the blockchain and earn additional rewards. The work in [90] uses FL and blockchain to address the security and trust issues to stimulate EV users to share their data for collaborative analysis for an improved driving experience using the Internet of Vehicles (IoV) scheme. Similarly, the work in [91] proposes FLchain using delegated PoS (DPoS) blockchain consensus protocol and FL to reduce network bandwidth and increase security. In [92], a consensus protocol called Proof of benefit (PoB) is proposed to stabilize the demand and supply of a P2P energy marketplace. In this mechanism, EVs are given a benefit value that determines the contribution of an EV toward efficient grid performance by optimizing charging and discharging schedules.

Despite enhancing security and trust in FRL, the usage of blockchain in the previous studies is limited to keeping records in a distributed database. Unlike existing work, we propose the PoSOC protocol, which aims to provide energy trading benefits during peak hours to EVs with higher energy balances compared to others in the system. Specifically, EVs try to take trading strategies that maximize their utility and proof score with regard to the time of the day. Table 2.2 summarizes the comparison between existing works and our study with several key technical properties that are closely associated with the above limitations. Unlike the existing work, our study includes all those essential components for a sustainable V2V energy trading system.

Table 2.2: Comparison between our work and existing V2X energy trading studies

Properties	Work [82]	Work [83]	Work [84]	Work [85], [86]	Work [12]	Work [89]	Work [87]	Work [88]	**Our Study**
Decentralization	✓	✓	✓	✓	✓	✓	✓	✓	✓
EV/V2V Energy Trading	-	-	-	✓	✓	✓	✓	✓	✓
Distributed LAG	-	-	-	-	-	-	-	-	✓
Security and Privacy	-	✓	-	-	✓	✓	✓	✓	✓
Enhancing Trust	-	-	-	-	✓	✓	✓	-	✓
Novel Consensus Protocol	-	-	-	-	-	-	-	-	✓
Mining Mechanism	-	-	-	-	-	-	-	-	✓
Mobile-edge Computing (MEC)	-	-	-	-	✓	-	-	-	-
Stackelberg Game	-	-	-	-	-	-	✓	-	-

The aforementioned research gaps, as explained in Table 2.1 and Table 2.2, lead us to propose a P2P energy trading system for small-scale prosumers and for EV users, respectively. The next section presents a trading system for locally connected prosumers using blockchain and cooperative game theory.

Chapter 3

Blockchain and Cooperative Game Theory for P2P Energy Trading

In this chapter, we propose a P2P energy trading system that combines blockchain technology and cooperative game theory that enables a more secure, efficient, and sustainable energy system in the smart grid while also providing economic benefits to producers and consumers of energy. In section 3.1, we discuss the background and the existing challenges of the proposed P2P trading systems. Then, section 3.2 discusses the P2P energy trading model based on blockchain and cooperative game theory. In section 3.3, we represent the coalition formation mechanism using the Proof of energy generation (PoEG) consensus protocol followed by the system performance analysis and results in section 3.4. Finally, the summary is discussed in section 3.5.

3.1 Introduction

Peer-to-peer (P2P) energy trading is emerging due to the growing interest in Renewable Energy Sources (RES) as a key strategy to fight the global warming crisis. Indeed, the adoption of distributed RES is paving the way to creating new forms of market players, where prosumers (producers and consumers of energy) can trade electricity amongst one another at lower costs [78]. With P2P energy trading paradigms, prosumers can play the role of sellers or buyers independent of the main grid. Furthermore, P2P energy trading supports participants who do not have any generation capacity by enabling them to purchase electricity from the open market at lower prices than those offered by utilities, making clean energy more accessible to people [93, 94].

3.1.1 Challenges and Motivations

Despite the opportunities brought by P2P energy trading models, they also face challenges that hinder their wide adoption. In the following, we include some of the technical challenges and the motivations for solving them.

- **Security:** The P2P models require a secure, efficient, and transparent system that fosters economic incentives for users [95, 3]. Technically, it is not trivial to trace each unit of electricity on the network and audit transactions without a system that is immune to tampering. To address this challenge, we propose the use of blockchain technology because it creates a distributed ledger (database) where each transaction is validated through consensus between multiple parties. Therefore, every participant is validated before being engaged in energy trading. Furthermore, the transactions can be verified without the need for a trusted intermediary (e.g., a broker or bank).

- **SPOF:** The majority of transaction platforms rely on centralized servers to manage and coordinate their networks, which presents a vulnerability to a single point of failure, potentially interrupting transaction authentication and payment services [96]. This problem is significant and requires an effective solution. To tackle this issue, blockchain technology could be a suitable alternative as it stores data on a distributed network of computers, ensuring high availability. In this system, every node maintains a copy of the blockchain and verifies new transactions for validity before adding them to the blockchain. With no central authority, no single node has absolute control over the network. Instead, all nodes work together on a mechanism called the mining process to reach a consensus on the state of the blockchain. Furthermore, the conventional mining mechanisms (e.g., Proof of Work (PoW)) need extensive resources, which affect the transaction throughput significantly. To address this issue, we developed a computationally inexpensive consensus protocol called Proof of Energy Generation (PoEG) to enhance the throughput of the P2P trading system.

- **Cooperative Game Theory:** The recent trend in energy trading systems is to develop a P2P trading network that empowers prosumers to make independent decisions and leverage their energy resources within the community, resulting in a closed-loop economy. However, this is challenging, especially in a distributed system without a mechanism that enables independent decision-making [10]. This is particularly significant because local communities mainly operate RES, and their cooperative decisions enable trading energy with their neighbor's parties to become independent of external sources, such as the main utility grid. Therefore, it is essential to allow prosumers to make cooperative decisions, fix their energy prices, and trade among themselves to improve the monetary benefit of a self-sustained community. Alternatively, they can prefer to trade with costly external sources when the cooperative members can not balance the demand/supply because of their volatile RES generation. Self-organization, using cooperative game theory, can be a suitable approach to enable users to collaborate and minimize distribution line loss. In this regard, self-organization represents a process by which a group of individuals organizes their energy resources and adapts to the demand and supply of participants to maximize their P2P trading profit.

In the case of the P2P energy trading in a smart grid, the existing studies [97, 98] neither consider the objective of transaction security nor the goal of maximizing individual prosumers' trading profit. In [99, 100], the authors enhance the security, privacy, and trust by employing blockchain; however, they did not consider individual users trading profit. Specifically, the research in [100] uses Blockchain-enabled Fog Computing Model (BFCM) to enhance security, transparency, and trust only. In [7, 59, 60], the authors propose blockchain-based game-theoretic trading systems to enhance rational prosumers' individual profit but do not focus on exploring the dynamics of cooperative strategies. In particular, the work in [7] uses non-cooperative game theory, which can not allow prosumers to adopt cooperative decisions to maximize their benefits. The above blockchain-based game-theoretic approaches have one common weakness, which is the usage of blockchain is limited to record-keeping in a distributed database, and the consensus protocol they use is computationally expensive. Specifically, the proposed system in [100] uses the Proof of Work (PoW) blockchain consensus protocol but fails to guarantee fast transaction verification.

Unlike the existing works, we tackle the gaps by proposing a novel system that combines blockchain technology and cooperative game theory to secure trading and stimulate users to maximize their profit. Particularly, to achieve the security and reliability of the system, we utilize blockchain technology where users can store renewable energy credits as assets, fix the price, trade them with others, and control their distribution through the community. Also, we modify the Proof of Energy Generation (PoEG) consensus protocol [101], where energy is stake as a form of asset to select a miner to reward them and increase the systems transaction throughput. However, in this study, we improve the PoEG consensus protocol significantly by including distribution line loss to provide a higher transaction advantage to participants. To enable self-organization among participants and maximize the economic benefit, this study proposes a coalition formation algorithm based on the extended consensus protocol to determine the winning coalition as a miner using Weighted

Random Sampling (WRS). Consequently, the economic benefit is twofold, firstly, users save energy by reducing distribution loss, and secondly, by receiving mining rewards. In addition, the Shapley value concept is used to fairly distribute the mining reward among the coalition members as it has desirable properties such as symmetry. Finally, to provide added security and trust, we build a technique to punish malicious miner who either verifies invalid transactions or refuses to mine the block to disrupt the services by slashing their stake and confiscating the reward.

3.1.2 Key Contributions

The combination of coalition formation games with blockchain as a means of empowering participants in a self-sustained community is a relatively unexplored subject in P2P energy trading. Unlike existing work, our mechanism ensures that rational prosumers in the P2P trading system increase their payoff while acting in coalitions. To the best of our knowledge, this is the first time that such a model is considered. The main contributions of this study are:

- We propose a blockchain-based P2P energy trading mechanism among prosumers and a novel coalition formation algorithm to determine the winning coalitions as block miners. In the coalition, an optimization technique is used to determine the list of optimal energy transactions to improve their utility.

- We improve the PoEG protocol by introducing distribution line loss with energy balance to provide a higher transaction advantage to participants and reduce the verification latency in the blockchain.

- We design a miner selection algorithm (based on the WRS method) and a reward mechanism to motivate prosumers to verify transactions and add blocks to the blockchain. The mining reward is then fairly distributed among coalition members by using the Shapeley value solution concept. In addition, we introduce a technique to punish the malicious prosumer who either refuses to mine the block or verifies invalid transactions in the blockchain.

- We implement the model using the Avalanche blockchain platform where prosumers are modeled as computing nodes to show the feasibility of the proposed model in real-world scenarios. Unlike the existing research, we implement the concepts of native and universal tokens using the Ethereums' Smart Contract to promote energy as a digital asset.

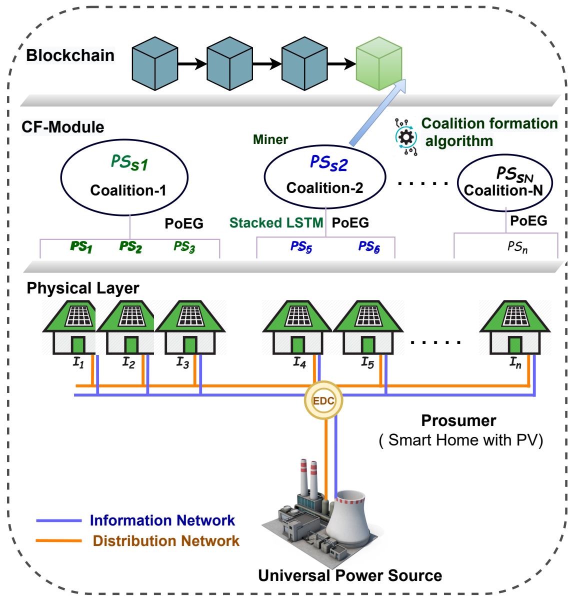


Figure 3.1: Community-based P2P energy distribution framework with blockchain and cooperative game theory.

3.2 System Model and Preliminaries

The proposed model is presented in Fig 3.1. The prosumers are connected to the electricity distribution center (EDC) of a utility company through a standard IEEE 14 bus electrical system [102]. Prosumers have roof-top solar panels as their RES, and it is assumed that they are equipped with prediction mechanisms to determine their energy consumption and production capacity. The prediction system can either use a model that can be trained locally or a pre-trained model supported by a third party. Hence, it is reasonable for a residential prosumer with low-scale computing resources to be equipped with a prediction mechanism obtained from a third party [103, 104, 105]. The model envisions that prosumers stabilize their energy demand and supply through P2P trading within local communities by coalition formation. Prosumers sell their surplus energy to members of the same coalition who are in an energy deficit. Once the process of P2P trading is settled within a coalition, the remaining energy balance is then traded with external parties, such as the EDC.

The coalition formation in this model stimulates prosumers to act together and investigates the rationality of profit allocation. Also, individual prosumers have the freedom to choose with whom they want to cooperate and trade renewables. By doing so, rational players in the energy network maximize their benefits by minimizing trading with costly utility grids and prosumers in distant locations. An example of a proposed coalition formation among prosumers is available in [106]. Also, a real-life project based on prosumers cooperation known as EnergyLab, Nordhavn in Copenhagen potentially attracts more stakeholders [107]. In addition, the model introduces an innovative approach to select a coalition as a miner that has the highest renewable energy contribution at time t. Thus, the mechanism not only increases the prosumer's chances of receiving mining rewards but also reduces the blockchain's computing complexity. In the proposed model, the blockchain provides distributed time-stamp blocks that record the transactions between trading parties without the need for a central authority. All trading transactions are replicated throughout the decentralized nodes of the blockchain using consensus protocols to prevent forgery. Mining in blockchain allows participants to validate new trading transactions and add them as new blocks in the chain. To be elected as a miner, a node must compete to solve a computational problem (e.g., Proof of Work) or show that it holds a certain amount of assets (e.g., Proof of Stake). This study proposes a PoEG algorithm as a ranking mechanism to determine which coalition is elected as a miner. Basically, the coalition with the highest rank, referred to throughout the thesis as proof score ($\mathcal{PS}$), will be elected as a miner, which yields the ability to add a new block to the blockchain network and receive a financial reward. The pricing mechanism per unit of electricity is not the core objective of this study. Therefore, to determine the value of the mining reward, we assume a market-clearing price that is more prosumer-centric compared to the utility company. The mining reward is then fairly distributed to all members of the winning coalition using the Shapley value mechanism. Furthermore, the model assumes that when all the financial statements are settled in the blockchain, the physical distribution layer is responsible for transferring the actual power. A detailed description of the main components of the proposed system is given below.

3.2.1 Decentralized Mechanism

In the proposed model, each prosumer's node consists of two main components, Coalition Formation Module (CF-Module) and blockchain. CF-Module supervises the execution of PoEG and coalition formation algorithm. Initially, every node executes an instance of the PoEG algorithm to calculate its singleton proof score $\mathcal{PS}_i$ of prosumer i, which is shared through the blockchain to avoid revealing the energy production $\mathcal{P}_i$, the energy consumption $\mathcal{C}_i$, and its geographical location l_i. This information is essential for building coalitions among the network participants of the cooperative game.

In the beginning, at time t, the prosumer who has the highest proof score, $argmax(\mathcal{PS}_i, \forall i \in \mathcal{I})$, is selected as a miner. The miner gains access to fetch the information from $Blockchain_{t-1}$.

The location information is utilized to produce a distance matrix, $\mathcal{DM}$, which holds the distance between all participants in the trading network. Next, the coalition formation algorithm executes and returns information regarding the partition of coalitions and a list of P2P transactions. The P2P transactions consist of prosumers' public keys and the amount of electricity that needs to be traded among them. There are two types of P2P transactions, the local energy transactions $T_x^{Coalition}$ that takes place inside each coalition and the external transactions T_x^{DC} of the remaining energy traded with the EDC. Finally, all the transactions are added to the smart contract of the blockchain to create $Block_t$. The miner is responsible for adding the newly formed $Block_t$ to the existing $Blockchain_{t-1}$.

At time $t+1$, the control to execute CF-Module is transferred to the prosumer, who is selected as the winning coalition. The winner is determined by a WRS technique based on their calculated proof score. This can be calculated by using $\mathcal{S}_{win} \leftarrow WRS(w\mathcal{P}_i, i \in I)$. The model follows the same technique similar to the steps at time t to execute CF-Module, and smart contract code. This process allows prosumers to achieve greater benefits than that of their singleton trading with EDC, as shown later in the thesis.

3.2.2 Physical Layer

The physical layer in our system consists of prosumers who have a single-phase connection and are connected with the three-phase Distribution System Operator (DSO) as shown in Fig. 3.2. In the system, consumers are also connected using a single-phase connection with the power distribution system. This phase difference between the participating prosumers/consumers and the distribution system creates an unbalance regime. Hence, a balancing service provider helps to make the entire system to be balanced. Essentially, the structure of the distribution network is complex [108, 109], and specifying such a system is not in the scope of this work. Therefore, we assume that every prosumer/consumer who wants to participate in the proposed trading system will have an agreement with the DSO. It is DSO's responsibility to deliver the traded amount of energy to the trading participants and provide an estimation of the loss to the trading parties. The loss of the distribution system depends on 1) network switching states (how many paths are connected at two ends: the complete path) and 2) load amount and power factor of the load. Hence, there will be an estimated variable loss allocation model as well as real-time loss calculation. DSO will measure the voltages and power flow in real time of all nodes and branches and determine the loss in real-time. To this end, we assume that the DSO has several key responsibilities, including 1) maintaining voltage quality, 2) balancing of AC network (up to a certain level of unbalance, it should tolerate), 3) power quality (harmonics, etc.), and 4) ensure minimum loss path at real-time. For the sake of simplicity, we do not consider

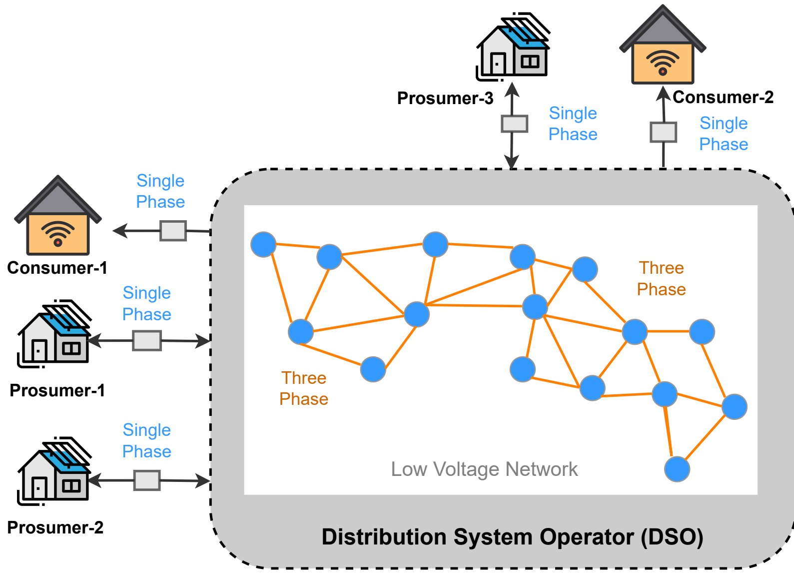


Figure 3.2: Distribution network in the Physical Layer.

specific physical layer constraints such as feeder and power flow congestion, which may hinder the efficient flow of energy [110]. We assume that the power lines are not congested and the power flow resulting from energy trading faces no limitations. However, the model can be easily extended to include constraints such as congestion, voltage drops, etc. [111].

There are several research works that focus on estimating the loss incurred on a distribution line such as those in [112, 109, 113], which demonstrates various methods for calculating power loss in a distributed network. In our work, we adopt the methods proposed in [108], where the study uses a graph-based loss allocation framework for transactive energy markets in the unbalanced distribution network. While calculating the loss, the method considers a distribution network consisting of a single-phase household prosumer and two-or three-phase loads. The most common setup entails three phase layers and a fourth neutral layer (3 Ph-4W) distribution network. The main objective of the framework is to allocate loss to each transaction based on its contribution to the real and reactive power flows of each line belonging to its path between trading parties. The transaction (Tr) path, assuming there is a line on phase layer k connecting nodes X_k and Y_k. Hence, the amount of loss ($AL_{Tr}^{X_kY_k}$) allocated for the transaction for this particular line can be determined as in the following equation [108],

$$AL_{Tr}^{X_kY_k} = \frac{P_{Tr}^{X_kY_k}}{JP_{X_kY_k}}.Lp_{X_kY_k} + \frac{Q_{Tr}^{X_kY_k}}{JQ_{X_kY_k}}.Lq_{X_kY_k} \tag{3.1}$$

Where X_kY_k represents the direction of the transaction using the line connecting node X_k and Y_k along the path. The ratio $\frac{P_{Tr}^{X_kY_k}}{JP_{X_kY_k}}$ is utilized to determine the relative contribution of the transaction's real power to the line real power flow. The term $\frac{Q_{Tr}^{X_kY_k}}{JQ_{X_kY_k}}$ determines the relative contribution of the transaction's reactive power flows to the line's real power flow. Here, $Lp_{X_kY_k}$ and $Lq_{X_kY_k}$ represent the loss due to the real power and reactive power flow on line X_kY_k on phase k, respectively. To this end, we assume that there are n lines that exist along the transaction path from one trading participant to another participant connected through the distribution line. The approximate allocated total power loss for this transaction (Tr) could be calculated as follows,

$$AL_{Tr}^{Total} = \sum_{\forall X_kY_k} \sum_{\forall k} AL_{Tr}^{X_kY_k} \tag{3.2}$$

The experimental result of the above method shows that the range of losses incurred is based on different test cases, loads, and distances between lines in an unbalanced distribution network [108].

3.2.3 Coalition Games

Let $\mathcal{I}$ be the set of all prosumers such that $\mathcal{I} = \{\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3, ..., \mathcal{I}_n\}$. A coalition, $\mathcal{S}$, is a nonempty subset of $\mathcal{I}$, and a collection, $\mathcal{H}$, is an arbitrary set of disjoint coalitions $\{\mathcal{S}_1, ...\mathcal{S}_l\} \in \mathcal{I}$ which does not necessarily need to include all prosumers of $\mathcal{I}$ [114]. If this collection is composed of all the prosumers of $\mathcal{I}$ ($U_{k=1}^{l}\mathcal{S}_k = \mathcal{I}$), the collection is called a partition Π of $\mathcal{I}$ [115]. Each coalition $\mathcal{S} \in \mathcal{I}$ has a value function $\nu(\mathcal{S})$ that determines the worth of the members' group forming the coalition. Also, each member $i \in \mathcal{S}$ has a payoff value $\nu_{i,i\in\mathcal{S}}(\mathcal{S})$. In a cooperative game, when the coalition value is a real number and can be fairly distributed among its members, it is called a game with transferable utility (TU). In this study, the distribution of gain among members of a coalition follows the *Shapley Value* concept since it considers the average of all marginal contributions of each member [116]. Therefore, the payoff of a prosumer $i \in \mathcal{S}$ could be calculated using the equation below,

$$\phi_i(u) = \sum_{\mathcal{S}' \subseteq \mathcal{S} \setminus \{i\}} \frac{|\mathcal{S}'|!(|\mathcal{N}| - |\mathcal{S}'| - 1)!}{|\mathcal{N}|!} [u(\mathcal{S}' \cup \{i\}) - u(\mathcal{S}')] \tag{3.3}$$

In (3.3), u is a function called the worth of coalition $\mathcal{S}'$, which explains the total expected payoff of the member of $\mathcal{S}'$ could obtain by cooperation. Therefore, the term $u(\mathcal{S}' \cup \{i\}) - u(\mathcal{S}')$ shows the marginal payoff contribution of prosumer i in coalition $\mathcal{S}'$. The Shapley value also satisfies several axioms as described in [115, 116]. We also adopt the following definitions.

DEFINITION 1. *(Essential):* The coalition formation game is said to be essential if the payoff of a prosumer i increases by joining coalition $\mathcal{S}$, i.e., $\nu(\mathcal{S}) > \sum_{\forall i \in \mathcal{S}} \nu(i)$.

DEFINITION 2. *(Pareto Order):* This implies that a group of prosumers prefer to be partitioned by a collection $\mathcal{H}$ instead of collection $\mathcal{Q}$ such that $\mathcal{H} \cap \mathcal{Q} = \emptyset$ [67].

DEFINITION 3. *(Stable Coalition):* A coalition formation game is said to be stable when no prosumer has the motivation to leave one coalition and join another for better payoff.

DEFINITION 4. *(Merge and Split Rules):* The merge operation implies that a coalition could form a bigger coalition if at least one of the prosumers improves its payoff without hurting others [117].

3.2.4 Proof of Energy Generation (PoEG)

The fundamental concept of the PoEG is that the prosumer who has a higher RES generation compared to their consumption in the energy market will have higher chances to be elected as a miner at time t. In this mechanism, the energy behavior of a prosumer is quantified by a Production-Consumption-Loss function, which yields the proof score [101]. Let $\mathcal{P} = \{\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3, ..., \mathcal{P}_n\}$ denotes the forecasted energy production and $\mathcal{C} = \mathcal{C}_1$, $\mathcal{C}_2$, $\mathcal{C}_3, ..., \mathcal{C}_n$ the forecasted energy consumption for each prosumer. The proof score $\mathcal{PS}_i$ of a prosumer i acting alone and trading with the EDC only, can be calculated as follows,

$$\mathcal{PS}_i = \begin{cases} (\mathcal{P}_i - \mathcal{C}_i) - E^{lo}_{id}, & \text{if} \quad \mathcal{P}_i - \mathcal{C}_i > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{3.4}$$

In equation (3.4), the proof score of a prosumer i depends on its excess of energy balance and energy distribution loss. The components E^{lo}_{id} represent the energy transfer loss from prosumer i to the EDC. If prosumer i does not have any energy credit to sell (i.e., $(\mathcal{P}_i - \mathcal{C}_i) <= 0$), its achieved proof score will be set to 0. Please note that the energy loss for transactions (e.g., E^{lo}_{id}) will be allocated by using the method as discussed in Section

3.2.2 following equation 3.2. Algorithm (1) illustrates the mechanism of formulating PoEG. The algorithm ensures that the value $\mathcal{PS}_i$ can be maximized either by increasing energy balance or by reducing distribution loss. The former may require adding more PV panels while the latter can be enhanced by simply forming a coalition with others, as shown in the next section.

Algorithm 1 Proof of Energy Generation, `PoEG`

Input: $\mathcal{I}$, $\mathcal{P}$, $\mathcal{C}$, Distance from EDC= $\mathcal{DM}$.
Output: Set of proof scores $\mathcal{PS}^*$= $\{\mathcal{PS}_1, \mathcal{PS}_2, ..., \mathcal{PS}_n\}$.

```
1: Initialisation :PS* ← ∅
2: for Each Prosumer i ∈ I do
3:    Calculate proof score by following equation (3.4). In this case, associated distribution
      loss will be measured by equations 3.1 and 3.2.
4:    if (P_i − C_i) > 0 then
5:       PS_i ← ((P_i − C_i) − E_id^lo)
6:    else
7:       PS_i ← 0
8:    end if
9:    PS* ← PS* ∪ {PS_i}
10: end for
       PS*
```

3.3 Coalition Formation Mechanism

In the proposed model, the two factors that propel prosumers to form a coalition are reducing distribution line loss and gaining the mining reward, $\mathcal{M}_r$. The proof score of a coalition $\mathcal{S}$ is denoted by $\mathcal{PS}_S$, and formulated by the following equations,

$$\mathcal{PS}_S = \sum_{\forall i \in \mathcal{S}} \mathcal{PS}_{S_i}, \quad \mathcal{S} \subseteq \mathcal{I} \tag{3.5}$$

$$\mathcal{PS}_{S_i} = \begin{cases} \mathcal{PS}_i + (E_i^{loss} - E_{S_i}^{loss}), & \text{if} \quad \mathcal{PS}_i > 0 \\ E_i^{loss} - E_{S_i}^{loss}, & \text{otherwise.} \end{cases} \tag{3.6}$$

In equation (3.5), $\mathcal{PS}_S$ represents the total attained proof score of all prosumers in coalition $\mathcal{S}$. Using equation (3.6), we calculate the proof score of prosumer i, $\mathcal{PS}_{S_i}$, by determining

E_i^{loss} and $E_{S_i}^{loss}$, the prosumer's distribution loss when acting alone and in a coalition $\mathcal{S}$, respectively. The main strategy here is that prosumer i who is a seller (i.e., $i \in \mathcal{S}_{seller} \subset \mathcal{S}$), reduces its loss by initially trade energy within its coalition before trading with the EDC. Also, the model allows prosumer j, who is a buyer (i.e., $j \in \mathcal{S}_{buyer} \subset \mathcal{S}$), to enhance its proof score based on the amount it contributes to save energy loss in the coalition. The distribution loss E_i^{loss} and $E_{S_i}^{loss}$ are calculated in equations (3.8) and (3.7) as follows.

$$E_{S_i}^{loss} = \begin{cases} E_{id}^{lo} + \sum_{\forall j}(E_{ij}^{lo}), & \text{if} \quad i \in \mathcal{S}_{seller}, j \in \mathcal{S}_{buyer} \\ E_{di}^{lo} + \sum_{\forall j}(E_{ji}^{lo}), & \text{if} \quad i \in \mathcal{S}_{buyer}, j \in \mathcal{S}_{seller} \end{cases} \tag{3.7}$$

where, $\qquad \mathcal{S} = \mathcal{S}_{seller} \cup \mathcal{S}_{buyer}$

$$E_i^{loss} = E_{id}^{lo} + E_{di}^{lo}, \quad i \notin \mathcal{S} \subseteq \mathcal{I} \tag{3.8}$$

We calculate E_i^{loss} by accumulating E_{id}^{lo} and E_{di}^{lo}, which determines the energy loss of prosumer i assuming it acts alone and its entire trading takes place with the EDC. Finally, $E_{S_i}^{loss}$ is determined based on the distribution loss inside the coalition E_{ij}^{lo} between a seller i and a buyer j, and the distribution loss between prosumer i and the EDC, E_{id}^{lo}. Similarly, we represent E_{di}^{lo} in the case when prosumer i acts as a buyer. Please note that the loss calculation of the distribution network is based on the method described in the previous section 3.2.2, followed by equation (3.2) and (3.1).

Coalition Proof Score Optimization

Every coalition seeks to maximize its proof score $\mathcal{PS}_S$ in order to win the mining reward and trade its energy in the market. Therefore, a coalition $\mathcal{S} \subseteq \mathcal{I}$ solves the following in each time slot t.

$$Maximize \quad [\mathcal{PS}_S]_t \tag{3.9}$$

subject to,

$$\mathcal{PS}_{\mathcal{S}_i} \geq \mathcal{PS}_{\mathcal{K}_i} \geq \mathcal{PS}_i, \quad \forall i \in \mathcal{S}, \quad \forall \mathcal{K} \in 2^I \setminus \{\mathcal{S}\} \tag{3.10}$$

The constraint in (3.10) ensures that at time t, a prosumer i in coalition $\mathcal{S}$ could obtain a proof score $\mathcal{PS}_{\mathcal{S}_i}$ higher than the proof score $\mathcal{PS}_{\mathcal{K}_i}$ if it joins any other coalition $\mathcal{K}$, or its proof score $\mathcal{PS}_i$ if acted alone. The above proof score maximization of a coalition $\mathcal{S}$ could be achieved by minimizing energy loss using the following equations,

$$Minimize \sum_{i \in \mathcal{S}_{buyer}} \sum_{j \in \mathcal{S}_{seller}} X_{ij}.C.dist_{ij} \tag{3.11}$$

subject to,

$$\sum_{i \in \mathcal{S}} E_i^{bal} = \sum_{i \in \mathcal{S}} E_{id} \tag{3.12}$$

$$-\sum_{i \in \mathcal{S}} E_i^{bal} = \sum_{i \in \mathcal{S}} E_{di} \tag{3.13}$$

$$E_{ij} \geq 0 \quad \forall i, j \in \{\mathcal{S} \cup d\} \tag{3.14}$$

$$\exists i \in \mathcal{S}_{seller} \subset \mathcal{S}, \quad \exists j \in S_{buyer} \subset \mathcal{S}, \quad \forall i, j \in \mathcal{S} \tag{3.15}$$

In the equation (3.11), X_{ij} represents the amount of energy traded between prosumer i to prosumer j, at time t, C represents per km line loss (estimated by the DSO), and $dist_{ij}$ represents the distance between prosumer i to prosumer j. The constraint in (3.12) ensures that surplus energy after trading within coalition $\mathcal{S}$ is sold to external parties, e.g., the EDC. Similarly, the constraint in (3.13) indicates that if a coalition S has an energy deficit, it purchases energy from the EDC. The constraint in (3.14) indicates that E_{ij} is greater than zero for the trade to happen. Furthermore, the constraint in (3.15) guarantees that the coalition formation is possible if at least one seller $i \in \mathcal{S}_{seller}$, and one buyer $j \in \mathcal{S}_{buyer}$ exists.

Miner Selection

The selection of a miner depends on the proof score that a coalition achieves at time t. The higher the proof score a coalition achieves, the more it has the chance to become a miner compared to other coalitions in the proposed system. We model the mechanism using a weighted random sampling (WRS) selection process with replacement [118]. In the WRS mechanism, the coalitions are weighted with their proof score, which determines the probability of being selected as a miner. Hence, we formulate the mining selection probability of a coalition $\mathcal{S}_i$ as follows,

$$\mathcal{P}_{\mathcal{S}_i} = \frac{\mathcal{PS}_{\mathcal{S}_i}}{\sum_{\forall j} \mathcal{PS}_{\mathcal{S}_j}} \tag{3.16}$$

In (3.16), the probability of coalition $\mathcal{S}_i$ can be obtained by the function of its achieved proof score $\mathcal{PS}_{\mathcal{S}_i}$ and the total proof score of all the coalitions $\sum_{\forall j} \mathcal{PS}_{\mathcal{S}_j}$, at time t. We show this miner selection mechanism in Algorithm (2). The algorithm ensures that all prosumers have a chance to become the miner regardless of their attained proof score. This technique is aligned with known practices of miner selection processes of the PoS consensus mechanisms, e.g., [119].

Algorithm 2 Miner Selection Mechanism using `WRS`

Input: Set of Coalitions $\mathcal{S}$, Set of coalition proof scores $\mathcal{PS}_S$.
Output: Coalition miner, S_{miner}

1: **for** Each Coalition $i \in \mathcal{S}$ **do**
2: Calculate coalition probability (3.16).
3: Let $\mathcal{P}_{\mathcal{S}i} \leftarrow \frac{\mathcal{PS}_{\mathcal{S}_i}}{\sum_{\forall j} \mathcal{PS}_{\mathcal{S}_j}}$ be the probability of coalition i to be selected as miner
4: $w\mathcal{P}^* \leftarrow w\mathcal{P}^* \cup \mathcal{P}_{\mathcal{S}i}$
5: **end for**
6: $S_{miner} \leftarrow$ Randomly select a coaliton based on the probability in $w\mathcal{P}^*$
S_{miner}

Coalition Utility

The utility of a coalition is determined by the energy saved from trading with its members and the mining reward $\mathcal{M}_r$ if it has the highest proof score. To this end, the utility of a coalition $\mathcal{S} \subseteq \mathcal{I}$ is defined as follows,

$$u(\mathcal{S}) = \mathcal{M}_r + p_t * E^{save}_{\mathcal{S}} \tag{3.17}$$

$$\mathcal{M}_r = 0; \quad \forall \mathcal{S}_k \in \neg \max\{\mathcal{PS}_{\mathcal{S}}\}, \quad \mathcal{S}_k \in \Pi_{stable} \tag{3.18}$$

$$E^{save}_{\mathcal{S}} = E^{loss}_{\mathcal{D}} - E^{loss}_{\mathcal{S}} \tag{3.19}$$

$$E^{loss}_{\mathcal{D}} = \sum_{\forall i \in \mathcal{S}_{seller}} E^{lo}_{id} + \sum_{\forall j \in \mathcal{S}_{buyer}} E^{lo}_{dj} \tag{3.20}$$

$$E^{loss}_{\mathcal{S}} = \sum_{\forall i \in \mathcal{S}} E^{loss}_{S_i} \tag{3.21}$$

In equation (3.17), we formulate the utility of a coalition $\mathcal{S}$ as a function of mining reward $\mathcal{M}_r$, and its total energy-saving $E^{save}_{\mathcal{S}}$. The parameter p_t represents the price per unit of electricity at time t. Equation (3.18) ensures that a coalition will not be allowed to receive $\mathcal{M}_r$ if its proof score is not the highest at time t. Equation (3.19) calculates the amount $E^{save}_{\mathcal{S}}$ as the difference between $E^{loss}_{\mathcal{S}}$ and $E^{loss}_{\mathcal{D}}$, the power loss due to in coalition trading and with the EDC, respectively. $E^{loss}_{\mathcal{D}}$ and $E^{loss}_{\mathcal{S}}$ are calculated in (3.20) and (3.21).

Prosumers' Utility

The payoff, x_i, of a prosumer $i \in \mathcal{S}$ is defined by the amount of energy it saves and the fraction of $\mathcal{M}_r$ that i receives. Therefore, the utility of prosumer i is calculated as follows,

$$x_i = (\mathcal{M}_r)_i + p_t * E^{save}_{\mathcal{S}_i} \tag{3.22}$$

Where, $E^{save}_{\mathcal{S}_i}$ represents the energy-saving of prosumer i in coalition $\mathcal{S}$, and $(\mathcal{M}_r)_i$ is its share from the reward after applying the Shapley Value in (3.3).

3.3.1 Coalition Formation Algorithm

Algorithm (3) illustrates the details of the proposed coalition formation. At time t, the process starts by executing the PoEG algorithm (1) to determine the proof score $\mathcal{PS}_i$ of every prosumer assuming they are acting alone. The system then separates prosumers into buyer and seller categories based on their energy balance. From lines (8) to (17), the algorithm (3) initiates coalitions with one buyer and one seller prosumer in close proximity. In this step, the algorithm guarantees that prosumers attain an improved proof score. This formation of initial coalitions represents the initial partition Π_{init}. Lines (19) to (23) optimize the proof score of coalition $\mathcal{S}$ by the function OptimizeLoss that takes the coalition structure, members' energy profile, and distance matrix as parameters. The function returns-optimized P2P energy transactions between parties of the coalition by using equations (3.9) - (3.15) implemented by the Gurobi Optimizer tool. The initial coalitions are processed through multiple merges and splits operations before the algorithm converges following lines (25) to (29). In each operation, $OET_{\mathcal{S}}$ and $\mathcal{PS}^*_{\mathcal{S}}$ are computed to reach the final stable partition Π_{stable}. In line (30), the algorithm determines the winner coalition, $\mathcal{S}_{win}$ by executing algorithm (2) based on the $\mathcal{PS}^*_{\mathcal{S}}$ at time t. In line (31), we assume that the value of $\mathcal{M}_r$ can be determined by a scaling factor δ_t of the energy savings of the entire market $\sum_{\forall i \in \forall \mathcal{S} \in \Pi_{stable}} E^{save}_{\mathcal{S}_i}$ for each time t. From lines (32) to (38), the algorithm computes the utility of the coalition $U_{\mathcal{S}}$ for $\forall \mathcal{S} \in \Pi_{stable}$ and distributes individual utility x_i at time t. Finally, in line (39), the algorithm includes all P2P transaction records in a new block $Block_t$, which is added to the existing $Blockchain_{t-1}$ and emerges as $Blockchain_t$. To this end, the steps of miner selection for the timeslot $t+1$ is as follows,

- At time t, the $miner_t$ executes the coalition formation algorithm (3) and returns information regarding the coalition structures and a list of P2P transactions based on the prosumer's supply and demand predicted by the proposed system for the next timeslot $t+1$.

- All the transactions are then added to the smart contract of the blockchain to create $Block_t$. At this point, the $miner_t$ is responsible for adding $Block_t$ to the existing $Blockchain_t$.
- Finally, the miner $miner_{t+1}$ is selected to control and execute the algorithm (3) to perform the above-mentioned two steps for the timeslot $t + 2$.

Hence, the verification of commitment is performed by a single computing node acting as a miner. Other prosumers will not receive any incentive and are not allowed to perform the task of a miner. As a result, there will be no redundancy in the mining process like PoW or PoS.

To use the proposed algorithm, a prosumer simply needs to subscribe to the proposed blockchain-based marketplace. Once they subscribe, two main components will be installed, Coalition Formation Module (CF-Module) and blockchain. CF-Module supervises the execution of PoEG and coalition formation algorithm as mentioned in section 3.2.1. These modules are responsible for the entire process.

The proposed strategy is offline. This is because the creation of coalitions depends on prosumers' energy profile and their location, and it may change every time the proposed Algorithm (3) executes. Also, the model requires the consumption and generation amount, which are normally a day or hour ahead [120]. Hence, when the energy network size is small, the computing resource required to execute the coalition formation algorithm might be trivial. In contrast, once the number of prosumers increases, the computing complexity increases, and executing in real-time would complicate the overall system. The reason is that performing a merge and split operation in real-time will be costly. However, the proposed mechanism could be used in real-time, if potential miners have resources to support executing the algorithm with the complexity as discussed in the later section.

3.3.2 Mining Reward

The proposed model utilizes the idea of an energy credit-based reward mechanism. This mechanism attempts to answer the following questions for every transaction, (1) Who pays who; (2) How to pay the reward; and (3) How much is the amount of the reward? The mechanism proposes to allocate a fraction of the energy savings $E^{save}_{\mathcal{S}_i}$ for a prosumer i who have joined a coalition at time t. The total mining reward $\mathcal{M}_r$ is calculated as follows,

Algorithm 3 `Coalition Formation`, `Proof score` and `Payoff` Calculation for a Single Time Period.

Input: $\mathcal{I}$, E_i^{bal}, $\mathcal{PS}_i$, Distance Matrix= $\mathcal{DM}$, $Blockchain_{t-1}$.
Output: Prosumers Payoff, $Blockchain_t$.

//Initialization
$\mathcal{I}_{seller} \leftarrow \emptyset$, $\mathcal{I}_{buyer} \leftarrow \emptyset$, $\Pi_{init} \leftarrow \emptyset$
for *Each Prosumer* $i \in \mathcal{I}$ **do**
  // Calculate proof score by following Algorithm (1).
  $\mathcal{PS}_i \leftarrow$ Execute Algorithm (1)
  Separate buyer ($\mathcal{I}_{buyer}$) and seller prosumers ($\mathcal{I}_{seller}$).
end for
//Creating initial coalitions.
while $\mathcal{I}_{seller} \neq \emptyset$ and $\mathcal{I}_{buyer} \neq \emptyset$ **do**
  $k = 0$
  // Create initial coalition $\mathcal{S}_k$ with neighbourhoods.
  if ($dist_{di} > dist_{ij}$ and $dist_{dj} > dist_{ij}$) **then**
    if ($i \in \mathcal{I}_{seller}$ and $j \in \mathcal{I}_{buyer}$) **then**
      $S_k \leftarrow S_k \cup \{i, j\}$
      $\mathcal{I}_{seller} \leftarrow \{\mathcal{I}_{seller} \setminus i\}$, $\mathcal{I}_{buyer} \leftarrow \{\mathcal{I}_{buyer} \setminus j\}$
    end if
  end if
  $\Pi_{init} \leftarrow \Pi_{init} \cup \{\mathcal{S}_k\}$, $\mathcal{S}_k \leftarrow \emptyset$, $k \leftarrow k+1$
end while
for *Each Coalition* $\mathcal{S} \in \Pi_{init}$ **do**
  // Optimize and calculate $\mathcal{PS}_{\mathcal{S}}$ of coaliton using equation (3.9) - (3.15).
  $OET_{\mathcal{S}} \leftarrow$ **OptimizeLoss** ($\mathcal{S}, \mathcal{P}, \mathcal{C}, \mathcal{DM}$)
  $\mathcal{PS}^*_{\mathcal{S}} \leftarrow$ Execute Algorithm (1), with equation (3.5).
end for
// Creating a stable partition.
while *not Converges* **do**
  Merge($U_{j=1}^{l}\mathcal{S}_j$), if { $U_{j=1}^{l}\mathcal{S}_j$ } $\triangleright$ $\{\mathcal{S}_1, \mathcal{S}_2, ..., \mathcal{S}_l\}$
  Split($\mathcal{S}$), if $\{\mathcal{S}_1, \mathcal{S}_2, ..., \mathcal{S}_l\}$ $\triangleright$ { $U_{j=1}^{l}\mathcal{S}_j$ }
  Calculate $OET_{\mathcal{S}}$, and $\mathcal{PS}^*_{\mathcal{S}}$
end while
$\mathcal{S}_{win}$ = Execute Algorithm(2) // Miner Selection
$\mathcal{M}_r = \delta_t * \sum_{\forall i \in \forall \mathcal{S} \in \Pi_{stable}} E_{\mathcal{S}_i}^{save}$
for *Each* $\mathcal{S}_{stable} \in \Pi_{stable}$ **do**
  Calculate coalition utility $U_{\mathcal{S}}$, and $\mathcal{M}_r$ using equation (3.17) to (3.21).
  for *each Prosumer* $i \in \mathcal{S}_{stable}$ **do**
    // Calculate prosumers payoff according to equation (3.3), and (3.22).
    $PayOff[i] =$ **CalculatePayoff**(i)
  end for
end for
$Blockchain_t \leftarrow Blockchain_{t-1} \cup Block_t$
$PayOff, Blockchain_t$

$$\mathcal{M}_r = \delta_t * \sum_{\forall i \in \forall S \in \Pi_{stable}} E^{save}_{S_i}, 0 < \delta \leq 1 \tag{3.23}$$

Here, δ_t is a scaling factor determined by the mutual agreement among the members of the miner coalition at each time t. The value of δ_t varies between 0 and 1, which means that a rational prosumer needs to pay part of its profit. Such a strategy is similar to any other conventional service provider where participating prosumers need to pay a fee to receive a service. However, instead of paying the fees from the prosumers' own pocket, the fee is part of their energy saving. Furthermore, if a prosumer i does not earn any profit ($E^{save}_{S_i}$=0), it does not require to pay anything. With such a system, parties do not lose anything, but their participation empowers their chances of increasing profit by receiving mining rewards. For simplicity of the model implementation, we start with a value of δ_t as 0.1. This portion of the total energy saving would be the mining reward, which will be divided fairly among the members of the winning coalition according to their contribution. Technically, the smart contract in the blockchain layer is responsible for distributing the reward as a form of energy credit transaction to the respective prosumers' digital wallets. In this blockchain layer, we also create a native token, EPOS (Energy proof score), to facilitate such transactions.

3.3.3 Mechanism Analysis

This section discusses essential proofs of some of the properties explained in the earlier section with respect to the proposed mechanism.

Lemma 1. *At time t, the final formed coalition from algorithm (3) is Pareto optimal and Stable.*

At time t in Algorithm (3), the decision of a prosumer $i \in I$ to form a coalition is coordinated by its increased proof score. For every round of merge and split, the optimized operation produces the highest coalition proof score for a particular round $\mathcal{PS}_S(\Pi_{itr_j}) = max[\mathcal{PS}_S]_{itr_j}$. This maximization is governed by Pareto rules, which means that a prosumer i would prefer to split from its current coalition and join into another as long as $\Pi_{itr_j} \triangleright \Pi_{itr_{j-1}}$ is satisfied. This means that in iteration itr_j the partition Π_{itr_j} is chosen by all prosumers in coalition $S \in \Pi_{itr_j}$ over that of the itr_{j-1}, if at least one of the prosumer could increase its proof score without hurting others. Hence, after the final iteration, itr_f, the proof score of prosumers would follow $\mathcal{PS}_{S_i}(\Pi_{itr_f}) \geq \mathcal{PS}_{K_i}(\Pi_{itr_{\{g\} \setminus f}}) \geq$

$\mathcal{PS}_i, \quad \forall i \in S, \quad \{g\} \setminus f \in itr_1, itr_2....itr_{f-1}$. This demonstrates that in the final iteration itr_f the proof score of the prosumer i, $PS_{S_i}(\Pi_{itr_f})$ is greater than or equal to all other iterations $(itr_{\{g\}\setminus f})$ proof scores $PS_{K_i}(\Pi_{itr_{\{g\}\setminus f}})$ and also its non-cooperative proof score PS_i. At this point, prosumers lose their motivation to merge and split further from the coalition, which is a part of the final partition Π_{itr_f}. Thus, the coalitions in the final partition achieve their stability.

Lemma 2. *An agreement of paying a part of the profit as a form of the mining reward, $\mathcal{M}_r(i)$ for a prosumer i is beneficial.*

In formula (3.23), we could infer that the $\mathcal{M}_r(i)$ is equal to a fraction δ of the energy savings $E_{S_i}^{save}$ of prosumer i at time t. The prosumer i requires to pay this amount as a form of transaction in the blockchain, and the amount could be determined by $\mathcal{M}_r(i) = \delta * E_{S_i}^{save}, i \in \mathcal{S} \in \Pi_{stable}, 0 < \delta \leq 1$. We have already proved that the algorithm is Pareto optimal when the partition structure is stable, Π_{stable} in the earlier section. Essentially, it follows, $PS_{S_i}(\Pi_{itr_f}) \geq PS_{K_i}(\Pi_{itr_{\{g\}\setminus f}}) \geq PS_i$, what it means is that the achieved proof score of prosumer i is higher than its singleton proof score. The amount of energy that represents the mining reward could be transformed to increase the proof score such that $\mathcal{M}_r(i) = \delta * [\mathcal{PS}_{S_i} - \mathcal{PS}_i], i \in S \in \Pi_{stable}, 0 < \delta \leq 1\ PS_i$. Hence, regardless of the value of δ, in the worst-case scenario, prosumer i will not lose any of its singleton proof score $\mathcal{PS}_i$. At this point, we could infer that a rational prosumer would prefer to join the coalition and agrees to pay the $\mathcal{M}_r(i)$.

Penalty Mechanism

There is a possibility that a prosumer may refuse to mine a block of transactions to disrupt the system even though it is selected as a miner at time t. So the obvious question is why the rational prosumer will disrupt the system? and what unfair advantage it may gain? This refusal strategy reduces the reputation of the system at stake when a counterparty initiates at the cost of more overhead so that the node can insert additional records in the block before mining. The prosumer may receive financial benefits from the third-party competitor.

To address this problem, we add a two-step mechanism to penalize such a malicious prosumer who declines to sign transactions and resists a block from being inserted in the blockchain. Firstly, the system slashed the locked-up stakes and withheld the reward. Secondly, the prosumers' coalition is removed as a miner, and the system will then select a new miner from the rest of the coalitions based on their relative weight by following Algorithm (2). This can be mathematically formulated as follows,

$$w\mathcal{P}^* \leftarrow \mathcal{P}^* \setminus \mathcal{P}_{\mathcal{S}_k}; \quad \text{if malicious prosumer} \quad k \in \mathcal{S}_k \tag{3.24}$$

3.3.4 Complexity Analysis

The complexity analysis answer the question "How long does it require to form a stable coalition structure with a number of prosumers considering their energy behavior at time t." In other words, how does the algorithm (1) and (2) scale when the number of prosumers increases in the network? To answer this question, we need to analyze the PoEG algorithm and coalition formation algorithm. In equations (3.25) and (3.26), we formulate the complexity analysis assuming that the number of prosumers in the network is n, coalitions in each iteration is k, and merge or split operation requires is l.

$$O(\text{PoEG}) = O(n) \tag{3.25}$$

$$O(\text{Coalition formation}) = O(n^2) + O(m^2kl) \tag{3.26}$$

If the number of prosumers in the network is n, the time complexity of the PoEG algorithm would be to compute the prosumers proof score for n times, which results in $O(n)$ formulated in equation (3.25). In the case of coalition formulation, firstly, the algorithm creates an initial coalition Π_{init} that will take $O(n^2)$. Secondly, every coalition optimizes its proof score with time complexity $O(m^2)$ where m denotes the number of the prosumer in each coalition, and the value will not be more than n. If there are k coalitions for a particular coalition structure, it will take $O(m^2k)$. Finally, this complexity continues to evolve for every coalition structure before the algorithm converges, and we assume that it requires l number of merge or split operations in total. Hence, the total time complexity to achieve a stable coalition would be $O(m^2kl)$.

3.3.5 Smart Meter Security

Smart meter experience security and privacy challenges, including data tampering and energy theft. The protection and safeguard of the smart meter data are based on tackling different types of attacks (e.g., doubles spending, transaction malleability, Sybil, etc. [121]). Although addressing these problems is not in the scope of our study, we assume that there are blockchain protection techniques to solve them. For example, the transaction

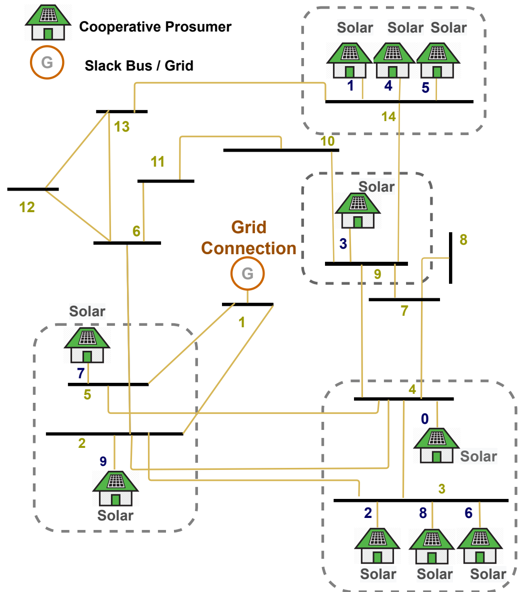


Figure 3.3: IEEE 14-bus system model. Ten (10) prosumers participate in coalition formation to maximize their PS.

malleability attack of a smart meter can be fortified by a consensus algorithm called proof-of-efficiency (PoEf). The mechanism considers current generation, rate, and consumption and then analyze with prosumers' previous history [122]. Also, the Sybil attack, which takes control of the blockchain network by 51% corrupted nodes [121], can be resisted by creating a unique timestamp block for verification and multi-signature based anonymous encrypted messaging streams as explained in [123], and [124]. Other mechanisms, such as those in [125], [126], and [127] can be used to protect smart meters from malicious attacks.

3.4 Model Evaluation and Discussion

We evaluate the proposed model using IEEE 14-bus system model given in Fig. 3.3. The IEEE 14-bus system represents a simple approximation of the US power grid system. A common approach to analyzing power system stability is to use a computing simulation because accessing the real-world distribution system proves challenging [128]. In this experiment, we employ the IEEE 14-bus system model to demonstrate the feasibility of integrating our model in real-world scenarios. Existing works such as [129], and [130] use the IEEE 14-bus test system in their experiments to show the accuracy of their proposed model. The IEEE 14-bus system is typically designed as a mesh network, which improves the reliability of a distribution system by providing multiple routes for power flow. However, constructing and maintaining this type of network is more complex and costly. In contrast, a radial network is a simpler network configuration where power flows via a single path or branch from a single source to multiple terminals, resembling a tree structure. However, as mentioned in [131], the future deployment of distributed networks will be able to deal with the inherited limitations of mesh networks thanks to the advancement of smart grid technology. To this end, we assumed a meshed distributed network in all our experiments. According to [132], future mesh networks are expected to provide better power reliability within an urban setting characterized by a high population density.

In our experiment, the entire simulation area is considered to be 20km x 20km, with the EDC in the middle. Prosumers are connected to the network nodes with the assistance of different buses, including 2, 3, 4, 5, 9, and 14, to transfer any amount of energy among all parties. We assume that the duration in each time slot of the algorithm (3) is 24 hours. The dataset used for implementing the proposed model is based on an open dataset published by Ausgrid, the largest electricity distributor on Australia's east coast [133]. The dataset consists of energy consumption and the generation of smart homes every half an hour. However, the generation from the installed rooftop solar panels is recorded from 7:30 to 17:00 hrs every day. Table 3.1 shows the details of the database and the description of its features. We use this time-series dataset to train a staked LSTM (Long Short-Term Memory Networks) learner [134] to forecast the energy consumption and generation for ten prosumers.

For the brevity of our proposed model's implementation, in our experiment, we use a variable energy loss with a range from 1.8% to 5.6% following the work in [108]. The energy transaction occurs between prosumers connected to different phases in an unbalanced distribution network similar to our proposed system model. We study another experiment based on a real-world distribution network based on Fujian Putian Power Supply Company [135], where the average energy loss is approximately 3.15% when the network consists of renewable generations with local consumption. For every experiment, we utilize the GUROBI optimizer to optimize the power transaction among prosumers to minimize the distribution loss.

3.4.1 Illustration of a 10 Prosumers Energy Network

The aim of this experiment is to analyze how prosumers behave in a simulated local distribution energy network that consists of 10 prosumers. In this setup, we use ten prosumers'

Table 3.1: Ausgrid dataset feature description.

Column	Field	Format / Description
1	Customer	Customer ID from 1 to 300
2	Postcode	Postcode location of the customer
3	Generator Capacity	Solar panel capacity recorded on the application for connection for each customer. Units are Kilowatt Peak (kWp), which is the solar panels peak power under full solar radiation.
4	Consumption Category	GC = General Consumption for electricity supplied all the time CL = Controlled Load Consumption (Off peak 1 or 2 tariffs) GG = Gross Generation for electricity generated by the solar system with a gross metering configuration.
5	Date	Date in DDMMMYYYY format
6 to 53	0:00 - 23:30	Kilowatt hours (kWh) of electrical energy consumed or generated in the half hour interval (eg. between 0:00 and 0:30).
54	Row Quality	(Blank) actual electricity recorded by the meter in the half hour NA = Estimates or substitutes of the electricity consumed or generated.

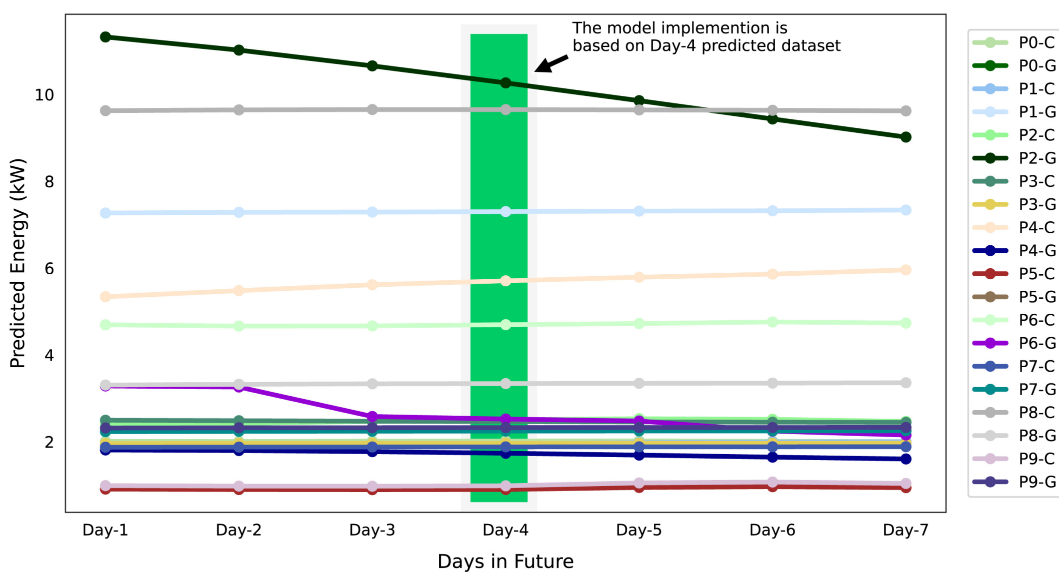


Figure 3.4: Ten prosumers energy forecast for 7 days generated from the stacked LSTM learner using Ausgrid dataset.

Table 3.2: Proof score distribution of coalitions and prosumer of a 10 prosumers energy network.

Itr	Partition	Coalition proof score	P0	P1	P2	P3	P4	P5	P6	P7	P8	P9
1	[[0], [1], [2], [3], [4], [5], [6], [7], [8], [9]]	[0, 3.22, 0, 4.12, 0, 0.88, 0, 2.67, 0, 0]	0	0	3.22	0	0	4.12	0	0.88	0	2.67
2	[[0, 7], [2, 8], [1], [5], [3, 9], [4], [6]]	[1.02, 3.35, 0, 4.12, 2.81, 0, 0]	0.01	0	3.27	0.07	0	4.12	0	0.89	0.08	2.74
3	[[0, 7, 1], [2, 8, 5, 6], [3, 9], [4]]	[1.04, 7.79, 2.81, 0]	0.08	0.07	3.28	0.07	0	4.22	0.18	0.89	0.11	2.74
4	[[0, 7, 1], [2, 8, 5, 6], [3, 9], [4]]	[1.07, 7.68, 2.84, 0]	0.08	0.07	3.28	0.07	0	4.22	0.18	0.89	0.11	2.74

energy forecast for seven days generated from the stacked LSTM learner using the Ausgrid dataset as shown in Fig. 3.4, where five prosumers have an energy demand, while others have an energy surplus.

Table 3.2 shows the proof score distribution of the coalitions for each iteration. At time t, after applying Algorithm (3), a total of four coalitions are formed. The algorithm converges after three merge and split operations. In the first iteration, the algorithm calculates its singleton proof score assuming no coalition formed. In the second iteration, the algorithm selects prosumers with the closest proximity, which yields the highest proof score of this round. Here, we observe that P_0, P_2, P_3, P_8, and P_9 improve their proof score without affecting other prosumers proof score. Liekwise, in the third iteration only prosumer P_0, P_1, P_5, and P_8 improve their proof score. As shown in table 3.2, in the fourth iteration, there were no further improvements observed, and the algorithm converges with four stable coalitions.

We show the stable coalitions, including their proof score, and the amount of P2P energy transactions among prosumers in coalitions and with EDC in Fig. 3.5. For example, Coalition-2 has the highest proof score. This is because the coalition initially reduces its loss by balancing prosumers' demand with the surplus energy produced by prosumer P_2 and P_5. Then, this coalition trades the rest of the energy with EDC.

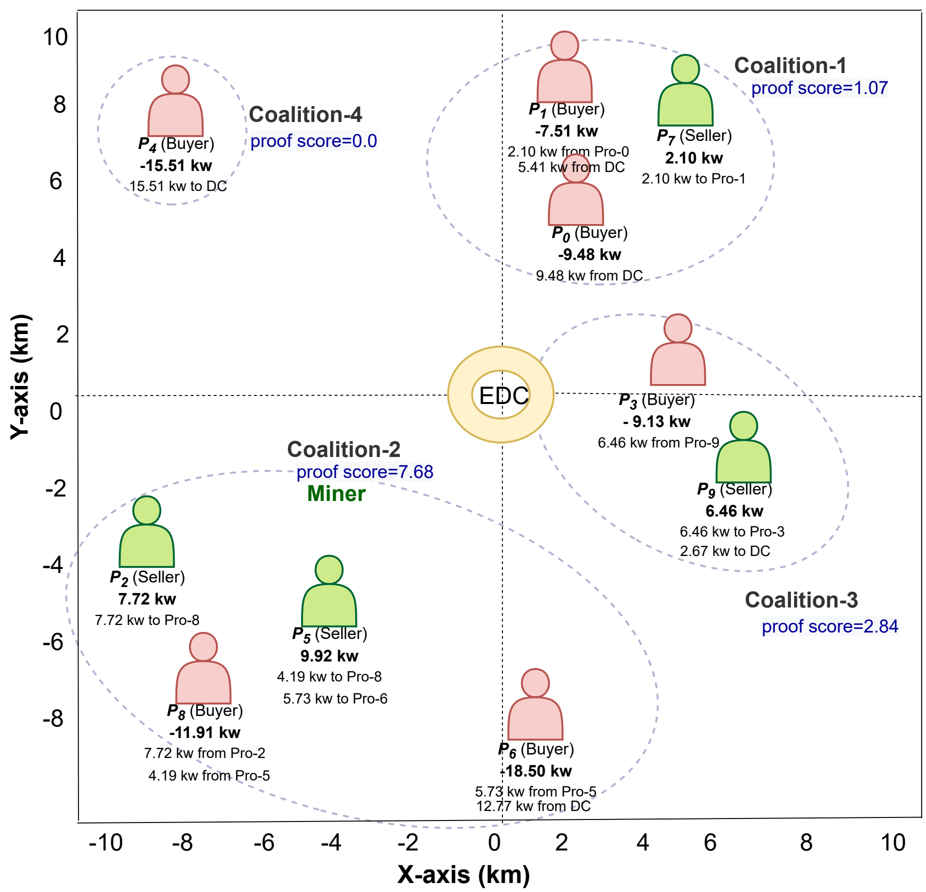


Figure 3.5: A stable coalition of 10-prosumers after employing the proposed algorithm at time t.

3.4.2 Payoff Analysis for a 10 Prosumers Energy Network

In Fig. 3.5, we observe that Coalition-2 achieves the highest proof score, which is approximately 7.68. Therefore, this coalition is selected as the miner to add the new block in the blockchain and receives $\mathcal{M}_r$ at time t. Fig. 3.6 shows the comparison of prosumers' energy loss by trading the same amount of energy associated with their cooperative and singleton behavior. Although one of the key criteria of our proposed algorithm is to minimize loss by creating coalitions, the loss of prosumer P_4 demonstrates to be constant. The reason for such a constant energy loss is that it trades all its demand with EDC, with whom the loss is minimum. In contrast, prosumer P_8 has approximately 5% loss reduction by joining Coalition-2. In the graph, the shaded area between the two lines represents the profit of

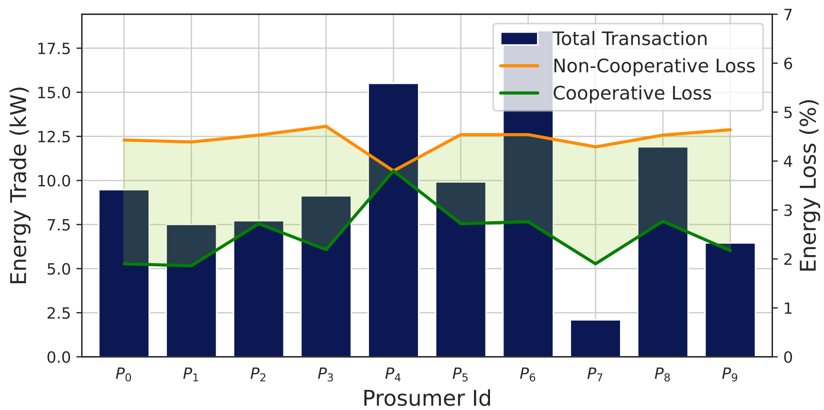


Figure 3.6: Prosumers' energy loss comparison between acting alone and cooperative behavior at time t.

prosumers considering their energy savings.

In Fig. 3.7, prosumers increase the payoff from their energy-saving and mining reward. However, only prosumers who are members of the winning coalition, P_2, P_5, P_6, P_8 receive the $\mathcal{M}_r$. We assumed that the value of the scaling factor δ is 0.1, which means that the M_r is 10% (δ) of the entire markets' energy savings (5.68kW) at time t. Every prosumer in the winning coalition further increases its payoff by receiving part of the $\mathcal{M}_r$. However, each prosumer receives a portion of the reward based on its contributed proof score in the winning coalition $\mathcal{S}_{win}$. For example, P_5 receives more reward (52% of $\mathcal{M}_r$) than P_8 (15% of $\mathcal{M}_r$) because P_5 contributes more proof score to $\mathcal{S}_{win}$. In contrast, prosumers in other coalitions receive zero mining rewards as they are not members of $\mathcal{S}_{win}$.

3.4.3 Stability Analysis of Coalitions

In this section, we analyze the stability of the final coalition formation. In this context, stability refers to the prosumer's motivation for not moving to another coalition for more benefit. Fig. 3.8 shows the probable critical outcomes of each action performed by prosumer P_2 when it joins a cooperative game and becomes a part of *Col-3* in the final partition Π_{stable}. We observe that initially when all the prosumers are acting alone, P_2 has the proof score of 3.28. In the proposed algorithm, we show that prosumers' proof scores started increasing when they chose to work in coalitions. Therefore, in the final coalition, each prosumer, including P_2 enhances their proof score compared to its initial and subsequent

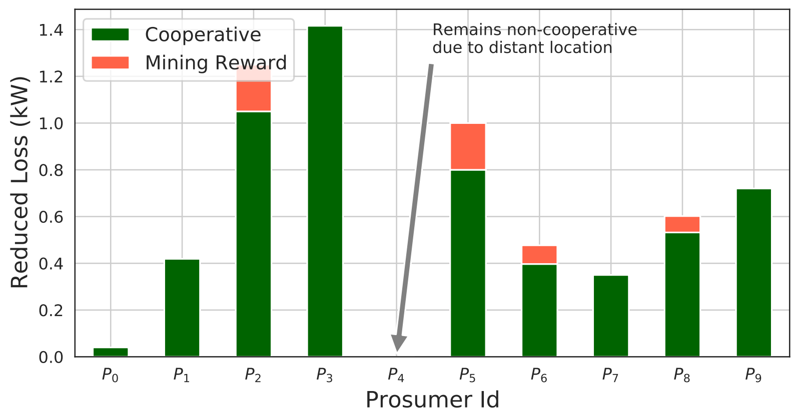


Figure 3.7: Prosumers increased payoff. Factors that attributed to this increased payoff include saving energy and mining reward. Here, S_{win} consists of prosumer P_2, P_5, P_6, and P_8.

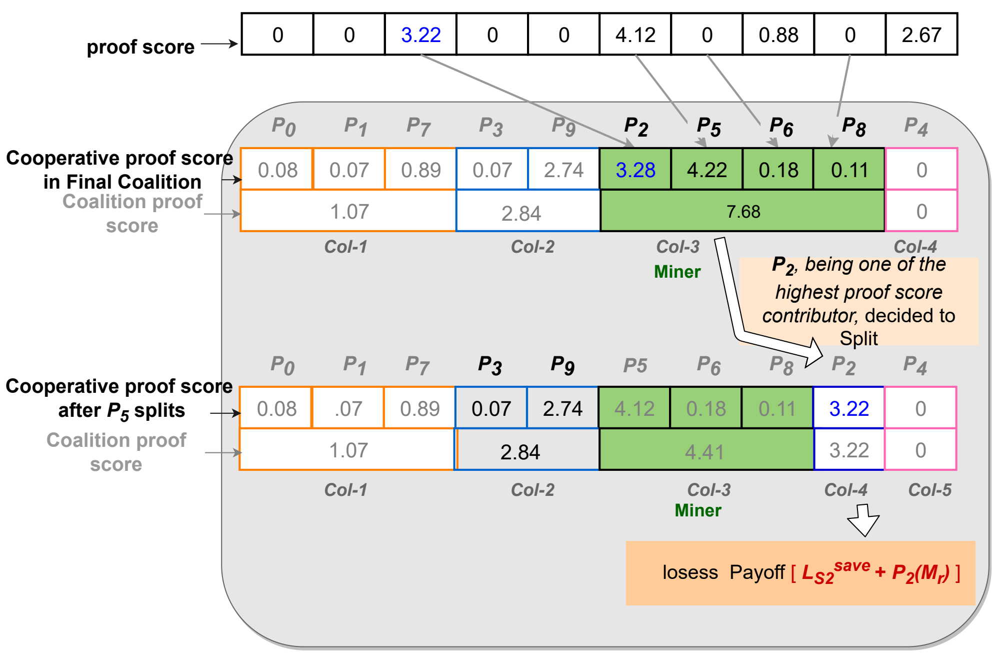


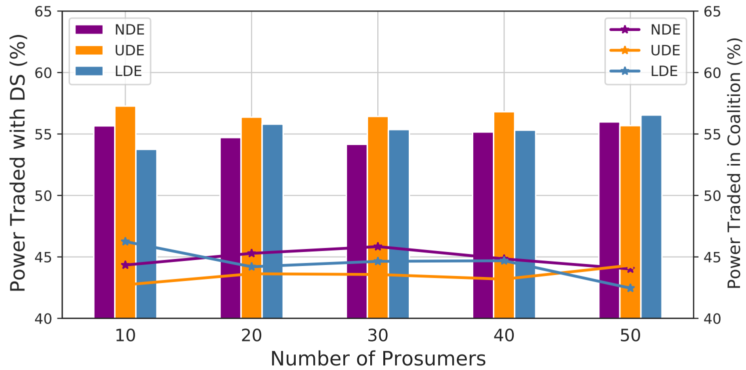
Figure 3.8: Payoff analysis of a rational prosumer P_5 after it splits from the final stable coalition at time t.

iterations of merge and split. Also, P_2 is part of the $\mathcal{S}_{win}$, therefore, this prosumer increases its proof score (0.06), which will reduce overall distribution loss $E^{loss}_{S_2}$ and receives mining reward $P_2(\mathcal{M}_r)$. If P_2 chooses to split from $\mathcal{S}_{win}$ to receive a better payoff by either acting alone or joining another coalition, we show that no other coalition is better off for P_2 because all other coalition members are far from P_2 except EDC. As a result, the impact is twofold, firstly, P_2 loses its leading proof score, and secondly, it does not receive any mining reward $P_2(\mathcal{M}_r)$. Therefore, a rational prosumer P_2 will choose to remain in *Col-3*. At this point, although *Col-3* does not lose its supremacy, other prosumers in this coalition observe a reduced proof score. Therefore, prosumer P_2's behavior affects not only its own payoff but also other prosumers' proof score in *Col-3* significantly.

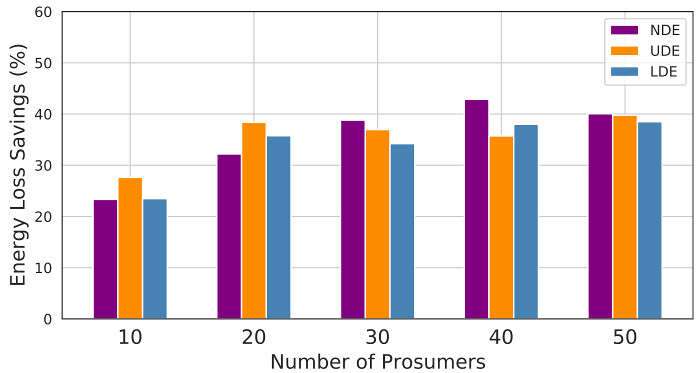
3.4.4 Model Analysis with Increased Number of Prosumers

In the following experiments, we increased the number of prosumers by 20, 30, 40, and 50. In every test, prosumers data are generated with the same simulation parameters as described in section 5. However, the energy profile is randomly generated by using the daily mean household consumption (12.41kWh) and generation (7.45kWh) of the Ausgrid Dataset [133]. For the entire experiment, we use this mean with the standard deviation of 8.0kWh as calculated from [133] to generate random datasets considering the distribution line loss with the range from 1.8% to 5.6% as in [108]. Moreover, the energy load profiles were randomly generated according to three distributions, namely Normal distribution energy (NDE), Uniform distribution energy (UDE), and Log-normal distribution energy (LDE).

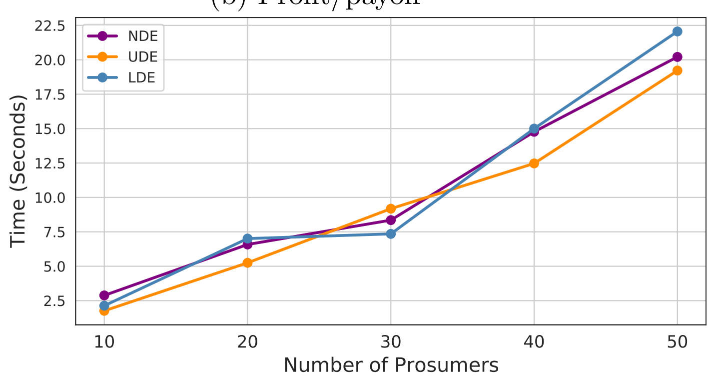
Fig. 3.9-(a) shows the ratio of energy traded within coalitions and EDC for all five energy networks using three data distributions. We observe that the proportion of power traded within the coalition for UDE is higher than the other two distributions. However, the results reveal that there are marginal differences among all energy networks. This is mainly because the prosumers first balance their energy with their coalition members, then trade the rest of the energy with EDC. On average, approximately 47% of each of the entire networks' energy has been traded among prosumers, thus avoiding costly trade with EDC. Fig. 3.9-(b) depicts the percentage of energy loss that is being saved by the cooperative prosumers compared to the same prosumers acting alone. When the number of prosumers in the network was 10, energy saving was at 22.5%. However, when the number of prosumers increases, energy-saving reaches (38%) for 50 prosumers. The main reason for this trend is that as the number of prosumers in the network increases, they will have a higher chance of exchanging power with prosumers in close proximity and fewer chances



(a) Energy Transaction (Tx)



(b) Profit/payoff



(c) Execution Time

Figure 3.9: (a) Energy transaction comparison with EDC and coalition prosumers (b) Profit/Payoff by saving energy loss in the cooperative game compare to acting alone (c) Model Execution time.

to exchange energy with the EDC. Therefore, it is evident that the increased number of prosumers provides better benefits for the entire P2P energy market.

To show the execution time of our algorithm, we performed the experiments by using Python programming on MS Windows Intel Xeon(R) 3.20 GHz processor with 8GB RAM. Fig. 3.9-(c) shows that execution time is linearly increasing with the increasing number of prosumers in the network. However, the proposed model assumes that the local computing system of prosumers executes the coalition formation algorithm(3) and stores the blockchain. As the number of prosumers grows, so does the number of mining nodes, potentially resulting in transaction delays. This is due to the requirement of broadcasting a new block to all mining nodes. This may have an impact on the scalability of the system, but with the advancement of the lightning network for blockchain, the speed of executing transactions is expected to enhance significantly. As a result, the scalability of the system will increase as the number of participants increases [136].

3.4.5 Blockchain Layer Implementation

In the blockchain layer, we implemented the open-source Avalanche blockchain platform [58], which enables system customization. The blockchain consists of X-chain, P-chain, and C-chain. Fundamentally, X-chain is the blueprint and an instance of the Avalanche virtual machine used for exchanging assets. The C-chain is the smart contract chain which is an Ethereum VM, therefore, the platform supports the Ethereum toolkit and contract applications written in Solidity. The P-chain is the administration chain that allows users to stake a certain amount of assets to become validators or delegators and get rewards from it.

Fig. 3.10 shows the details of the model implementation in the Avalanche blockchain. We created a blockchain sandbox that contains five Avalanche computing nodes. Any five prosumers in the energy market may either set up a single node by using their existing low-end computing hardware or rent cheap cloud resources. Then, we created wallets for each of the ten prosumers, which can be accessed by simply using HTTP protocol coordinated by an application called Avax Wallet. At time t, the prosumer in a coalition which has the highest proof score would be added as a potential validator of a particular node. The rest of the prosumers are set as a delegator who have the freedom to choose a validating node with whom they want to stake their proof score. Furthermore, we created a native token, namely Energy proof score (EPOS), in the blockchain network. The proof score of prosumers generated from Algorithm (3) are credited as EPOS to the respective prosumers' X-chain address in the wallet. The proof score is then transferred to the P-chain for staking,

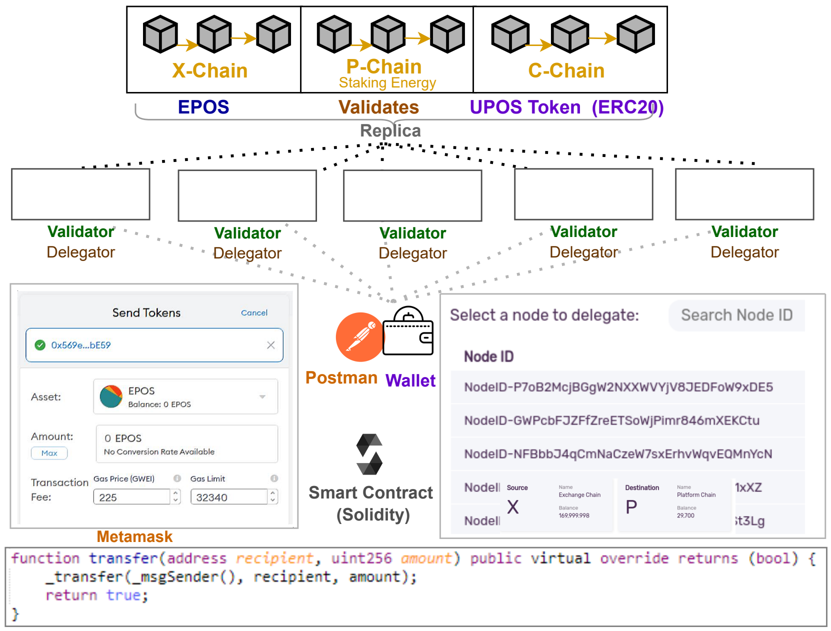


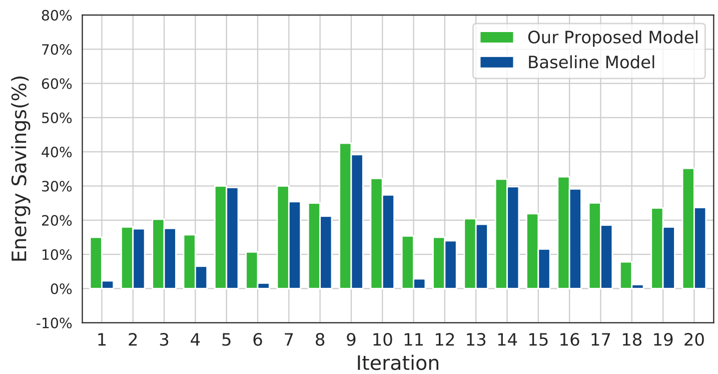
Figure 3.10: Blockchain Layer implementation (Avalanche).

similar to the PoEG consensus process. Once all the energy transaction in the physical layer is processed, the prosumers' mining reward is then distributed to the validators and delegators based on the proof score they staked. Finally, we created a UPOS token in the C-chain following Ethereum's ERC20 standard that unlocks assets interoperability on the blockchain layer and makes our energy market self-sufficient with the ability to sell the surplus energy to other autonomous energy markets.

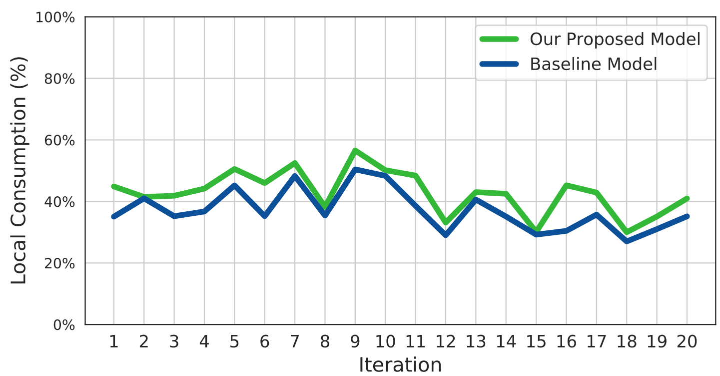
3.4.6 Comparative Experiments

In this section, we present some of the key criteria to demonstrate the difference between ours and the existing works. We consider the work in [7] as a benchmark that is closely related to our proposed system. This is because they propose a decentralized blockchain-based energy management system for residential prosumers within a microgrid to reduce its peak hour energy usage. Also, they use game theory to maximize the utility of individual users.

Fig. 3.11 (a) compares the energy savings between the baseline scenario and our proposed model for 20 iterations. In each iteration, we generate the dataset based on the



(a) Energy Savings



(b) Local Consumption

Figure 3.11: (a) Energy-saving comparison between our model and the baseline model. (b) The comparative energy consumption by prosumers in the trading system.

prediction model as discussed in section 3.4 and using parameters as in Table 3.1. Then, we use this dataset for executing our proposed algorithm (3) as well as the baseline model [7]. In this case, the energy network follows the same parameters as in Table 3.1. The energy savings in our model are higher (by an average of 6.5%) than the baseline model considering all the experiments assuming both of the networks consist of the same number of prosumers. The reason is that prosumers in our model cooperate and trade among themselves to reduce line loss, while prosumers in the baseline model trade energy non-cooperatively. On the other hand, Fig. 3.11 (b) shows the percentage of energy that is consumed by users by locally produced renewable energy produced by other users in the trading system. The experimental results demonstrate that our proposed model performs well regarding self-consumption as a higher amount of energy is consumed by neighboring prosumers with cooperation.

To compare the throughput, we consider three criteria: block insertion time, transaction per second, and safety threshold. The proposed system has a block insertion time of at most 2 seconds, while the baseline requires approximately 35 seconds. With respect to the transactions per second (tps), our model can scale much higher because the tps increases when the number of subnets increases compared to the baseline (max. tps $\geq$ 2000) models. Finally, our proposed model shows a higher security threshold as PoEG in the Avalanche platform requires at least 80% nodes to be corrupted to alter information in the blockchain [137]. However, in Fig. 3.11 (a), we notice that the proportion of energy savings in the proposed model, compared to the baseline model, is vastly different in some iterations. This is because the energy savings in our system depends on the coalition structure, which contributes to the difference in local consumption between our proposed model and the baseline model, as shown in Fig. 3.11 (b).

3.5 Summary

In this chapter, we consider blockchain-based P2P energy trading and coalition formation among decentralized prosumers. The proposed model provides the foundation for future integration of RES and trading within communities using a secure and transparent infrastructure. The proposed PoEG provides a feasible model for the prosumers to reduce distribution losses and increase payoff. We show that the final coalition formation algorithm converges to a stable structure of rational prosumers who strive to be elected as miners in the underlying blockchain. Also, we implement the mechanism on the industry-standard Avalanche blockchain platform to represent its potential for real-world implementation.

Chapter 4

Blockchain and Federated Reinforcement Learning based Vehicle-to-Everything Energy Trading System

In this chapter, we proposed a system that combines three different technologies: blockchain, Federated reinforcement learning (FRL), and vehicle-to-everything (V2X) communication. The system's goal is to enable energy trading between Electric vehicles (EVs) and the power grid through the use of a decentralized, secure, and intelligent platform. In section 4.1, we discuss the background and existing challenges of proposed systems. Then, section 4.2 discusses the V2V energy trading model based on blockchain, federated reinforcement learning, and the PoSOC mining protocol. In section 4.3, we represent the blockchain-based FRL system for V2V Energy Trading followed by the system performance analysis and results in section 4.4. Finally, the summary is discussed in section 4.5.

4.1 Introduction

Vehicle-to-Everything (V2X) is an emerging paradigm where EVs equipped with bidirectional charging technology can exchange energy with grids (V2G) [28], with buildings (V2B) [138], and with other EVs (V2V) [28]. The exchange of energy in a V2X scheme can occur anytime and anywhere and thus provides more flexibility and benefits to EV owners

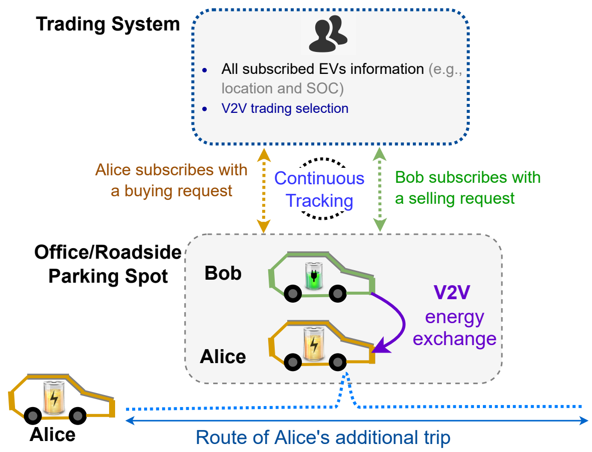


Figure 4.1: Example of a V2V energy trading scenario.

[11]. This study focuses on one aspect of V2X, namely V2V. Specifically, we propose a novel Federated Reinforcement Learning (FRL) system combined with blockchain technology to maximize EV users' utility while preserving the security and privacy of trading transactions in V2V.

Currently, most popular EVs have a range of 53 to 270 miles per charge [139], which allows a user to travel 2 to 6 times an average commuting distance [140]. However, when users are outside of their normal commuting distance and need to visit additional places, the EV may require an extra charge along the way. The immediate solution is to visit a Charging Station (CS) within proximity, but there are situations when CSs are not available in the area or existing CSs are booked by other EVs leading to an undesirable waiting time. In such a situation, the user may look for an alternative charging means such as V2V charging which allows one EV to charge another EV using appropriate charging/discharging hardware technologies [141]. In figure 4.1, for example, if Alice is traveling beyond her daily commute and needs to charge her EV, then a V2V application may match Alice with Bob who happened to have a fully charged EV nearby, given that Bob is willing to sell some of his battery's energy to Alice. For such a match to succeed, Alice and Bob must be subscribers of a V2V system and both must permit the application to track their trip and EV information (e.g., location, State of Charge (SOC)). Then, Alice and Bob can meet at a certain location and perform the V2V energy exchange. The above example shows the benefits of V2V energy trading systems, but in reality, there are many challenges related to privacy, security, and energy trading strategies that need to be addressed to incentivize users to participate.

4.1.1 Challenges and Motivations

In the following, we include some of the technical challenges and the motivations for solving them.

1. **Trust:** The trading participants, Alice and Bob, do not know each other, and there is no implicit trust between them. The challenge here is to design a secure trading system that allows them to interact without necessarily trusting each other. To address this challenge, we propose the use of blockchain technology because it creates a distributed ledger (database) where each transaction is validated through consensus between multiple parties. Therefore, Alice and Bob are validated before being engaged in energy trading. Furthermore, the transactions can be verified without the need for a trusted intermediary (e.g., a broker or bank).

2. **Selfish user:** In the V2X trading system, some users may be inclined to avoid paying their fees, which presents a challenge for the system. To address this challenge, we propose the use of a self-executing contract called a smart contract on the blockchain. This contract serves as an agreement between electric vehicle (EV) users and the V2X trading system, which must be adhered to by all parties at the time of subscription. In practical terms, the smart contract will confiscate any profit earned by users from their energy trading transactions for a designated period of time until the fee is paid in full. This measure will ensure that every user contributes their agreed share to the systems' success and sustainability.

3. **FRL:** To maximize the utility of an EV (agent) within the V2X trading system, an intelligent mechanism must be in place to enable an optimal decision-making strategy. This strategy must consider a range of factors, including the various locations of the EVs and the trading prices. However, this is a challenging task as the agents may lack previous trade data to inform their decision-making process. To address this issue, we propose the use of FRL for two main reasons. Firstly, FRL is a trial-and-error process that allows users to adopt the best action to take with minimum or no prior experience. Secondly, the model can learn collaboratively and adapt to the dynamic environment without necessarily sharing users' data with others.

4. **Computing Complexity with FRL:** Training a constantly evolving intelligent system is costly in terms of computing and communication overhead. The current model-based methods such as linear regression employ deductive procedures where an overall comprehension of the system is necessary to get an ideal result. However, these approaches cannot reduce the overhead because data must be transmitted from

all distributed sources to a central server for efficient processing [11]. To address this issue, FRL would be a suitable technology to consider as it enables distributed learning without explicitly transferring the data from all the local nodes to a central server.

4.1.2 Major Limitations and Proposed Approach

The typical approach in the literature to solve an EVs' optimal energy transaction problem is to use constrained optimization techniques such as mixed-integer programming (MIP) and deterministic and stochastic model predictive control (MPC) [79, 142]. MPC has problems because it is model-based (and therefore not end-to-end) and requires expensive computation because it solves an entire optimization problem in each timestep. On the contrary, data-driven approaches like Deep Neural Networks (DNNs) and Deep Reinforcement Learning (DRL) have emerged as promising alternatives to handle these problems [143, 144]. However, DNN-based techniques require excessive data and extensive computation for training. To address these issues, FRL enables decentralized training and sharing of a tiny amount of data without expensive communication overhead [145]. Technically, each FRL agent communicates its local updates to a centralized server, referred to as a Local aggregator (LAG). The LAG combines the updates from each FRL agent and responds back with the updated global model. However, in our situation, we identify a key issue with conventional FRL. Specifically, FRL requires a centralized aggregator to combine the local updates, which may lead to a single point of failure. Furthermore, malicious users may violate users' privacy and transaction security which may hinder the potential of attracting EV users to a V2V trading platform. This is a core reliability issue because it may compromise transaction authentication and payment services. To address this issue, the studies in [12], [146], and [89] use a combination of blockchain and FRL to perform on-device training collaboration with a centralized aggregator for different use-cases. The focus of those studies, however, is restricted to enhancing the security of the system only.

In contrast to the existing works, we address the above gaps by proposing a novel V2V energy trading system that combines blockchain and FRL to secure transactions, protect user privacy, enhance trust, maximize economic benefit, and reduce communication overhead. Specifically, we address the security, privacy, and trust issues, by using blockchain technology to enable EV users to store energy as an asset, trade with untrusted parties, and enhance the models' dynamic data protection. In addition, we propose PoSOC consensus protocol that renders a score by considering EVs SOC and energy price (On-peak, Mid-peak, Off-peak) for each timestep (e.g., 1 hour). An EV achieves a higher score when selling energy during On-Peak and Mid-Peak hours compared to the Off-peak hour. The total score is calculated for each timestep for an entire episode consisting of 24 timesteps. The score of individual EVs is staked to the blockchain for users to access and determine miners. The proposed FRL allows the training of each agent individually using locally collected data. Then, each FRL agent shares the model's updated parameters with the LAG for parameter aggregation. Consequently, the benefit of using such a technique is that it enables the evolution of the RL model without privacy leakage and reduces data communication overhead. The conventional FRL systems consider LAG as a single node that is vulnerable to cyber-attacks and SPOF. Hence, we propose a decentralized LAG where every node has the chance to become a mining LAG and validate and insert a block to the blockchain.

4.1.3 Key Contributions

The integration of blockchain and FRL as a means of empowering all EV participants in an energy trading system using V2V schemes is a relatively unexplored field. To this end, the main contributions of this study are:

1. We propose a novel blockchain-based V2V energy trading mechanism with FRL to enable EV users to select optimal energy trading strategies for maximizing trading benefits while protecting security, privacy, trust, and transparency.

2. We design the Proof of State of Charge (PoSOC) consensus protocol as a function of EVs SOC and the amount of energy traded during peak hours. The mechanism selects a miner for a specific interval which ensures transaction validity and transparency. The significance of this protocol is that EV users in the system are motivated to sell energy more during peak hours to maximize their chances of being selected as miners and receiving rewards.

3. We propose the approach of decentralized LAGs where every node has the chance to become a mining LAG (mLag) and validate and insert a block to the blockchain. This approach prevents the system from SPOF and malicious attacks. Also, we propose a mining reward mechanism using the smart contract technique to deter the selfish behavior of users.

4.2 System Model

The proposed system model is illustrated in Figure 4.2. This model assumes that Alice and Bob met at Bob's working location, where he parks his EV. The topology that governs the trading system in this model focuses on V2V and V2B. This is a realistic setup of known real-life projects such as Fermata Energy, Boulder in the USA [147]. In the proposed model, an EV may decide to sell its surplus energy to other participants at a higher price during peak hours and refill its battery at times when the electricity price is lower. The model consists of three layers: the optimization and distributed learning layer, the security layer, and the physical layer. The optimization and distributed learning layer enable EVs to analyze their expected profit according to the current SOC and probable future actions in the environment. The preference of an agent (each EV) is governed by the intelligence of the system considering several environmental parameters, such as time of the day, price, availability, etc., to maximize its profit. The model uses an FRL scheme to train the agents collaboratively for EVs to adopt optimal trading strategies. The process of learning is coordinated by a LAG, which is responsible for combining the learning updates (gradients) received from the local learners, and then the LAG broadcasts the updated global model back to each of the agents.

In the security layer, the blockchain uses a consensus mechanism to certify a block of energy transaction data before replicating it to all the decentralized computing nodes in the network to prevent potential counterfeit. Mining a block of data allows a selected EV to verify each transaction and insert it into a new block in the existing blockchain. Conventionally, the chances of becoming a miner depend on the participants' computational capability (e.g., PoW) or the number of assets they stake (e.g., PoS). This thesis proposes an innovative PoSOC protocol to select a user to be the miner. Essentially, the algorithm is a ranking mechanism to determine which EV will be selected as a miner. The higher the attained proof score ($\mathcal{PS}$), the higher the chance for an EV to become the miner at time t and becomes mLAG. In this sense, our system is decentralized since no central aggregator is needed. Finally, once the energy transactions in the system are settled, the physical layer is responsible for supporting the physical transmission of energy among participants in the network.

4.2.1 Proof of State of Charge (PoSOC)

The fundamental concept of PoSOC is that EV users who discharge more than they have charged in a time period will have a higher chance of becoming the mLAG. This behavior

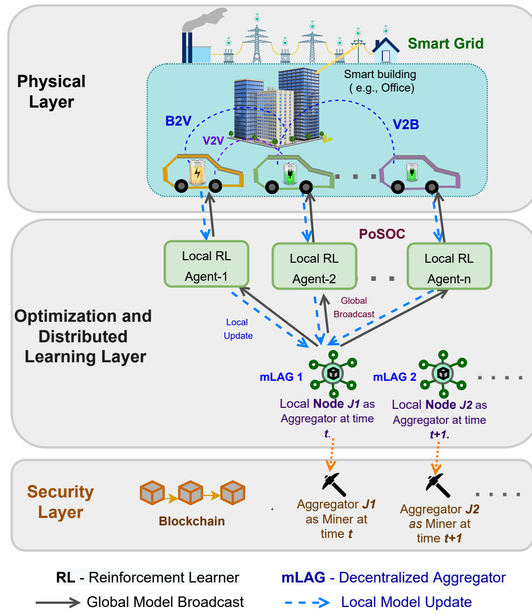


Figure 4.2: EV based V2V energy trading network with blockchain and Federated Reinforcement Learning.

is quantified by a function of the amount of electricity traded and the SOC of an EV [148]. Let $\mathcal{I}$ be the set of all participating EVs such that $\mathcal{I} = \{\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3, ..., \mathcal{I}_n\}$. At time t, $\mathcal{D} = \{\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, ..., \mathcal{D}_n\}$ denotes the amount of energy discharged while $\mathcal{C} = \{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3, ..., \mathcal{C}_n\}$ represents the amount of energy charged for each EV. The state of charge of EVs can be denoted as $SOC = \{SOC_1, SOC_2, SOC_3, ..., SOC_n\}$ and their corresponding maximum SOC is $SOC^{max} = \{SOC_1^{max}, SOC_2^{max}, SOC_3^{max}, ..., SOC_n^{max}\}$. Hence, the proof score $\mathcal{PS}_i$ of an EV i can be mathematically represented as,

$$\mathcal{PS}_i = \begin{cases} (SOC_i^{max} - SOC_i) + \delta(\mathcal{D}_i - \mathcal{C}_i); & \text{if departing} \\ \delta(\mathcal{D}_i - \mathcal{C}_i) & \text{otherwise} \end{cases} \tag{4.1}$$

$$\delta \in (\delta_{mid}, \delta_{on}); \quad 0 < \delta < 1 \quad \text{and} \quad \delta_{on} > \delta_{mid} > \delta_{off} \tag{4.2}$$

In equation (4.1), the proof score of an EV i depends on its SOC and the amount of energy it charges or discharges when it departs from its current location. The component $(SOC_i^{max} - SOC_i)$ represents the portion of the battery of EV i kept empty when the EV is expected to depart and will not participate in the charging/discharging process anymore.

Therefore, to maximize the above component and increase the proof score, users need to sell their extra energy as much as possible before departing. On the other hand, the part $\delta(\mathcal{D}_i - \mathcal{C}_i)$ represents how the trading amount influences the attained proof score. Particularly, the EVs action of charging reduces $\mathcal{PS}_i$, while discharging will increase $\mathcal{PS}_i$. Therefore, the mechanism motivates EV users to sell more and buy less energy at peak hours. This motivation is magnified in the severity by δ, which will be greater during periods of higher energy prices. The dynamics of δ can be demonstrated with the variations of energy price at On-peak (λ_{on}) or Mid-peak (λ_{mid}). This could be mathematically formulated by the following equations,

$$\delta_{mid} = \left[\frac{\lambda_{mid} - \lambda_{off}}{\lambda_{off}}\right] \tag{4.3}$$

$$\delta_{on} = \left[\frac{\lambda_{on} - \lambda_{off}}{\lambda_{off}}\right] \tag{4.4}$$

In equation (4.3) and (4.4), δ_{mid}, and δ_{on} represents the δ value at Mid-peak and On-peak hours, respectively. Essentially, the values are calculated by the increase in per unit of electricity at peak hours relative to their Off-peak hour baseline.

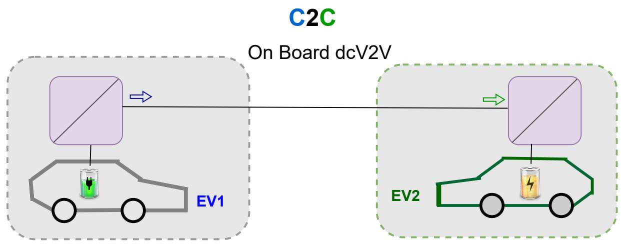
4.2.2 Decentralize Stake Control

In the proposed model, an EVs' computing node consists of a Control Module (CM) and blockchain. The task of the CM is to supervise the execution of the PoSOC protocol, store the proof score records on the blockchain, and control its access. Initially, at time t, every EV executes an instance of PoSOC protocol and shares the proof score $\mathcal{PS}_i$ of EV i to the existing blockchain instead of sharing their sensitive charging/discharging information. Moreover, every EV (FRL Agent) share their local learning updates to the blockchain. Next, the mLAG, essentially an FRL agent at time t, fetches all the information, aggregates the learning updates, and broadcasts back to all the nodes through the blockchain. Finally, a weighted random selection (WRS) technique is used based on FRL agents' attained proof score to select the miner mLAG for time slot $t+1$. At the time $t+1$, the control of the CM module is transferred to the node mLAG selected at time t. The proposed model repeats the above steps to execute the CM module at time $t+1$, to select mLAG for time $t+2$.

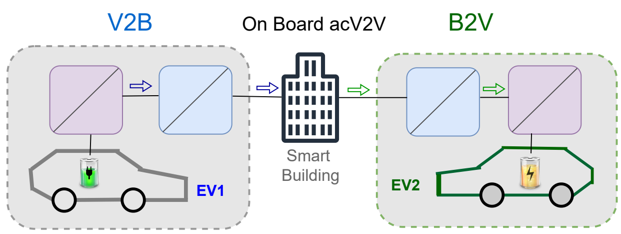
Blockchain enables users to audit all the transactions, which is the key to preserving the transparency of EV users and building trust in the proposed system. The public and private key encryption techniques in the blockchain can obscure the true identity of an EV user. However, it is demonstrated that complete anonymity can not be ensured in the blockchain due to transaction linkability, private-keys management, and recovery issues [149]. For example, one can link all public key addresses to a user by keeping track of the history of the transaction graph in the blockchain. One way to address this issue is by connecting to another overlay network first, for example, TOR(onion routing) or I2P [150] used by Moreno blockchain. Moreover, Zero-knowledge proof (ZKP) and Zero-knowledge Succinct Non-interactive Arguments of Knowledge (ZK-SNARKS) provide full anonymity [151]. However, they are not efficient for responsive scenarios because of the high computational time. On the other hand, users' privacy could be enhanced by employing compound identity [87], which generates multiple asymmetric key-pair related to the owners, storage, and dataset. This could be a potential research direction that is not in the scope of our work. However, interested readers can refer to [152, 87] and [149] for detailed techniques and their relevant explanations.

4.2.3 V2V Topology

Traditionally, the concept of V2V combines the operating mechanism of V2G and G2V to transfer energy when EVs are connected to a local grid using a bidirectional charger [26]. In this scheme, an Electric Vehicle Aggregator (EAG) controls the entire energy transfer and the grid performs as a third-party infrastructure provider. Furthermore, with the combination of V2H and G2V, V2V could be achieved. On the other hand, it is possible to perform a direct V2V power transfer by utilizing various interfaces as in [27] to reduce the number of dc-ac and ac-dc conversions. This could be attained by utilizing dc-dc converters using CHaedeMO ports that permit bidirectional current flow [141]. Inspired by the aforementioned topologies, we assume that our proposed system model achieves V2V by transferring energy directly as in Fig. 4.3-(a), or a combination of V2B and B2V with a local building energy network as in Fig. 4.3-(b). Hence, the proposed system can integrate a variety of EVs with distinct types of charging techniques, which provides seamless energy trading services to all participants. Furthermore, Dynamic Wireless Power Transfer (DWPT) technology has emerged as an appealing solution that enables power transfer when the EV is on driving with the help of a roadside charging pad. For the brevity of model implementation, we do not consider wireless charging using third-party vendors. The interested reader can check [26] and [153] for more EV charging options and technologies for potential future research direction.



(a) Direct power transfer.



(b) Indirect power transfer.

Figure 4.3: (a) Power transfer using an external dc-dc converter. (b) Power transfer with the combination of V2B and B2V mode.

4.2.4 Preliminaries of the FRL Model

Markov Decision Process (MDP) Model

To increase the profit of an EV, we aim to optimize the electricity trading amount, schedule, and location. An EV can either buy energy from other EVs or from the main grid. As shown in Fig. 4.4, the EV could be in four different states: driving, charging, discharging, and simply observing the environment. Let us assume that the amount of energy required for charging the EV i is $E_{t,i}^{ch}$ at a particular time t. This demand could be balanced from three sources: Local Energy Market ($E_{t,i}^{V2H}$), other EVs ($E_{t,i}^{V2V}$), and main grid ($E_{t,i}^{V2G}$). The availability of energy from those sources, their price, and the time of the day are a few parameters that could significantly influence the EV's trading decision.

In reinforcement learning (RL), an agent, i.e., EV, learns optimal policies (through numerous trials and errors) by observing state parameters from the environment and selects the next action to maximize its expected return. Fundamentally, the core of RL is a mathematical framework for a sequential decision-making system known as the Markov Decision Process (MDP) which could be represented by a tuple $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \mathcal{T})$. Here, $\mathcal{S}, \mathcal{A}$ represents the finite set of states and actions, respectively [85]. $\mathcal{P}$ is a transition

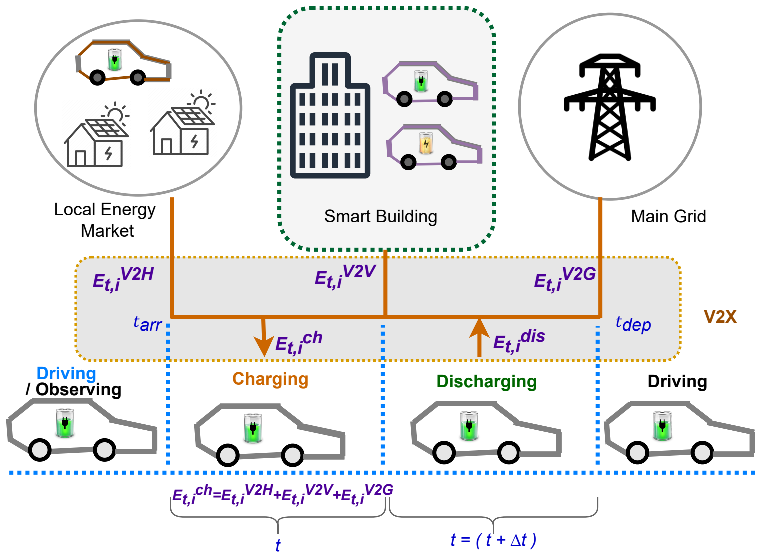


Figure 4.4: The mechanism of an EV for a specific charging/discharging control period.

probability function that maps an agent's current state s_t and action a_t into a probability of transitioning into state s_{t+1}. The function $\mathcal{R}$ is a reward function that represents the benefit of an agent when moving from state s_t to state s_{t+1} by following the action a_t. Also, $\mathcal{T}$ represents the finite time interval.

Notably, Q learning is one approach of model-free RL mechanisms [154], as opposed to the model-based approaches discussed in Section I, where the policy relies on the expected return represented by Q-value. Here, a static weight γ known as discounted factor ensures the preference of a recent event compared to future events in terms of receiving rewards. The value determines the optimality (goodness) of an action using the Bellman equation as follows,

$$Q^{\pi}(s_t, a_t) = E\left[R_{t+1} + \gamma \max_{a'} Q^{\pi}(s_{t+1}, a^{'})\right]; \quad \gamma \in [0, 1] \tag{4.5}$$

The main objective of a Q learning algorithm is to identify an optimal policy π^*_θ by maximizing its Q-value by playing actions using the lookup table as follows,

$$\pi^*_\theta(a_t|s_t) = \arg\max_{a_t} Q(s_t, a_t) \tag{4.6}$$

One of the challenges of Q learning is that it does not follow the interaction sequence when computing the Q values [154]. Rather it selects the highest Q value for its next time step, which leads to an overestimation. Q learning performs well as long as the state and action space is discrete. To solve a continuous action spaces problem similar to the V2V energy trading, the action spaces required to be discretized which leads to the curse of dimensionality [155]. Also, it can not perform well when the discretization range is too large because it may lead to unaccepted control output. Therefore, the computing complexity increases exponentially when the dimension of the state space or action space increases, which is not an ideal technology to consider for a V2V energy trading system. This problem could be resolved by using policy gradient-based methods because they use probability density function on actions which can be either continuous or discrete [156]. The actor-critic method is one way to address the above problem that refers to combined learning of policy and a value function. The actor creates policies, selects actions, and interacts with the environment. The updates of an actor in this method are calculated by following the mathematical equations (4.7) and (4.8). The critic assesses the value function of the generated actors policy at every timestep [156].,

$$\nabla_\theta J(\theta) = E_{\pi_\theta(s)}\left[\sum_a Q(s,a)\nabla_{\theta\pi_\theta}(a|s)\right] \tag{4.7}$$

$$\theta = \theta + \alpha\nabla_\theta J \tag{4.8}$$

Asynchronous Advantage Actor-Critic (A3C) algorithm is an extension of the actor-critic method, which is more efficient to use in high dimensional state space and continuous action space. Unlike the conventional policy gradient methods, the A3C agent collects data, and the agent interacts with the environment simultaneously using multiple threads. Once the thread completes the training using the collected data separately, it then updates the global model parameters asynchronously. A3C, proposed by OpenAI, has higher performance than the well-known Advance Actor-Critic (A2C) because A3C efficiently uses GPU resources [156]. To enhance the exploration ability of the A3C model, the entropy of the policy is included in the objective function. Therefore, the policy gradient equation becomes,

$$\nabla_\theta J(\theta) = \frac{1}{T}\sum_t^T \nabla_\theta log\pi(\theta)\left(\sum_{i=1}^n \gamma^{i-1} r_{t+1} + v(s_{t+n}) - v(s_t)\right) + \beta\nabla_\theta H(\pi(s_t;\theta)) \tag{4.9}$$

Here, $\sum_{i=1}^n \gamma^{i-1} r_{t+1} + v(s_{t+n}) - v(s_t)$ demonstrate the advantage for n-steps return, and β is the weight. Most of the state-of-the-art RL algorithms are based on actor-critic methods and then expand with more sophisticated techniques. For example, when we optimize the policy using (4.7) and (4.8), the symbol α controls the step size of the update. The selection of α is important because it impacts how fast the algorithm will converge. To ensure that the policy will move steadily, we insert a constraint on the optimization problem by ensuring that the updated policy would lie in the trust region. This technique is called Trust Region Policy Optimization (TRPO) and is responsible for improving the actor. One of the recent algorithms is Proximal policy optimization (PPO) which combines the techniques of A3C and TRPO. This algorithm has a built-in mechanism (foster clipping objective function) to prevent large gradient updates that outperforms A3C on many continuous control environments such as V2V energy trading scenarios [157].

To perform well, Deep Q learning-based models, as discussed above, require a centralized computing server with massive trading experience records that need to be shared continuously to evolve. With regard to the RL agent, the learning mechanism increases the computing complexity and risk to users' data privacy. Those issues could be resolved by using an FRL algorithm.

Federated Reinforcement Learner

In FRL, local machines train a global reference model using their locally generated dataset. The LAG is responsible for aggregating the learning updates to form a new global model, which is then broadcasted to all other nodes in the network. Let $\mathcal{M} = \{\mathcal{M}_1, \mathcal{M}_2, ..., \mathcal{M}_n\}$ be the set of n local machines. At time t, assume that each local machine $\mathcal{M}_i$ trains the global model w_t^G using its local trading experiences referred to as a local dataset to create a locally updated model w_t^i. Once the local updates are transmitted to the LAG, it estimates a new global model denoted as $w_{t+1}^G = f(w_t^1, w_t^2, ...w_t^n)$. Then, the newly formed model w_{t+1}^G is send back to all local devices to replace their global reference model to train at time $t+1$.

4.3 Proposed Blockchain Based FRL System for V2V Energy Trading

The EV charging mechanism in our proposed model is adopted initially from the study in [158] that utilizes a deep Q network to schedule an EV battery for optimal charging control. In this regard, we consider several stochastic variables including EV i's arrival time (t_a), departure time (t_d), state of charge (SOC^{EV_i}), and the energy price. With those variables, we define the state space, action space, and reward function of an MDP environment.

4.3.1 State

The state s_t is composed of six variables which could be defined as follows,

$$s_t = \left\{\delta_t, SOC_t^{EV_i}, p_t^{v2v}, p_t^{Gbuy}, p_t^{Gsell}, \Omega_t\right\} \tag{4.10}$$

$$t = t + \Delta t \tag{4.11}$$

The first state variable δ_t is a scaling factor that could be determined by using equation (4.3)-(4.4) in Section II.A. The second state variable $SOC_t^{EV_i}$ represents the state of charge of the battery of EV i which should not exceed its maximum limit $SOC_{max}^{EV_i}$. Also, p_t^{v2v} determines the price per unit of electricity when trading in a V2V fashion, at time t. On the other hand, the variables p_t^{Gbuy} and p_t^{Gsell} represent grid buying and selling prices, respectively. We assume $p_t^{Gsell} > p_t^{v2v} > p_t^{Gbuy}$, which means that trading energy in the V2V scheme is more beneficial than trading with the main grid. Lastly, the variable Ω_t, represents the charging infrastructure cost while transferring per unit of electricity using modes described in Section 4.2.3, i.e., the cost of direct V2V versus assisted V2V, which may influence the EV's trading decisions, choices, and the reward. In equation (4.11), we formulate the dynamics of t as a discrete variable with regular time intervals that resets (t=0) at midnight.

4.3.2 Action Space

The action space consists of the amount of energy that an EV i trades during a specific time period t. This could be defined mathematically by,

$$\mathcal{A}_t^{EV_i} = \left\{E_{t,i}^{ch/dis} \,|\, E_{max,i}^{dis} \geq E_{t,i}^{ch/dis} \leq E_{max,i}^{ch}\right\} \tag{4.12}$$

$$-(SOC_i - SOC_{min}) \leq \frac{\mathcal{A}_t^{EV_i}}{E_{max,i}} \leq (SOC_{max,i} - SOC_i) \tag{4.13}$$

In equation (4.12), The variable $E_{t,i}^{ch/dis}$ can be either positive or negative, where a negative value indicates the battery was discharging, while positive values indicate the battery was charging. Notably, an EV is required to maintain the minimum and maximum energy levels as in equation (4.13). The strategy of selecting the trading amount depends on the expected optimal reward for an entire episode. Particularly, the episode starts when an EV checks in to a location with a trading request and ends when it departs or at midnight. For simplicity, we consider a timestep is an hour, but in reality, it could be even less than that. In the next section, we formulate the reward mechanism that can compute the expected return by the following action, $\mathcal{A}_t^{EV_i}$.

4.3.3 Reward Function Specification

The optimal energy control could be achieved by scheduling EVs actions as in equation (4.12) at a time when trading energy is cost-effective. We use the term departing to represent that an EV i will not participate in the trading system and is expected to leave the current MDP after the current time slot ends. To this end, the reward of an EV i could be defined in two ways: when an EV is departing and when it is not. The reward function can be defined mathematically as,

$$\mathcal{R}_t = \begin{cases} -\delta_t \, A_t^{EV_i} \left(p_t^{v2v} + \Omega_t\right) + \eta RA_i, & \text{if} \quad \text{depart} \\ -\delta_t \, A_t^{EV_i} \left(p_t^{v2v} + \Omega_t\right) & \text{otherwise.} \end{cases} \tag{4.14}$$

$$RA_i = \left(SOC_t^{EV_i} - SOC_{max}^{EV_i}\right)^2 \tag{4.15}$$

In equation (4.14), the value of δ_t, similar to equation (4.2), is multiplied by the amount of trading energy ($A_t^{EV_i}$) and the cost of per unit of electricity ($p_t^{v2v} + \Omega_t$) to compute the reward as long as the EV is not departing. $A_t^{EV_i}$ could be either negative or positive, and therefore, it ensures that discharging is more rewarding than charging during peak hours. The cost per unit of electricity consists of the price of per unit of electricity, and the charging service cost Ω_t. The value of Ω_t depends on the V2V topology as mentioned in subsection IV.A. In the case of direct V2V, the value of Ω_t is set to zero, but in the case of assisted V2V additional cost may be incurred. We capture such distinction in the following equation

$$\Omega_t = \begin{cases} Cost_{V2V}^{assisted}, & \text{if} \quad \text{assisted V2V} \\ 0 & \text{direct V2V.} \end{cases} \tag{4.16}$$

In (4.16), $Cost_{V2V}^{assisted}$ is the service cost value corresponding to the assisted V2V mode. For the brevity of our system implementation, we only consider direct V2V energy transfer as discussed in Section 4.2.3, and assume that there are no extra charges involved for the charging mode, so the Ω_t will be set to zero. Equation (4.15) represents the case when the EV departs, the reward mechanism includes range anxiety RA_i, which is mathematically represented by equation (4.15). Essentially, range anxiety quantifies a user's level of satisfaction resulting from the available SOC and the maximum capacity. Hence, it is important for a user to maintain a trade-off between the level of comfort and the expected reward, which can be determined by a coefficient η. Specifically, they can choose a value of η that could either increase or decrease the influence of range anxiety on the expected reward.

4.3.4 Decentralized Aggregator/ Miner Selection Mechanism

This study proposes a decentralized aggregator called mLAG that combines the LAG and the blockchain miner as a single entity, as shown in Fig. 4.2. The mLAG is designed to perform two tasks: (1) in the optimization and distributed learning layer, it combines the local learning updates, and (2) in the security layer, it adds a new block of energy transaction records into the blockchain. Basically, mLAG is a computing node that represents an EV user, who is randomly selected based on their weighted proof score computed by the PoSOC consensus mechanism. At midnight, the model calculates the cumulative proof score for the last 24 hours (denoted as $\mathcal{H}$), ${\mathcal{PS}^{\mathcal{H}}}_i$ for the EV i by following the below equation (4.17),

$$ {\mathcal{PS}^{\mathcal{H}}}_i = \sum_{\forall h \in \{0,..23\}} (SOC_i^{max,h} - SOC_i^h) + \delta(\mathcal{D}_i^h - \mathcal{C}_i^h) \tag{4.17} $$

$$ \mathcal{N}_t^{mLAG} = \underset{i}{WeightedRandom}(\mathcal{PS}_i, \forall i \in \mathcal{I})_t \tag{4.18} $$

At time t, in this case t=24, the computing node that has the highest proof score will have a higher chance to become the mLAG compared to others. Assume that $\mathcal{N}_t^{mLAG}$ is selected as mLAG as in (4.18), which is responsible for creating global model w_t^G by following equation (4.19) and broadcast to all other nodes in the network.

$$ w_t^G = \frac{1}{N} \sum_{n=1}^{N} w_{t-1}^n \tag{4.19} $$

At time $t + \mathcal{H}$, all nodes start to train the referenced global model w_t^G and share their updates to the mLAG. Then, the mLAG selects another node $\mathcal{N}_{t+\mathcal{H}}^{mLAG}$ to transfer the control/task of mLAG by following equation (4.17)-(4.18).

4.3.5 Block Mining Reward Technique

The mLAG receives a reward denoted as the mining reward for performing its service. The entire mining reward mechanism can be explained by answering several questions, including who pays the reward and to whom?; How to pay the reward?; and what is the reward amount? To this end, the mechanism proposes that every user agrees to pay a part of their profit to the mining node at time t as a mining reward (similar to the service fee). In the following equation (4.20), we present the formulation of computing mining reward,

$$\mathcal{M}_r = \sum_{\forall i \in I} \beta_t * R_t^{EV_i}, 0 < \beta_t \leq 1 \tag{4.20}$$

Where β_t determines the portion of the profit that requires to be allocated for mining rewards from all the participants. In this case, the value of β_t can be selected by the mutual agreement among the participating EVs in the energy market at every time step. The value of β_t is restricted between 0 and 1 so that it does not impact the payoff of a user negatively. However, there might be a situation where the EV users are selfish and are not willing to pay part of their profit. To address this problem, we propose to include a self-executing contract (smart contract) of blockchain that will confiscate the user's profit until the user send its share of the mining reward (part of its profit) to the miners' wallet. This is an agreement between an EV and the V2V trading system when they subscribe to the system. The entire procedure for this mechanism to prevent selfish EV users is shown in Fig. 4.5.

4.3.6 Algorithm

Algorithm (4) demonstrates the details of the proposed mechanism that enables EVs to trade energy with each other using the V2V scheme. In the proposed framework, at time t, EVs and mLAG interact with each other and train the global model locally in a distributed fashion. To this end, we utilize the Proximal policy optimization (PPO2) method of OpenAI [157] for the learning process, which is basically an extension of the well-known A2C mechanism. We select PPO2 due to its popularity and effectiveness in environments with continuous action spaces. The model assumes that all the agents start their learning process together. When the local learning process is finished, all agent sends their local model update to mLAG. The overall learning process, mLAG selection, blockchain transactions, and calculation of the reward mechanism is explained as follows,

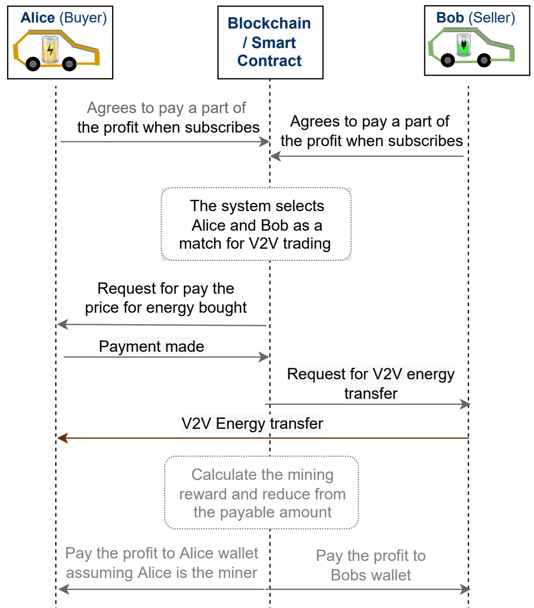


Figure 4.5: The flowchart of preventing selfish users who do not want to pay the mining reward by using a smart contract.

- At time t, when an EV subscribes to the trading system and checks in with a trading request, the algorithm starts with EV's SOC, energy demand, and expected departure time. Then, it initializes the parameters of global models that consist of value and policy networks by downloading them from the mLAG server following lines (1)-(3).
- In a local training step, every EV tries to achieve an optimal charging control to maximize their expected return for an entire episode (as in equation 4.14) by following lines (5)-(10).
- Then, the algorithm separates an action $A_t^{EV_i}$ of the EV i into buying and selling categories following line (11)-(15). With this, the algorithm computes the proof score $\mathcal{PS}_i$ of EV i for each timeslot (line (16) - (21)).
- The mLAG computes the global aggregation of the local updated weights received from individual agents, and the new global model is then broadcasted to all the participated agents following line (25)-(26).
- Next, the model selects the mLAG from the list of EVs by randomly selecting an EV based on their weighted proof score by trading V2V energy transaction following

lines (27)-(32). Then, the mining reward is calculated in line (33).

- Finally, once the V2V energy transaction details are added to a new block, it is then added to the existing blockchain following line (34).

4.3.7 Complexity Analysis

The complexity analysis determines how long it takes for an EV to decide on a trading action after observing a state s_t. This can be calculated by measuring the computational time, $O(P)$, required for a forward pass in our deep RL model. In equations (4.21) and (4.22), we formulate the complexity analysis for n EVs in the network.

$$O(\text{PoSOC}) = O(n) \tag{4.21}$$

$$O(\text{action selection - FRL}) = O(nP) \tag{4.22}$$

The time complexity of the PoSOC algorithm would be to compute EVs proof score for n times, which results in $O(n)$ followed by equation (4.1). Therefore, for an EV to execute the PoSOC and select an action using the proposed distributed V2V trading system, the computational complexity is $O(nP) + O(n)$. In other words, the increase in the number of vehicles linearly affects our proposed algorithm.

4.4 Model evaluation and Discussion

We evaluate the performance of the proposed system model by conducting several experiments considering the interactions between EV agents and the dynamic MDP environment. We assume that the entire simulation setup is based on a social hotspot (e.g., roadside office parking spot), and the EVs are equipped with a bidirectional charging setup. This enables V2V energy transfer using either direct V2V transfer or a combination of V2B and B2V technology. In this experiment, we use online reinforcement learning as the base learner of the FRL [159] in the optimization and distributed learning layer in our proposed model. We use two methods to determine how well the proposed FRL model learns: 1) extensively train the local learners and send the training update to mLAG in each iteration, and 2) test the trained model. Finally, the model is implemented on the Avalanche blockchain platform to demonstrate the real-world feasibility of the proposed model.

Algorithm 4 `Blockchain and FRL`-based Energy Trading Algorithm, and Calculation of `Proof Score` and `Payoff` for `EV`s.

Input: SOC, λ_{off}, λ_{on}, λ_{mid}, B_{e-1}.
Output: EV Payoff, B_e.

Initialize each EV with their forecasted amount of demands, supply, arrival time, and initial SOC.
Initialize the initial global model w_t^G and batch size ϕ.
Initialize the value (θ) and policy (ϕ) network parameters for the deep reinforcement agent.
    // Local and Global learning for an optimal EV charging/discharging schedule so that it could maximize its reward.
for *Each EV i* **do**
    for *Each timeslot until departure* **do**
        Select an action $A_t^{EV_i}$ with $p(s_t, A_t^{EV_i})$ for each state s_t.
        Calculate the reward R_t, $Q(s_t, A_t^{EV_i})$, and $V(s_t)$ for choosing $A_t^{EV_i}$.
        Compute advantage estimates $\hat{\text{A}}$.
        Update θ and ϕ with respect to objective function via stochastic gradient and mean-squared error of value-function.
        Update the weight with optimizer:
        $\nabla w_t = \phi(L_t^{actor}(\theta) + L_t^{critic}(\theta))$ $w_{t+1}^i \leftarrow w_t^i + \nabla w_t$
        //Calculate Proof Score
        if ($A_t^{EV_i} < 0$) **then**
            $\mathcal{D}_i \leftarrow A_t^{EV_i}$, $\mathcal{C}_i \leftarrow 0$
        else
            $\mathcal{C}_i \leftarrow A_t^{EV_i}$, $\mathcal{D}_i \leftarrow 0$
        end if
        $PS_t{=}\delta(\mathcal{D}_i^h - \mathcal{C}_i^h)$
        if (EV_i is departing) **then**
            $\mathcal{PS}_i \leftarrow \mathcal{PS}_i + (SOC_i^{max,h} - SOC_i^h) + PS_t$
        else
            $\mathcal{PS}_i \leftarrow \mathcal{PS}_i + PS_t$
        end if
    end for
    $\mathcal{PS}^* \leftarrow \mathcal{PS}^* \cup \mathcal{PS}_i$
end for
Compute the federated average using weights of n EVs in each episode: $w_{t+1}^G = \frac{\sum(w_t^1, w_t^2, w_t^3 \ldots w_t^N)}{n}$
Broadcast w_{t+1}^G to all connected EVs at $t+1$ timeslot.
// Selecting mLAG by using WRS.
for Each EV i **do**
    Let $\mathcal{PS}_i \leftarrow \frac{\mathcal{PS}_{S_i}}{\sum_{\forall j} \mathcal{PS}_j}$ be the probability of EV i to be selected as mLAG.
    $w\mathcal{P}^* \leftarrow w\mathcal{P}^* \cup \mathcal{PS}_i$
end for
$\mathcal{N}_{t+1}^{mLAG} \leftarrow$ Randomly select an EV based on $w\mathcal{P}^*$
Computing Mining reward:
$\mathcal{M}_r = \sum_{\forall i \in I} \beta_t * R_t^{EV_i}$
Add the new block to the existing blockchain by the mLAG:
$Blockchain_t \leftarrow Blockchain_{t-1} \cup Block_t$
    $PayOff, Blockchain_t$

Table 4.1: Model Parameters

Model Parameters	Distribution	Constraints
Arrival time	t_a $\mathcal{N}(7, 1^2)$	$4 \leq t_a \leq 10$
Departure time	t_d $\mathcal{N}(18, 1^2)$	$8 \leq t_d \leq 22$
Battery SOC	t_A $\mathcal{N}(0.5, 0.1^2)$	$0.4 \leq t_A \leq 0.9$
Off-peak Price/kWh (Grid)	λ_{off} $\mathcal{N}(6.8, 2.6^2)$	
Mid-peak Price/kWh (Grid)	λ_{mid} $\mathcal{N}(10, 1.5^2)$	
On-peak Price/kWh (Grid)	λ_{on} $\mathcal{N}(12.9, 2^2)$	
Off-peak hours	7pm - 7am	
Mid-peak hours	11am - 5pm	
On-peak hours	1am - 11am	
Max Battery Capacity	42 kWh (RAV4)	

4.4.1 Data Preparation

To implement the proposed model, we generate online data which is based on a real-world dataset in Ontario Electric Board, Canada [160]. The concept of charging behavior is significantly inspired by the approach taken in [158]. In Table 4.1, we include the initial parameter specifications of this experiment. The electricity price is sampled from the distribution $\mathcal{N}(0.82, 0.1^2)$, $\mathcal{N}(1.13, 0.1^2)$, and $\mathcal{N}(1.7, 0.1^2)$ for Off-peak, Mid-peak and On-peak hours, respectively, similar to the time of use rates of Toronto Hydro [161]. The arrival time of EV on the parking lot is sampled from $\mathcal{N}(7, 1^2)$ while the departure time is sampled from $\mathcal{N}(18, 1^2)$. Furthermore, we sampled the current SOC of an EV from the distribution $\mathcal{N}(0.5, 0.1^2)$. The minimum and maximum recommended SOC for each EV is considered as fractions of 0.2 and 0.8 of the battery capacity, respectively. For the simplicity of our experiment, we assume that the maximum battery capacity for all EVs is kept at 42 kWh, similar to a Toyota RAV4 [162].

4.4.2 Learning Evolution

The federated local RL model is trained for 300K episodes using OpenAI Gym and online data that creates an MDP environment in each timestep. Once an episode ends, the local model updates are shared with the mLAG to aggregate and broadcast to the network to train for the next episode. In Fig. 4.6, we show the training curve of all episodes and its expected average episode reward. In particular, we use the running average for

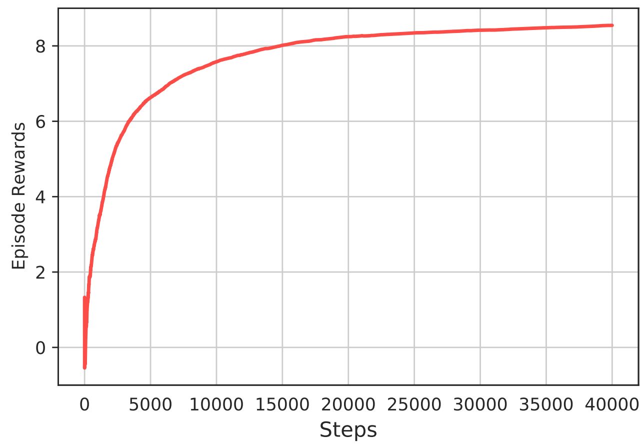


Figure 4.6: The evolution of average episode reward during the training of the model.

evaluating the episode reward. Initially, we observe that the model learns slowly with numerous trials and errors. Once the number of episodes is more than 25K, the model learns significantly and thus improves the expected return. This trend continues until the number of episodes reaches 50K (approximately) and after that, no further improvement is observed considering the current state and action space of the participating EVs.

4.4.3 Charging Control for a Single Local Learner

To demonstrate the performance of the proposed system, we select an EV that has the latest trained model for an entire episode. Fig. 4.7-(a) shows the charging control of the trained agent for each timeslot (hours) of a randomly generated episode sampled from Table 4.1 specification. The positive and negative amount represents the buying and selling action of EVs, respectively. We noticed that the EV joins the energy market with a SOC of 0.65. We observe that the EV sells its energy when the price per unit of electricity is high and buy when the price is low. Notably, the EV does not trade any energy (12:00 to 15:00 hours) when the model advises that the expected return will not be maximized with any trading.

Fig. 4.7-(b), represents the profit of the EV by performing the trading as shown in Fig. 4.7-(a). Here, we quantified the profit by calculating the difference between the amount earned by trading using the V2V scheme compared to the main grid. In this regard, every action has an impact on the returns of the EV, irrespective of whether it is buying or

selling. For example, we observe that the EV receives a profit of 27% (at 10:00 hours) and 15% (at 11:00 hours) by buying and selling energy using V2V methods, respectively.

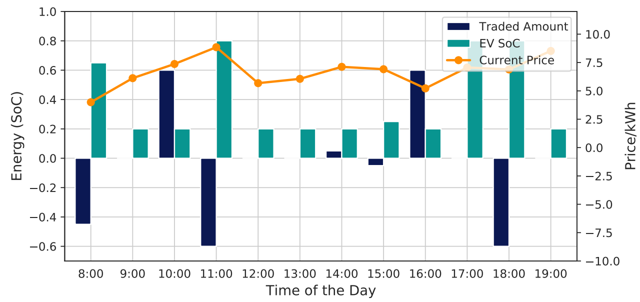
In Fig. 4.7-(c), we show the evolution of the proof score of the EV with respect to the time of the day by taking actions as in Fig. 4.7-(a) in each timeslot. Specifically, the value of δ determines how effective an EVs' action is during On-peak and Mid-peak hours. At 11:00 hours, we observe that the proof score is the highest because the EV discharge energy during this timeslot, as in Fig. 4.7-(a), to acquire more proof scores. The proof score is maximized in this instance because the technique to calculate the proof score is $\delta(\mathcal{D}_i - \mathcal{C}_i)$.

4.4.4 Illustration of a 10 EVs Energy Network

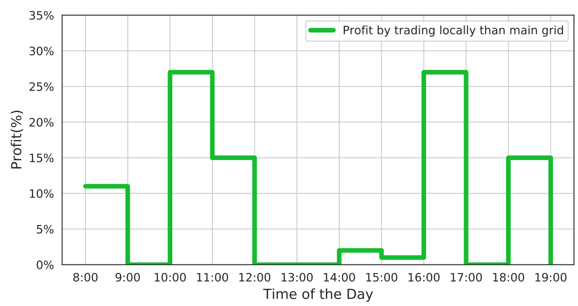
The objective of this experiment is to test and analyze how multiple agents control their energy trading in the proposed system that consists of ten EVs. In Fig. 4.8-(a), we show the start and end times of 10 EVs in the parking facility for their respective episodes. It can be inferred that different EVs have different timelines. This variation is critical to demonstrate the performance of the trained model with variable episode length.

With this timestamp, Fig. 4.8-(b) represents the corresponding EV's energy control. In fact, EVs with longer episodes exhibit more trading amount compared to the EVs who has shorter episodes. However, there are some exceptions, for example, although the episode length of EV5 is longer than that of EV2, it indicates a lower trading amount. This is because the trading depends on the initial SOC, dynamic MDP environment, and episode return. In this regard, the return depends on the value of δ, and the expected reward for future timeslots, and EVs may refrain from frequent trading actions. As a result, the benefit is twofold, firstly, the EV maximizes the proof score that enhances the probability of becoming mLAG in the blockchain, and secondly, it reduces the number of conversions (e.g., ac-dc, etc.) required in a V2V scheme.

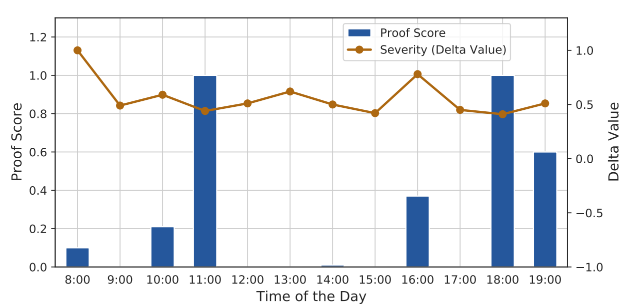
Fig. 4.8-(c) demonstrates the attained profit of the EVs in this network. The green bars represent the average profit gained by adopting trading strategies as in Fig. 4.8-(b). The EV that has the highest proof score for an episode will have a higher chance of being selected as the mLAG in the blockchain. In this episode, we observe that EV5 is selected as a miner. Hence, EV5 increases its payoff by receiving the mining reward computed by equation (4.20).



(a) Charging control.

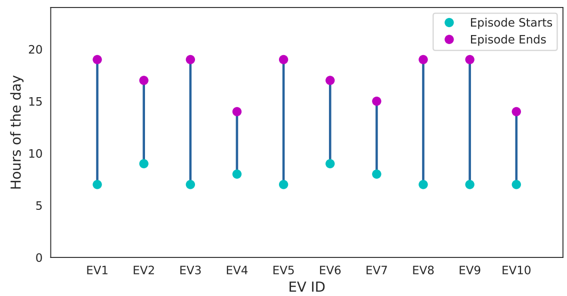


(b) Profit/payoff

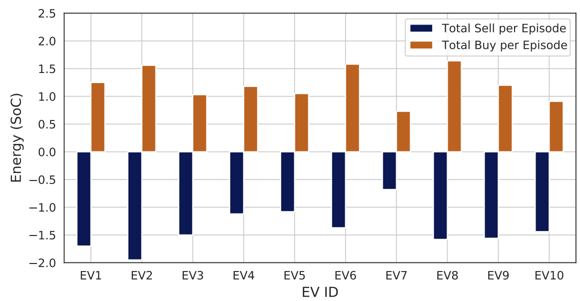


(c) Proof score distribution

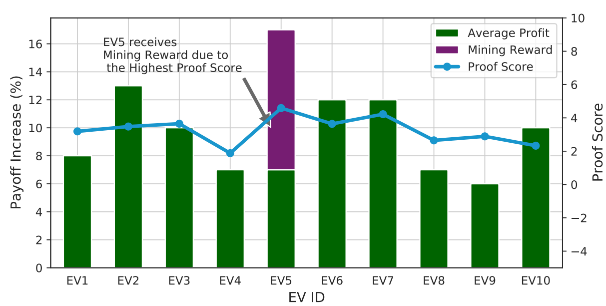
Figure 4.7: (a) The charging control of an EV for one episode using the trained model. (b) Payoff of trading using V2V scheme. (c) Proof score distribution of the EV for each time step.



(a) EV timeline.



(b) Energy Charge/Discharge



(c) Profit/payoff

Figure 4.8: (a) The timeline of the EVs in the parking facility. (b) The predicted charging control of an EV for one episode using the trained model. (c) Profit/Payoff using V2V trading and mining reward.

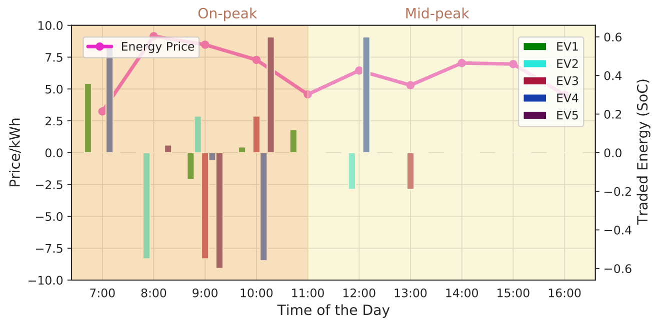


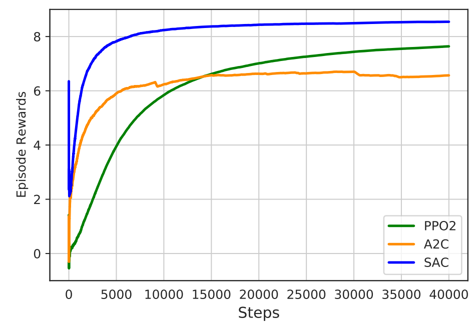
Figure 4.9: The effect of energy price on the charging/discharging control of 5EVs for a single episode (having different episode lengths).

4.4.5 Impact of Variation of Energy Price on EVs Energy Trading

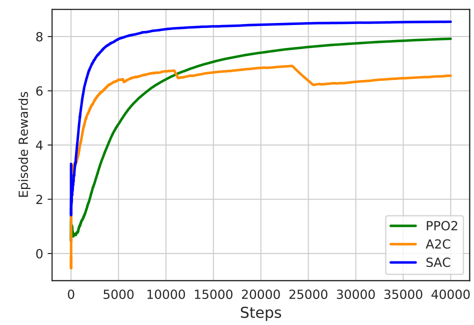
To show the effect of energy price on the performance of the proposed system, we consider 5 EVs' energy network and observe their actions (buy energy/sell energy/observing), as shown in Fig. 4.9. Under this setup, the participating EVs' episode lengths are different. We observe that the price of electricity during the On-peak hour is more than the Mid-peak hour. Therefore, the amount of charging during this timestamp is less compared to the discharging amount. This is because EVs receive less profit during On-peak hours for charging every unit (kWh) of electricity. At Mid-peak hour, we observe that the EVs energy trading reduces significantly and mostly charging compared to On-Peak hour, for example, EV2, EV3, and EV5.

4.4.6 Performance Analysis using Different Algorithms for various Data Distribution

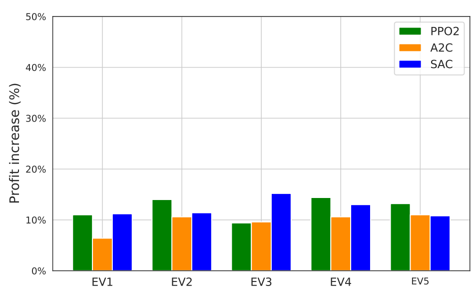
Fig. 4.10 shows the training reward convergence and their relative increase of the profit for EV users using the proposed blockchain-based FRL scheme with three state-of-the-art algorithms. Fig. 4.10-(a) shows the training evolution when the data is in a normal distribution (ND) where Soft actor-critic (SAC) shows rapid convergence with the highest average reward. On the other hand, with regards to the PPO2, the evolution of training reward gradually increases with the increase of the number of episodes. In this case, the training reward is slightly lower than that of SAC while higher than the A2C. Likewise, the reward convergence of all those three models is closely similar when the data is Uniform



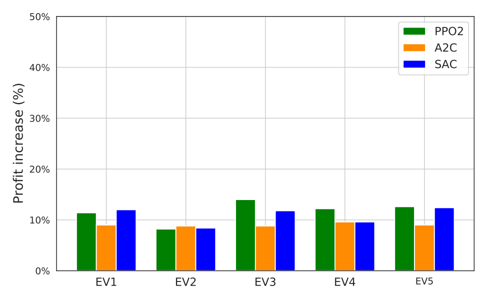
(a) Training evolution for ND.



(c) Training evolution for UD



(b) Profit for ND



(d) Profit for UD

Figure 4.10: (a) The evolution of average episode reward during the training of the model using Normal Distribution (ND) of data (b) Testing the model for the 5 EVs using ND of data (c) The evolution of average episode reward during the training of the model using Uniform Distribution (UD) of data (d) Testing profit for the 5 EVs using UD of data.

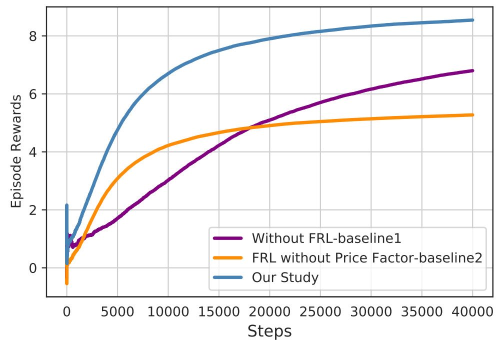


Figure 4.11: Comparison of average reward convergence using two baseline and our proposed system model.

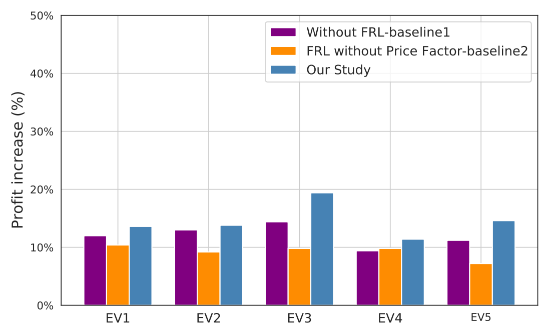


Figure 4.12: Average reward convergence for the two baselines and our proposed model.

distribution (UD) as shown in Fig. 4.10-(c). In terms of profit, we test the algorithm with an energy network consisting of five EVs and predict their probable charging control which generates profit. In Fig. 4.10-(b), we observe that all EVs receive a higher profit when the proposed model is trained using ND data by PPO2 and SAC compared to the A2C algorithm. The profit trend is similar when the proposed algorithm is trained using UD data as in Fig. 4.10-(d). Therefore, we can infer that the proposed model is consistent even though the data is in different distributions.

4.4.7 Performance Comparison between the Existing Studies and the Proposed Approach

In this experiment, we show some of the key criteria to represent the performance comparison between our proposed work and the existing studies. The work in [13] is closely similar to our proposed work, and we consider it as a benchmark that uses RL but not FRL, we refer to this model as baseline1. Moreover, this baseline study does not employ blockchain to enhance security and privacy. Next, our proposed model is compared with another system model which uses FRL but without the price factor (δ) [85]. The price factor, in this case, influences miner selection in PoSOC protocol and profit enhancement. We refer to the model as baseline2. In Fig. 4.11, we present the comparison of the running average of reward convergence for the above-discussed models. Our proposed study shows superior rewards with rapid convergence. This is because we use distributed learning with price factor (δ) results in faster convergence while enhancing profit. Fig. 4.12 shows an increase in profit for five EVs using the above-mentioned baseline models and our proposed work. However, to perform this experiment, we consider the average reward of 10 episodes for each EV. We observe that the proposed approach is most profitable for all five EVs. For example, EV3 achieves 10% and 5% more benefit compared to baseline1 and baseline2 models, respectively.

4.4.8 Security Layer Implementation

In this section, we discuss the implementation of the security layer in our proposed model using the open-source Avalanche blockchain platform. We use the Avalanche Virtual Machine (AVM) to create the entire blockchain network. As shown in Fig. 4.13, we create the blockchain network consisting of five nodes represented by five EVs in the trading system. The EVs can rent cheap cloud resources in order to become a node and act like mLAG. Avalance has several chains, namely, X-chain, P-Chain, and C-chain. Essentially, X-chain holds the AVM instance while C-chain executes the smart contract. The activity of the node can be verified by using the Postman HTTP client, as shown in Fig. 4.13. At the time t, the EV selected as a miner by following the PoSOC protocol will validate transactions in the blockchain. The proof score of the individual EVs will be credited as a token, which will be credited to the individual user's X-chain address in the wallet. Then, the proof score will be transferred to the P-chain for staking so that the EVs can participate in the mining process for the $t + 1$ timeslot.

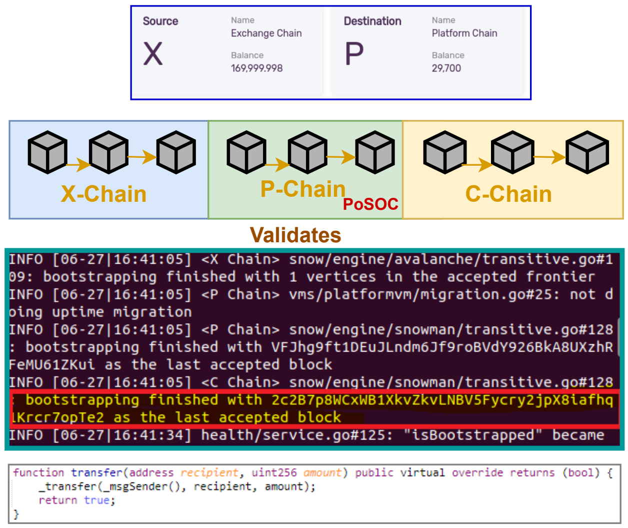


Figure 4.13: Security layer implementation on Avalanche platform.

4.5 Summary

This chapter proposes a blockchain and FRL-based V2X intelligent energy trading mechanism that enables EVs to trade energy with other parties anytime and anywhere. The proposed model presents the foundation for EVs' intelligent energy trading decisions using a secure and transparent infrastructure. The proposed PoSOC provides a feasible protocol for EVs for efficient charging control and increases financial benefit while reducing the main grid's energy management during peak hours. We show that the proposed model learns decentrally without prior knowledge to optimize users' actions for financial benefit and increase their chances to become a miner for additional rewards. Also, we implement the proposed system on the industry-standard Avalanche blockchain platform to show its potential for real-world implementation.

Chapter 5

Conclusion and Future Works

5.1 Conclusion

In this thesis, we focused on developing a secure P2P energy trading architecture that enables small-scale prosumers and EVs to participate in an open market and trade energy among themselves in a P2P fashion for economic benefit.

First, we propose a blockchain-based peer-to-peer energy trading system and coalition formation strategy among prosumers. The model offers a solid basis for integrating renewable energy sources and trading among community members, using a secure and transparent infrastructure. Through our proposed Proof of Energy Generation (PoEG) mechanism, prosumers can effectively minimize distribution losses and maximize their payoffs. Then, we demonstrate that the coalition formation algorithm ultimately leads to a stable structure of rational prosumers who aim to be chosen as miners in the underlying blockchain platform. Furthermore, we have successfully implemented the mechanism on the industry-standard Avalanche blockchain platform, thereby showcasing its potential for real-world implementation.

Second, we proposed a V2X intelligent energy trading mechanism that utilizes blockchain and FRL technology, allowing EVs to trade energy with other parties at any time and anywhere. The proposed model establishes a secure and transparent infrastructure, serving as the foundation for intelligent energy trading decisions made by EVs. With the proposed PoSOC protocol, EVs can efficiently control their charging and increase their financial benefit while reducing the energy management burden on the main grid during peak hours. Our proposed model can learn decentrally without prior knowledge to optimize users' actions for financial gain and enhance their chances of becoming a miner for additional rewards. We have also demonstrated the system's potential for real-world implementation by implementing it on the industry-standard Avalanche blockchain platform.

5.2 Future Works

This section presents some promising ideas for improving and expanding this research.

- A future research would include exploring other various types of renewable energy sources and examining the effect on the way prosumers collaborate with each other and form coalitions as the generation entirely depends on several factors (e.g., temperature, wind, price, etc.). Additionally, we did not consider non-cooperative game theory for determining optimal trading strategies, which presents a promising future direction without requiring modifications to other aspects of the model. Essentially, non-cooperative models analyze situations when participants make decisions independently. It would be interesting to compare the outcome of such a model with the combined social welfare of the group when they cooperate.

- The proposed system did not consider the residential storage system of smart homes. However, studying this system further is important as it has the potential to benefit both prosumers and the grid. Implementing it requires careful planning while taking into account technical and regulatory challenges. This requires further investigation to show the models' real-life feasibility.

- The proposed V2X trading mechanism did not consider an explicit energy pricing mechanism while determining the energy price. The pricing mechanism must be fair and transparent, and take into account factors such as the time of day, the location, and the demand for energy. This requires further study to ensure that the pricing mechanism is flexible and able to adjust to changing market dynamics and EV users' behavior.

- The proposed V2X energy model assumes that the local computing system of EVs can execute the FRL application and store the blockchain. Nevertheless, as the size of the blockchain grows over time, it can lead to a rise in computational complexity, which may jeopardize the model's feasibility. To address this issue, one possible solution is to explore the potential of a cloudlet-based system or lightning network for blockchain, which appears to be a promising idea. However, further investigation is needed to demonstrate the scalability of the model.

References

[1] International Renewable Energy Agency. Peer-to-peer electricity trading. https://www.irena.org/. (Accessed: 2023-4-2). 1

[2] Esther Mengelkamp, Johannes Gärttner, Kerstin Rock, Scott Kessler, Lawrence Orsini, and Christof Weinhardt. Designing microgrid energy markets: A case study: The brooklyn microgrid. *Applied Energy*, 210:870–880, 2018. 2, 30, 31

[3] Lei Wang, Yichao Ma, Liuzhu Zhu, Xuli Wang, Hao Cong, and Tiancheng Shi. Design of integrated energy market cloud service platform based on blockchain smart contract. *International Journal of Electrical Power & Energy Systems*, 135:107515, 2022. 2, 31, 35

[4] Olivier Van Cutsem, David Ho Dac, Pol Boudou, and Maher Kayal. Cooperative energy management of a community of smart-buildings: A blockchain approach. *International Journal of electrical power & energy systems*, 117:105643, 2020. 2, 31

[5] Magda Foti, Costas Mavromatis, and Manolis Vavalis. Decentralized blockchain-based consensus for optimal power flow solutions. *Applied Energy*, 283:116100, 2021. 2, 31

[6] Zhetao Li, Jiawen Kang, Rong Yu, Dongdong Ye, Qingyong Deng, and Yan Zhang. Consortium blockchain for secure energy trading in industrial internet of things. *IEEE transactions on industrial informatics*, 14(8):3690–3700, 2017. 2, 30, 31

[7] Sana Noor, Wentao Yang, Miao Guo, Koen H van Dam, and Xiaonan Wang. Energy demand side management within micro-grid networks enhanced by blockchain. *Applied energy*, 228:1385–1398, 2018. 2, 6, 27, 30, 31, 36, 64, 66

[8] Felipe Leno Da Silva, Cyntia EH Nishida, Diederik M Roijers, and Anna H Reali Costa. Coordination of electric vehicle charging through multiagent reinforcement learning. *IEEE Transactions on Smart Grid*, 11(3):2347–2356, 2019. 2, 31

[9] MyungJae Shin, Dae-Hyun Choi, and Joongheon Kim. Cooperative management for pv/ess-enabled electric vehicle charging stations: A multiagent deep reinforcement learning approach. *IEEE Transactions on Industrial Informatics*, 16(5):3493–3503, 2019. 2, 31

[10] Thomas Bauwens, Boris Gotchev, and Lars Holstenkamp. What drives the development of community energy in europe? the case of wind power cooperatives. *Energy Research & Social Science*, 13:136–147, 2016. 3, 35

[11] Mohammed Shurrab, Shakti Singh, Hadi Otrok, Rabeb Mizouni, Vinod Khadkikar, and Hatem Zeineldin. An efficient vehicle to vehicle (v2v) energy sharing framework. *IEEE Internet of Things Journal*, 2021. 4, 68, 70

[12] Dinh C Nguyen, Ming Ding, Quoc-Viet Pham, Pubudu N Pathirana, Long Bao Le, Aruna Seneviratne, Jun Li, Dusit Niyato, and H Vincent Poor. Federated learning meets blockchain in edge computing: Opportunities and challenges. *IEEE Internet of Things Journal*, 2021. 4, 32, 33, 70

[13] Feiye Zhang, Qingyu Yang, and Dou An. Cddpg: A deep-reinforcement-learning-based approach for electric vehicle charging control. *IEEE Internet of Things Journal*, 8(5):3075–3087, 2020. 6, 31, 95

[14] Morgen E Peck and David Wagman. Energy trading for fun and profit buy your neighbor's rooftop solar power or sell your own-it'll all be on a blockchain. *IEEE Spectrum*, 54(10):56–61, 2017. 11

[15] BloombergNEF. Top 10 energy storage trends in 2023. https://about.bnef.com/blog/top-10-energy-storage-trends-in-2023/. (Accessed: 2023-3-30). 12

[16] Andreas Poullikkas, George Kourtis, and Ioannis Hadjipaschalis. A review of net metering mechanism for electricity renewable energy sources. *International Journal of Energy and Environment (Print)*, 4, 2013. 12

[17] Feed-in tariff program. https://www.ieso.ca/sector-participants/feed-in-tariff-program/overview. (Accessed: 2021-07-16). 12

[18] Chenghua Zhang, Jianzhong Wu, Yue Zhou, Meng Cheng, and Chao Long. Peer-to-peer energy trading in a microgrid. *Applied Energy*, 220:1–12, 2018. 12, 13

[19] Esther Mengelkamp, Johannes Gärttner, Kerstin Rock, Scott Kessler, Lawrence Orsini, and Christof Weinhardt. Designing microgrid energy markets: A case study: The brooklyn microgrid. *Applied Energy*, 210:870–880, 2018. 12

[20] Thomas Morstyn, Niall Farrell, Sarah J Darby, and Malcolm D McCulloch. Using peer-to-peer energy-trading platforms to incentivize prosumers to form federated power plants. *Nature Energy*, 3(2):94–101, 2018. 12

[21] Wayes Tushar, Tapan Kumar Saha, Chau Yuen, David Smith, and H Vincent Poor. Peer-to-peer trading in electricity networks: An overview. *IEEE Transactions on Smart Grid*, 11(4):3185–3200, 2020. 13

[22] Energy prices are rising, postpone your switch. Available online: `https://vandebron.nl/`. (Accessed: 2022-09-19). 14

[23] iea. How rapidly will the global electricity storage market grow by 2026? Available online: https://www.iea.org/articles/how-rapidly-will-the-global-electricity-storage-market-grow-by-2026. 15

[24] Andrew W Thompson and Yannick Perez. Vehicle-to-everything (v2x) energy services, value streams, and regulatory policy implications. *Energy Policy*, 137:111136, 2020. 15

[25] Femina Mohammed Shakeel and Om P Malik. Vehicle-to-grid technology in a microgrid using dc fast charging architecture. In *2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE)*, pages 1–4. IEEE, 2019. 15

[26] Tiago JC Sousa, Vítor Monteiro, JC Aparício Fernandes, Carlos Couto, Andrés A Nozueiras Meléndez, and João L Afonso. New perspectives for vehicle-to-vehicle (v2v) power transfer. In *IECON 2018-44th Annual Conference of the IEEE Industrial Electronics Society*, pages 5183–5188. IEEE, 2018. 16, 75

[27] BS Umesh, Vinod Madhavrao Khadkikar, Hatem Zeineldin, Shakti Singh, Hadi Otrok, and Rabeb Mizouni. Direct electric vehicle to vehicle (v2v) power transfer using on-board drivetrain and motor windings. *IEEE Transactions on Industrial Electronics*, 2021. 16, 75

[28] Chunhua Liu, KT Chau, Diyun Wu, and Shuang Gao. Opportunities and challenges of vehicle-to-home, vehicle-to-vehicle, and vehicle-to-grid technologies. *Proceedings of the IEEE*, 101(11):2409–2427, 2013. 16, 67

[29] Daita Kobayashi, Takehiro Imura, and Yoichi Hori. Real-time coupling coefficient estimation and maximum efficiency control on dynamic wireless power transfer for electric vehicles. In *2015 IEEE PELS workshop on emerging technologies: wireless power (2015 WoW)*, pages 1–6. IEEE, 2015. 16

[30] Seyednima Khezr, Md Moniruzzaman, Abdulsalam Yassine, and Rachid Benlamri. Blockchain technology in healthcare: A comprehensive review and directions for future research. *Applied Sciences*, 9(9):1736, 2019. 18

[31] A. Stanciu. Blockchain based distributed control system for edge computing. In *2017 21st International Conference on Control Systems and Computer Science (CSCS)*, pages 667–671, May 2017. 18

[32] Rosslin John Robles, Tai-hoon Kim, D Cook, and S Das. A review on security in smart home development. *International Journal of Advanced Science and Technology*, 15, 2010. 18

[33] TJ OConnor, Reham Mohamed, Markus Miettinen, William Enck, Bradley Reaves, and Ahmad-Reza Sadeghi. Homesnitch: Behavior transparency and control for smart home iot devices. In *Proceedings of the 12th Conference on Security and Privacy in Wireless and Mobile Networks*, WiSec '19, pages 128–138, New York, NY, USA, 2019. ACM. 18

[34] Dae-Man Han and Jae-Hyun Lim. Design and implementation of smart home energy management systems based on zigbee. *IEEE Transactions on Consumer Electronics*, 56(3):1417–1425, 2010. 18

[35] Jian Shen, Chen Wang, Tong Li, Xiaofeng Chen, Xinyi Huang, and Zhi-Hui Zhan. Secure data uploading scheme for a smart home system. *Information Sciences*, 453:186–197, 2018. 18

[36] Muhammad Raisul Alam, Marc St-Hilaire, and Thomas Kunz. Peer-to-peer energy trading among smart homes. *Applied Energy*, 238:1434–1443, 2019. 18

[37] A. Dorri, S. S. Kanhere, R. Jurdak, and P. Gauravaram. Blockchain for iot security and privacy: The case study of a smart home. In *2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pages 618–623, March 2017. 18

[38] S. S. I. Samuel. A review of connectivity challenges in iot-smart home. In *2016 3rd MEC International Conference on Big Data and Smart City (ICBDSC)*, pages 1–4, March 2016. 18

[39] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. *Decentralized Business Review*, page 21260, 2008. 18, 19

[40] Danny Bradbury. The problem with bitcoin. *Computer Fraud & Security*, 2013(11):5–8, 2013. 18

[41] Merlinda Andoni, Valentin Robu, David Flynn, Simone Abram, Dale Geach, David Jenkins, Peter McCallum, and Andrew Peacock. Blockchain technology in the energy sector: A systematic review of challenges and opportunities. *Renewable and Sustainable Energy Reviews*, 100:143–174, 2019. 18, 19

[42] Juanjuan Li, Yong Yuan, Shuai Wang, and Fei-Yue Wang. Transaction queuing game in bitcoin blockchain. In *2018 IEEE Intelligent Vehicles Symposium (IV)*, pages 114–119. IEEE, 2018. 18, 21

[43] Sunny King and Scott Nadal. Ppcoin: Peer-to-peer crypto-currency with proof-of-stake. *self-published paper, August*, 19, 2012. 19

[44] Shubhani Aggarwal, Rajat Chaudhary, Gagangeet Singh Aujla, Neeraj Kumar, Kim-Kwang Raymond Choo, and Albert Y Zomaya. Blockchain for smart communities: Applications, challenges and opportunities. *Journal of Network and Computer Applications*, 2019. 19

[45] Adam Back et al. Hashcash-a denial of service counter-measure. 2002. 21

[46] A (short) guide to blockchain consensus protocols. https://www.coindesk.com/short-guide-blockchain-consensus-protocols. (Accessed: 2021-07-16). 22

[47] Fan Yang, Wei Zhou, QingQing Wu, Rui Long, Neal N Xiong, and Meiqi Zhou. Delegated proof of stake with downgrade: A secure and efficient blockchain consensus algorithm with downgrade mechanism. *IEEE Access*, 7:118541–118555, 2019. 22

[48] Christopher D Clack, Vikram A Bakshi, and Lee Braine. Smart contract templates: foundations, design landscape and research directions. *arXiv preprint arXiv:1608.00771*, 2016. 23

[49] Vitalik Buterin et al. Ethereum white paper: a next generation smart contract & decentralized application platform. *First version*, 2014. 23, 25

[50] Loi Luu, Duc-Hiep Chu, Hrishi Olickel, Prateek Saxena, and Aquinas Hobor. Making smart contracts smarter. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, CCS '16, pages 254–269, New York, NY, USA, 2016. ACM. 23

[51] Global Blockchain Compliance Hub. Usa: Smart contracts definition and legality. https://neo-project.github.io/global-blockchain-compliance-hub//united-states-of-america/USA-smart-contracts.html. (Accessed: 2023-4-25). 23

[52] Alfonso Panarello, Nachiket Tapas, Giovanni Merlino, Francesco Longo, and Antonio Puliafito. Blockchain and iot integration: A systematic survey. *Sensors*, 18(8):2575, 2018. 24

[53] The hyperledger greenhouse. Available online: `https://www.hyperledger.org/`. (Accessed: 2019-09-19). 24

[54] Vitalik Buterin et al. A next-generation smart contract and decentralized application platform. *white paper*, 2014. 24, 25

[55] Xiwei Xu, Ingo Weber, and Mark Staples. *Existing Blockchain Platforms. In: Architecture for Blockchain Applications.* Springer, 2019. 25

[56] S. Huh, S. Cho, and S. Kim. Managing iot devices using blockchain platform. In *2017 19th International Conference on Advanced Communication Technology (ICACT)*, pages 464–467, Feb 2017. 25

[57] Christian Cachin. Architecture of the hyperledger blockchain fabric. In *Workshop on distributed cryptocurrencies and consensus ledgers*, volume 310, page 4, 2016. 25

[58] Kevin Sekniqi. Ava labs white paper. 26, 63

[59] Yanni Jiang, Kaile Zhou, Xinhui Lu, and Shanlin Yang. Electricity trading pricing among prosumers with game theory-based model in energy blockchain environment. *Applied Energy*, 271:115239, 2020. 27, 30, 36

[60] Arnob Ghosh, Vaneet Aggarwal, and Hong Wan. Exchange of renewable energy among prosumers using blockchain with dynamic pricing. *arXiv preprint arXiv:1804.08184*, 2018. 27, 30, 36

[61] Yunpeng Wang, Walid Saad, Zhu Han, H Vincent Poor, and Tamer Başar. A game-theoretic approach to energy trading in the smart grid. *IEEE Transactions on Smart Grid*, 5(3):1439–1450, 2014. 29

[62] Z Yang, S Peng, Q Liao, D Liu, Y Xu, and Y Zhang. Non-cooperative trading method for three market entities in integrated community energy system. *Autom. Electr. Power Syst*, 42(14):32–39, 2018. 29

[63] Ge Wang, Qi Zhang, Hailong Li, Benjamin C McLellan, Siyuan Chen, Yan Li, and Yulu Tian. Study on the promotion impact of demand response on distributed pv penetration by using non-cooperative game theoretical analysis. *Applied energy*, 185:1869–1878, 2017. 29

[64] Mahdi Motalleb and Reza Ghorbani. Non-cooperative game-theoretic model of demand response aggregator competition for selling stored energy in storage devices. *Applied Energy*, 202:581–596, 2017. 29

[65] Giuseppe Belgioioso, Wicak Ananduta, Sergio Grammatico, and Carlos Ocampo-Martinez. Operationally-safe peer-to-peer energy trading in distribution grids: A game-theoretic market-clearing mechanism. *IEEE Transactions on Smart Grid*, 2022. 29, 31

[66] Liaqat Ali, SM Muyeen, Hamed Bizhani, and Marcelo Simoes. Economic planning and comparative analysis of market-driven multi-microgrid system for peer-to-peer energy trading. *IEEE Transactions on Industry Applications*, 2022. 29, 31

[67] Walid Saad, Zhu Han, and H Vincent Poor. Coalitional game theory for cooperative micro-grid distribution networks. In *2011 IEEE international conference on communications workshops (ICC)*, pages 1–5. IEEE, 2011. 29, 31, 43

[68] Jie Mei, Chen Chen, Jianhui Wang, and James L Kirtley. Coalitional game theory based local power exchange algorithm for networked microgrids. *Applied energy*, 239:133–141, 2019. 29, 31

[69] Sweta Malik, Maeve Duffy, Subhasis Thakur, Barry Hayes, and John Breslin. A priority-based approach for peer-to-peer energy trading using cooperative game theory in local energy community. *International Journal of Electrical Power & Energy Systems*, 137:107865, 2022. 29, 31

[70] Woongsup Lee, Lin Xiang, Robert Schober, and Vincent WS Wong. Direct electricity trading in smart grid: A coalitional game analysis. *IEEE Journal on Selected Areas in Communications*, 32(7):1398–1411, 2014. 29

[71] Liyang Han, Thomas Morstyn, and Malcolm D McCulloch. Scaling up cooperative game theory-based energy management using prosumer clustering. *IEEE Transactions on Smart Grid*, 12(1):289–300, 2020. 29

[72] Tarek AlSkaif, Jose L Crespo-Vazquez, Milos Sekuloski, Gijs van Leeuwen, and João PS Catalão. Blockchain-based fully peer-to-peer energy trading strategies for residential energy systems. *IEEE Transactions on Industrial Informatics*, 18(1):231–241, 2021. 30

[73] Shen Wang, Ahmad F Taha, Jianhui Wang, Karla Kvaternik, and Adam Hahn. Energy crowdsourcing and peer-to-peer energy trading in blockchain-enabled smart grids. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 49(8):1612–1623, 2019. 30

[74] Adam Hahn, Rajveer Singh, Chen-Ching Liu, and Sijie Chen. Smart contract-based campus demonstration of decentralized transactive energy auctions. In *2017 IEEE Power & energy society innovative smart grid technologies conference (ISGT)*, pages 1–5. IEEE, 2017. 30

[75] Vinod Pangracious. Blockchain based energy trade. In *23rd International Conference on Distributed Computing and Networking*, pages 278–283, 2022. 30

[76] Chao Liu, Kok Keong Chai, Xiaoshuai Zhang, and Yue Chen. Peer-to-peer electricity trading system: smart contracts based proof-of-benefit consensus protocol. *Wireless Networks*, 27(6):4217–4228, 2021. 30

[77] Yinliang Xu, Hongbin Sun, and Wei Gu. A novel discounted min-consensus algorithm for optimal electrical power trading in grid-connected dc microgrids. *IEEE Transactions on Industrial Electronics*, 66(11):8474–8484, 2019. 30

[78] Waqas Amin, Qi Huang, Khalid Umer, Zhenyuan Zhang, M Afzal, Abdullah Aman Khan, and Syed Adrees Ahmed. A motivational game-theoretic approach for peer-to-peer energy trading in islanded and grid-connected microgrid. *International Journal of Electrical Power & Energy Systems*, 123:106307, 2020. 31, 34

[79] Marina Dorokhova, Yann Martinson, Christophe Ballif, and Nicolas Wyrsch. Deep reinforcement learning control of electric vehicle charging in the presence of photovoltaic generation. *Applied Energy*, 301:117504, 2021. 30, 70

[80] Stoyan Dimitrov and Redouane Lguensat. Reinforcement learning based algorithm for the maximization of ev charging station revenue. In *2014 International Conference*

on Mathematics and Computers in Sciences and in Industry, pages 235–239. IEEE, 2014. 31

[81] Tao Ding, Ziyu Zeng, Jiawen Bai, Boyu Qin, Yongheng Yang, and Mohammad Shahidehpour. Optimal electric vehicle charging strategy with markov decision process and reinforcement learning technique. *IEEE Transactions on Industry Applications*, 56(5):5811–5823, 2020. 31

[82] Xiaofei Wang, Yiwen Han, Chenyang Wang, Qiyang Zhao, Xu Chen, and Min Chen. In-edge ai: Intelligentizing mobile edge computing, caching and communication by federated learning. *IEEE Network*, 33(5):156–165, 2019. 31, 33

[83] Howard H Yang, Zuozhu Liu, Tony QS Quek, and H Vincent Poor. Scheduling policies for federated learning in wireless networks. *IEEE transactions on communications*, 68(1):317–333, 2019. 31, 33

[84] Yuris Mulya Saputra, Dinh Thai Hoang, Diep N Nguyen, Eryk Dutkiewicz, Markus Dominik Mueck, and Srikathyayani Srikanteswara. Energy demand prediction with federated learning for electric vehicle networks. *arXiv e-prints*, pages arXiv–1909, 2019. 31, 33

[85] Sangyoon Lee and Dae-Hyun Choi. Federated reinforcement learning for energy management of multiple smart homes with distributed energy resources. *IEEE Transactions on Industrial Informatics*, 2020. 31, 33, 76, 95

[86] Sangyoon Lee and Dae-Hyun Choi. Dynamic pricing and energy management for profit maximization in multiple smart electric vehicle charging stations: A privacy-preserving deep reinforcement learning approach. *Applied Energy*, 304:117754, 2021. 31, 33

[87] Yuntao Wang, H Tom Luan, Zhou Su, Ning Zhang, and Abderrahim Benslimane. A secure and efficient wireless charging scheme for electric vehicles in vehicular energy networks. *IEEE Transactions on Vehicular Technology*, 71(2):1491–1508, 2021. 31, 32, 33, 75

[88] Yuntao Wang, Zhou Su, Jiliang Li, Ning Zhang, Kuan Zhang, Kim-Kwang Raymond Choo, and Yiliang Liu. Blockchain-based secure and cooperative private charging pile sharing services for vehicular networks. *IEEE Transactions on Vehicular Technology*, 71(2):1857–1874, 2021. 32, 33

[89] Omaji Samuel, Nadeem Javaid, Ahmad Almogren, Muhammad Umar Javed, Umar Qasim, and Ayman Radwan. A secure energy trading system for electric vehicles in smart communities using blockchain. *Sustainable Cities and Society*, page 103678, 2022. 32, 33, 70

[90] Yunlong Lu, Xiaohong Huang, Ke Zhang, Sabita Maharjan, and Yan Zhang. Blockchain empowered asynchronous federated learning for secure data sharing in internet of vehicles. *IEEE Transactions on Vehicular Technology*, 69(4):4298–4311, 2020. 32

[91] Yunlong Lu, Xiaohong Huang, Ke Zhang, Sabita Maharjan, and Yan Zhang. Communication-efficient federated learning for digital twin edge networks in industrial iot. *IEEE Transactions on Industrial Informatics*, 17(8):5709–5718, 2020. 32

[92] Chao Liu, Kok Keong Chai, Xiaoshuai Zhang, and Yue Chen. Proof-of-benefit: A blockchain-enabled ev charging scheme. In *2019 IEEE 89th Vehicular Technology Conference (VTC2019-Spring)*, pages 1–6. IEEE, 2019. 32

[93] Wayes Tushar, Tapan K Saha, Chau Yuen, David Smith, and H Vincent Poor. Peer-to-peer trading in electricity networks: an overview. *IEEE Transactions on Smart Grid*, 2020. 34

[94] Arezoo Hasankhani, Seyed Mehdi Hakimi, Mojtaba Bisheh-Niasar, Miadreza Shafie-khah, and Hasan Asadolahi. Blockchain technology in the future smart grids: A comprehensive review and frameworks. *International Journal of Electrical Power & Energy Systems*, 129:106811, 2021. 34

[95] Mohamed R Hamouda, Mohammed E Nassar, and MMA Salama. A novel energy trading framework using adapted blockchain technology. *IEEE Transactions on Smart Grid*, 2020. 35

[96] Yue Zhou, Jianzhong Wu, Chao Long, and Wenlong Ming. State-of-the-art analysis and perspectives for peer-to-peer energy trading. *Engineering*, 6(7):739–753, 2020. 35

[97] Wayes Tushar, Chau Yuen, Hamed Mohsenian-Rad, Tapan Saha, H Vincent Poor, and Kristin L Wood. Transforming energy networks via peer-to-peer energy trading: The potential of game-theoretic approaches. *IEEE Signal Processing Magazine*, 35(4):90–111, 2018. 36

[98] Ning Wang, Weisheng Xu, Zhiyu Xu, and Weihui Shao. Peer-to-peer energy trading among microgrids with multidimensional willingness. *Energies*, 11(12):3312, 2018. 36

[99] Faizan Safdar Ali, Moayad Aloqaily, Oznur Ozkasap, and Ouns Bouachir. Blockchain-assisted decentralized virtual prosumer grouping for p2p energy trading. In *2020 IEEE 21st International Symposium on" A World of Wireless, Mobile and Multimedia Networks"(WoWMoM)*, pages 385–390. IEEE, 2020. 36

[100] Saurabh Shukla, Subhasis Thakur, Shahid Hussain, and John G Breslin. A blockchain-enabled fog computing model for peer-to-peer energy trading in smart grid. In *International Congress on Blockchain and Applications*, pages 14–23. Springer, 2021. 36

[101] Md Moniruzzaman, Abdulsalam Yassine, and Rachid Benlamri. Blockchain-based mechanisms for local energy trading in smart grids. In *2019 IEEE 16th International Conference on Smart Cities: Improving Quality of Life Using ICT & IoT and AI (HONET-ICT)*, pages 110–114. IEEE, 2019. 36, 43

[102] LL Freris and AM Sasson. Investigation of the load-flow problem. In *Proceedings of the institution of electrical engineers*, volume 115, pages 1459–1470. IET, 1968. 38

[103] Shailendra Singh and Abdulsalam Yassine. Big data mining of energy time series for behavioral analytics and energy consumption forecasting. *Energies*, 11(2):452, 2018. 38

[104] Bin Zhou, Yunfan Meng, Wentao Huang, Huaizhi Wang, Lijun Deng, Sheng Huang, and Juan Wei. Multi-energy net load forecasting for integrated local energy systems with heterogeneous prosumers. *International Journal of Electrical Power & Energy Systems*, 126:106542, 2021. 38

[105] Krzysztof Gajowniczek and Tomasz Zabkowski. Electricity forecasting on the individual household level enhanced based on activity patterns. *PloS one*, 12(4):e0174098, 2017. 38

[106] Liyang Han, Thomas Morstyn, and Malcolm McCulloch. Incentivizing prosumer coalitions with energy management using cooperative game theory. *IEEE Transactions on Power Systems*, 34(1):303–313, 2018. 39

[107] EnergyLab. Energylab nordhavn. 39

[108] Alexandros I Nikolaidis, Charalambos A Charalambous, and Pierluigi Mancarella. A graph-based loss allocation framework for transactive energy markets in unbalanced radial distribution networks. *IEEE Transactions on Power Systems*, 34(5):4109–4118, 2018. 40, 41, 42, 55, 61

[109] Enrico Carpaneto, Gianfranco Chicco, and Jean Sumaili Akilimali. Loss partitioning and loss allocation in three-phase radial distribution systems with distributed generation. *IEEE Transactions on Power Systems*, 23(3):1039–1049, 2008. 40, 41

[110] Mohsen Khorasany, Amin Shokri Gazafroudi, Reza Razzaghi, Thomas Morstyn, and Miadreza Shafie-khah. A framework for participation of prosumers in peer-to-peer energy trading and flexibility markets. *Applied Energy*, 314:118907, 2022. 41

[111] F Sheidaei, A Ahmarinejad, M Tabrizian, and M Babaei. A stochastic multi-objective optimization framework for distribution feeder reconfiguration in the presence of renewable energy sources and energy storages. *Journal of Energy Storage*, 40:102775, 2021. 41

[112] J Mutale, G Strbac, S Curcic, and N Jenkins. Allocation of losses in distribution systems with embedded generation. *IEE Proceedings-Generation, Transmission and Distribution*, 147(1):7–14, 2000. 41

[113] Obrad M Mikic. Variance-based energy loss computation in low voltage distribution networks. *IEEE Transactions on Power Systems*, 22(1):179–187, 2007. 41

[114] Krzysztof R Apt and Tadeusz Radzik. Stable partitions in coalitional games. *arXiv preprint cs/0605132*, 2006. 42

[115] Walid Saad, Zhu Han, Mérouane Debbah, Are Hjorungnes, and Tamer Basar. Coalitional game theory for communication networks. *Ieee signal processing magazine*, 26(5):77–97, 2009. 42, 43

[116] Liyang Han, Thomas Morstyn, and Malcolm McCulloch. Incentivizing prosumer coalitions with energy management using cooperative game theory. *IEEE Transactions on Power Systems*, 34(1):303–313, 2018. 42, 43

[117] Krzysztof R Apt and Andreas Witzel. A generic approach to coalition formation. *International game theory review*, 11(03):347–367, 2009. 43

[118] Pavlos S Efraimidis and Paul G Spirakis. Weighted random sampling with a reservoir. *Information processing letters*, 97(5):181–185, 2006. 46

[119] Cong T Nguyen, Dinh Thai Hoang, Diep N Nguyen, Dusit Niyato, Huynh Tuong Nguyen, and Eryk Dutkiewicz. Proof-of-stake consensus mechanisms for future blockchain networks: fundamentals, applications and opportunities. *IEEE Access*, 7:85727–85745, 2019. 46

[120] Rosemary E Alden, Huangjie Gong, Cristinel Ababei, and Dan M Ionel. Lstm forecasts for smart home electricity usage. In *2020 9th International Conference on Renewable Energy Research and Application (ICRERA)*, pages 434–438. IEEE, 2020. 49

[121] Sarwar Sayeed and Hector Marco-Gisbert. Assessing blockchain consensus and security mechanisms against the 51% attack. *Applied Sciences*, 9(9):1788, 2019. 53, 54

[122] Juan Carlos Olivares-Rojas, Enrique Reyes-Archundia, José A Gutiérrez-Gnecchi, Jaime Cerda-Jacobo, and Johan W González-Murueta. A novel multitier blockchain architecture to protect data in smart metering systems. *IEEE Transactions on Engineering Management*, 67(4):1271–1284, 2019. 54

[123] Nurzhan Zhumabekuly Aitzhan and Davor Svetinovic. Security and privacy in decentralized energy trading through multi-signatures, blockchain and anonymous messaging streams. *IEEE Transactions on Dependable and Secure Computing*, 15(5):840–852, 2016. 54

[124] Nurzhan Zhumabekuly Aitzhan and Davor Svetinovic. Security and privacy in decentralized energy trading through multi-signatures, blockchain and anonymous messaging streams. *IEEE Transactions on Dependable and Secure Computing*, 15(5):840–852, 2018. 54

[125] Dipankar Dasgupta, John M Shrein, and Kishor Datta Gupta. A survey of blockchain from security perspective. *Journal of Banking and Financial Technology*, 3(1):1–17, 2019. 54

[126] Energy Web Foundation. We build operating systems for energy grids. 54

[127] powerledger. The power behind new energyn. 54

[128] John Albert Boudreaux. Design, simulation, and construction of an ieee 14-bus power system. 2018. 54

[129] Ankur Srivastava, Saikat Chakrabarti, Joao Soares, and Sri Niwas Singh. An optimization-based topology error detection method for power system state estimation. *Electric Power Systems Research*, 209:107914, 2022. 54

[130] Jayant Sharma, CK Sundarabalan, R Sitharthan, C Balasundar, and NS Srinath. Power quality enhancement in microgrid using adaptive affine projection controlled medium voltage distribution static compensator. *Sustainable Energy Technologies and Assessments*, 52:102185, 2022. 54

[131] Zihao Li, Wenchuan Wu, Xue Tai, and Boming Zhang. A reliability-constrained expansion planning model for mesh distribution networks. *IEEE Transactions on Power Systems*, 36(2):948–960, 2020. 55

[132] Marco RM Cruz, Desta Z Fitiwi, Sérgio F Santos, Sílvio JPS Mariano, and João PS Catalão. Prospects of a meshed electrical distribution system featuring large-scale variable renewable power. *Energies*, 11(12):3399, 2018. 55

[133] AusGrid. Solar home electricity datar. 55, 61

[134] Alex Graves, Abdel-rahman Mohamed, and Geoffrey Hinton. Speech recognition with deep recurrent neural networks. In *2013 IEEE international conference on acoustics, speech and signal processing*, pages 6645–6649. Ieee, 2013. 55

[135] Bojian Chen, Kangli Xiang, Li Yang, Qingmei Su, Daoshan Huang, and Ting Huang. Theoretical line loss calculation of distribution network based on the integrated electricity and line loss management system. In *2018 China International Conference on Electricity Distribution (CICED)*, pages 2531–2535. IEEE, 2018. 55

[136] Joseph Poon and Thaddeus Dryja. The bitcoin lightning network: Scalable off-chain instant payments, 2016. 63

[137] dwwealth. What is avalanche (avax)? https://www.drwealth.com/what-is-avalanche-avax/. (Accessed: 2023-4-25). 66

[138] Giovanni Barone, Annamaria Buonomano, Francesco Calise, Cesare Forzano, and Adolfo Palombo. Building to vehicle to building concept toward a novel zero energy paradigm: Modelling and case studies. *Renewable and Sustainable Energy Reviews*, 101:625–648, 2019. 67

[139] EVGO fast charging. Types of electric vehicles. Available online: https://www.evgo.com/ev-drivers/types-of-evs/. (Accessed: 2022-6-19). 68

[140] R Smith, S Shahidinejad, D Blair, and EL Bibeau. Characterization of urban commuter driving profiles to optimize battery size in light-duty plug-in electric vehicles. *Transportation Research Part D: Transport and Environment*, 16(3):218–224, 2011. 68

[141] Eyuphan Bulut and Mithat C Kisacikoglu. Mitigating range anxiety via vehicle-to-vehicle social charging system. In *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, pages 1–5. IEEE, 2017. 68, 75

[142] Arthur Richards and Jonathan How. Mixed-integer programming for control. In *Proceedings of the 2005, American Control Conference, 2005.*, pages 2676–2683. IEEE, 2005. 70

[143] Glenn Ceusters, Román Cantú Rodríguez, Alberte Bouso García, Rüdiger Franke, Geert Deconinck, Lieve Helsen, Ann Nowé, Maarten Messagie, and Luis Ramirez Camargo. Model-predictive control and reinforcement learning in multi-energy system case studies. *Applied Energy*, 303:117634, 2021. 70

[144] Michele Arnold, Rudy R Negenborn, Göran Andersson, and B De Schutter. Distributed predictive control for energy hub coordination in coupled electricity and gas networks. In *Intelligent Infrastructures*, pages 235–273. Springer, 2010. 70

[145] Hankz Hankui Zhuo, Wenfeng Feng, Qian Xu, Qiang Yang, and Yufeng Lin. Federated reinforcement learning. 2019. 70

[146] Shiva Raj Pokhrel and Jinho Choi. Federated learning with blockchain for autonomous vehicles: Analysis and design challenges. *IEEE Transactions on Communications*, 68(8):4734–4746, 2020. 70

[147] FermataEnergy. The nation's leader in proven v2x systems. Available online: https://www.fermataenergy.com/. (Accessed: 2022-1-11). 72

[148] Md Moniruzzaman, Abdulsalam Yassine, and Rachid Benlamri. Blockchain-based mechanisms for local energy trading in smart grids. In *2019 IEEE 16th International Conference on Smart Cities: Improving Quality of Life Using ICT & IoT and AI (HONET-ICT)*, pages 110–114. IEEE, 2019. 73

[149] Jorge Bernal Bernabe, Jose Luis Canovas, Jose L Hernandez-Ramos, Rafael Torres Moreno, and Antonio Skarmeta. Privacy-preserving solutions for blockchain: Review and challenges. *IEEE Access*, 7:164908–164940, 2019. 75

[150] Bassam Zantout, Ramzi Haraty, et al. I2p data communication system. In *Proceedings of ICN*, pages 401–409. Citeseer, 2011. 75

[151] Benedikt Bünz, Jonathan Bootle, Dan Boneh, Andrew Poelstra, Pieter Wuille, and Greg Maxwell. Bulletproofs: Short proofs for confidential transactions and more. In *2018 IEEE symposium on security and privacy (SP)*, pages 315–334. IEEE, 2018. 75

[152] Qi Feng, Debiao He, Sherali Zeadally, Muhammad Khurram Khan, and Neeraj Kumar. A survey on privacy protection in blockchain system. *Journal of Network and Computer Applications*, 126:45–58, 2019. 75

[153] Sarah LaMonaca and Lisa Ryan. The state of play in electric vehicle charging services–a review of infrastructure provision, players, and policies. *Renewable and Sustainable Energy Reviews*, 154:111733, 2022. 75

[154] Dan Zhang, Zehua Ye, Gang Feng, and Hongyi Li. Intelligent event-based fuzzy dynamic positioning control of nonlinear unmanned marine vehicles under dos attack. *IEEE Transactions on Cybernetics*, 2021. 77

[155] Timothy P Lillicrap, Jonathan J Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*, 2015. 78

[156] Xu Wang, Sen Wang, Xingxing Liang, Dawei Zhao, Jincai Huang, Xin Xu, Bin Dai, and Qiguang Miao. Deep reinforcement learning: a survey. *IEEE Transactions on Neural Networks and Learning Systems*, 2022. 78

[157] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017. 79, 83

[158] Zhiqiang Wan, Hepeng Li, Haibo He, and Danil Prokhorov. Model-free real-time ev charging scheduling based on deep reinforcement learning. *IEEE Transactions on Smart Grid*, 10(5):5246–5257, 2018. 79, 87

[159] Sascha Lange, Thomas Gabel, and Martin Riedmiller. Batch reinforcement learning. In *Reinforcement learning*, pages 45–73. Springer, 2012. 85

[160] OEB. Electricity rates. Available online: https://www.oeb.ca/consumer-information-and-protection/electricity-rates. (Accessed: 2022-06-22. 87

[161] synergyNorth. Toronto hydro - business rates and charges. Available online: https://www.torontohydro.com/for-business/rates. (Accessed: 2022-1-24). 87

[162] EVcompare.io. Toyota rav4 ev ii. Available online: https://evcompare.io/cars/toyota/toyota$_r av4_e v_2 012 - 2014$/. 87