**RELATED WORK**

**Energy Trading between multiple microgrids**

We mentioned earlier that the electricity demand is increasing and that it should balance with the constant need to limit global warming, that is why we introduced microgrids. The unidirectional design of national power grids did not allow for surplus microgrid generated power to be redistributed and sold to the primary national grid, so the process of energy trading was devised to accommodate for this problem in low voltage networks between microgrids.

In some cases, the microgrid demand is satisfied through the traditional utility grid when the connection is on-grid. Nevertheless, that means we will be paying an extra cost for them to overcome that shortage. Microgrids stand for economic optimality, which leads us to have efficient backups at reasonable and profitable cost and solve some of the technical problems we mentioned. These problems include better power quality, reduced voltage fluctuations, and a reliable system that is not affected by the utility grid outage through energy trading.

Understanding the problems facing microgrids takes us to the core of our research, the "energy transition between microgrids." The term energy trading is understood as the importing and exporting of energy in a market of retailers, producers, and vendors, including the large industrial consumers in the utility grid. It was then reinterpreted as the local energy trading that happens amongst users in the microgrid.

Energy trading can be approached from different perspectives. We can refer to it as an optimization problem where we can look at it as we look to the microgrid control either through centralized or decentralized approaches. In centralized, we have one central controller responsible for solving this problem, where it looks into a way to minimize the generation and the transportation cost of the microgrid. In contrast, we have a decentralized approach that looks into studying all the participants and the benefits that fall for each inclination.

In energy trading, what affects one affects all that is any action performed by any participant in the market will affect all. That is why we got introduced to the Game Theory (GT) technique in energy trading that in itself has a different concept to cover or try to solve the energy trading problem. GT acts in competitive situations where it takes into account that one's strategy will affect all the other strategies. We mostly use it on the decentralized structure of our microgrid, which makes it easier to check for each one's behavior. GT games are classified into Direct and Indirect games where one aims to find the ideal policy, and the other is concerned with planning a game that satisfies particular objectives. The former does not have any effect on energy trading so far. We will focus on direct games such as non-cooperative games in which each individual is looking for their benefit.

The work of Pilz %pilz2017] covers many studies regarding GT energy trading with its different properties. We find one study taking energy storage with the background of the schedule and reducing the peak to the average ratio where they plan a cost function under certain conditions that lead the system to be balanced. We then assume that each consumer's total load is the sum of the external power delivered from the primary grid. After setting up, he then proposed two different approaches, one is a static non-cooperative game where the utility sets a cost function, and the player plays a scheduled game in order for him to minimize the respective cost. Here we have a user with the advantage of selling energy back to the utility grid, known as a reverse peak. A second game was introduced that takes the utility grid as part of the game (participant) and adjusts the prices and schedules the trade a typical leader-follower structure defined by the "Stackelberg game." which proves that Stackelberg equilibrium is equivalent to minimizing the peak-to-average ratio.

Another study takes a look at situations where the traditional power station could not meet the high demand at some point, so it buys the needed energy from energy consumers (electric vehicles, renewable energy farms, and any participants involved with the central power station as an individual). The researchers proposed a non-cooperative Stackelberg game where we do not deal with each component alone. Instead, they operated a solution that serves the social benefit assuring that each component benefits from participating in energy trading. They introduced a price model where the price can differ for different energy consumers. Henceforth, the authors applied an iterative algorithm to minimize the cost for the central power station and, at the same time, maximize the sum of utility functions of energy consumers.

Another research covered the transition of energy among the MGs. The trade does not happen directly with each other but instead tries to trade surplus energy with the market and request the deficiency as well. This multileader-multifollower Stackelberg game proposed, the sellers act as leaders and the buyers as followers in which the surplus energy is proposed by the leaders to the followers proportionally to the bids each buyer has placed. This method leads us to know that the best solution for this scenario depends on bids given and the number of players in the game. Because of the expanding rivalry between the purchasers, the worth monotonically diminishes when the quantity of purchasers increases. Simultaneously, the aggregate of the utility qualities for the dealers' increases, since more costumers permit them to sell more.

Another Stackelberg seller-buyer structure among MGs was taken into consideration as the ones before. However, to make the model more expressive, the author encompasses the known structure to the Bayesian game. In this type of game, our knowledge is incomplete, and we do not have full awareness of the game aspects and players' states, meaning that each player is private about their information. In this case, we take the players as normal or abnormal, the emergency state in which the sellers are less profound to sell energy and value the stored energy. From the buyers' point of view, they tend to bid more to ensure the requested energy delivery. We build on the last study by proposing a communication link between respective MGs where a weighting variable is used to express the relation between them. Precisely the conditional probability distribution over the condition of the player is classified as a two-stage technique. In stage one, each Mg estimates the state based on the players' given messages; the second stage updates the estimates based on information gathered from the close neighbors in the structure, looking into increasing the trust within the network showing and the partial trusted information. In the end, a debate is held questioning whether this will increase the power quality but was left for further work.

Unlike the late researches here, the central unit does not only communicate, but it works as a distributor or gatherer for the energy that is traded among the MGs. Also, no scheduling scheme is proposed to pay the sellers. By providing energy to the system, the respective MG collects points that increase the contribution value. If this MG runs into a deficit of energy, its high contribution value will give it a more significant chunk of energy given by the rest. The distributor sets that in order to maximize the social welfare function. Knowing the distribution mechanism, the game here deals with the remark of how much energy to request directly proportional to this and inversely proportional to the contribution value it gets. Furthermore, each buyer is given a stage in the queue in which h should try to be served earlier to minimize what is requested from the utility grid. In this case, not enough surplus energy to serve, we have nash equilibrium property that even if participants deviate from it, the other does not be impacted negatively.

For security reasons, all correspondences are composed of the central unit. Seen from any of the MGs, this prompts a fragmented data game, as no one thinks about the systems and settlements of the others. All the more speciﬁcally, the author divides the MGs into merchants and purchasers and plans a two-phase Stackelberg game in which every one of these gatherings attempts to ﬁnd their best activities by methods or reinforcement learning algorithms. The same classification and giving principle based on proportionality are added; this implies there are two utility capacities, one for each group of buyers and sellers without the knowledge of the other players. It shows that the learning algorithm here converges to the best reply, which is the same as the solution to the sellers' and buyers' optimization, respectively. In comparison, the iteration solutions earlier this take 100 times more to converge to Nash Equilibrium.

If we looked at the energy exchange proficiencies, we find some focused on selling the energy back to the conventional grid. At the same time, others took into account two types of participants, the sellers and the buyers. In most cases, the energy transition happens in secondary structure as it does not happen between individuals but happens through an operator, a third party that leads us to not fully decentralized scenario.

On the other hand, most of the utility functions taken by the games are focused on the monetary function perspectives from looking into the cost of storing the energy to the cost of transmission of energy between different parties. On the other hand, the utility function did not look into the price function but instead looked into the ratio between allocated energy and requested energy. Other approaches acted upon an auction algorithm where the buyers and the sellers are balanced in the market.

In all scenarios, the customers were referred to as sellers or buyers despite having surplus energy or deficit. In the above models, there was a shortage in models that combine a high-quality demand analysis with the RE generation in energy trading. Most of this researcher proposed what they call blue-sky approaches with "reinforcement learning" and "contribution-based" energy trading. Furthermore, all those authors lacked in the long-term assessable suggestions opposing the merely one-day ahead analyses in energy trading. We use a reinforcement learning algorithm that works on solving the situation without prior information about the microgrid. As achieving Supply-Demand equilibrium is complicated when considering the non-formality of the RES's and many studies were proposed in market-based energy trading among microgrids to utilize DERs across the network fully %ali2009electricity]

The idea of microgrids replacing conventional power grids in rural areas has been the subject of research. B. M. Sivapriya et al. %mothilal2018pv] worked with the problem of microgrid design using the center of moment approach to the placement of PV panels on the network providing case studies for their designs on villages in India. Murenzi et al. %murenzi2015case] worked in Africa, introducing Microgrids as a viable method to electrify sub-Saharan Africa. They showed that in a typical Rwandan village, the installation of a microgrid with PV, batteries, and a micro-hydro is a better financial alternative than extending the national power grid transmission to reach the village.

Applications of Reinforcement Learning in smart grids and microgrids vary, A smart building energy management algorithm%kim2018reinforcement] that uses a Markov decision process to model the smart building. The algorithm controlled included interactions with the utility grid and internal RES. The algorithm used Q-Learning to make decisions on energy dispatch actions achieved better energy costs in the building against multiple pricing policies. Mocanu et al. %mocanu2016unsupervised] created a deep belief network that improved the performance of standard reinforcement learning algorithms. They namely worked on SARSA and Q-learning, in the context of predicting energy in a smart building, the algorithm can generalize a learned behavior model into any other building without any specific history of that building. Leo Raju et al.% raju2015reinforcement] proposed a model-free reinforcement learning algorithm (Q-learning) to solve the optimal dispatch problem, which concerns finding the best combination of available power resources to provide the required load with minimal cost. Their algorithm converged to the optimal solution and provided adaptability in dynamic situations and unforeseen load management.

Fabrice et al.% lauri2013managing] proposed an algorithm to fully control power flow between a multi-storage Microgrid mapping it as a Multi-Agent System (MAS) and using Multi-Agent Reinforcement Learning to solve the problem. They produced results showing that a centralized control; unit for the microgrid is not needed. The algorithm can achieve the minimal cost of drawing power from the primary grid and achieve most grid independence. Finally, Xiao et al. %xiao2018reinforcement] proposed an energy trading game between different microgrids intending to achieve the Nash Equilibrium without knowing the generation and load demand of the other microgrids using a DQN-based energy trading strategy achieving an improvement of 22.3% in the utility of the microgrid.

**References**

%1]Sutton S. Richard, Barto G. Andrew Reinforcement Learning an Introduction 2nd edition 2018

%2]Watkins, 1989

%3] Matthias Pilz, and Luluwah Al-Fagih “Recent Advances in local energy trading in the smart grid based on game-theoretic Approaches “ arXiv:1702.02915v2 %cs.GT] 28 Sep 2017

%4] S. M. Ali, “Electricity trading among microgrids,” Master’s thesis, The  
University of Strathclyde, Glasgow, UK, 2009. %Online]. Available:  
http://www.esru.strath.ac.uk/Documents/MSc 2009/Ali.pdf

%5]PV Microgrid Design for Rural Electrification

Sivapriya Mothilal Bhagavathy 1 and Gobind Pillai

%6] The Case for Microgrids in Electrifying Sub-Saharan Africa

Jean Pierre Murenzi, Taha Selim Ustun

%7] Reinforcement Learning Based Energy Management Algorithm for Smart Energy Buildings

Sunyong Kim and Hyuk Lim

%8] Unsupervised energy prediction in a Smart Grid context usingreinforcement cross-building transfer learning

Elena Mocanu∗, Phuong H. Nguyen, Wil L. Kling1, Madeleine GibescuDepartment

%9]. Reinforcement Learning in Adaptive control of power system generation (Leo Raju, Milton R S, Swetha Suresh, Sibi Sankar)

%10] Managing power flows in microgrids using multi-agent reinforcement learning (Fabrice Laurie, Gillian Basso, Jiawei Zhu, Robin Roche, Vincent Hilaire, Abderrafiaa Koukam)

%10] Reinforcement Learning-based Energy Trading for Microgrids

Liang Xiao\_, Xingyu Xiao\_, Canhuang Dai\_, Mugen Peng†, Lichun Wang‡ and H. Vincent Poor§,