**RELATED WORK**

**Energy Trading between multiple microgrids**

We mentioned earlier that the demand for electricity is incrementing and that it should balance with the constant need to limit global warming that’s why we introduced microgrids.The unidirectional design of national power grids didn’t allow for surplus microgrid generated power to be redistributed and sold to the main 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 demand of the micro grid is satisfied through the traditional utility grid when connection is on-grid. But that means we will be paying extra cost for the to overcome the shortage . Microgrids stand for the economic situations which leads us to have efficient backups at reasonable and profitable cost and solve some of the technical problems we mentioned we have like better power quality ,reduction of the voltage fluctuations as well as a reliable system that isn’t effected by the utility grid outage through energy trading.

This takes us to the core of our research the “energy transition between microgrids”; The term energy trading is taken as the importing and exporting of energy in a market of retailers ,producers and vendors including the large industrial consumers in the conventional utility grid it was then taken to refer to the local energy trading that happens amongst users in micro grid.

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

In energy trading what effects one effects all i.e any action performed by any participant in the market will effect the all that’s why we got introduced to the Game theory(GT) technique in energy trading that in itself has different concept to cover or try to solve the the energy trading problem . GT act in competitive situations were it takes into account that the strategy of one will effect all the other strategies.We mostly use it on decentralized structure of our microgrid that makes it easier to check for the behavior of each individual on its on.GT games are classified into i) Direct games ii) indirect games where one aims to find the ideal policy and the other is concered with planning a game that satisfy certain objectives.the former didn’t have any effect on the energy trading so far we will focus on the direct games.such as non-cooperative games in which each individual is looking for his own benefit.

Game theory Taxonomy leads us to the four most concerned with our study properties i) frequency of play ii) chronology of play iii)Awareness of players iv) Knowledge of players.

1. Frequency of play: we usually compare the games that are played once and the once that are played iteratively as we repeat the game with the same apponent the actions may differ .
2. Chronology of play: takes us to sequential or simultaneous games ;where sequential we have the players moving in turn till the end and the outcome is defined and action differ in each palyer game while in the simultaneous games we don’t react to the opponent all payers choose same action at same time and its known as (one-shot ).
3. Awareness of players :perfect and imperfect game were one the players knows all the moves that are about to be done and the past actions as well while the imperfect is not knowing and not being aware of the move before it happens.
4. Knowledge of players:last we mentioned the imperfect game and that not showing full information of the game and such a game where you don’t know the strategies of the game we refer to as Bayesian game .

In [pilz2017] a lot of studies were covered in regard to GT energy trading with its different properties ;we find one study taking energy storage with background of the schedule and reduction of the peak to average ratio.were they plan a cost function under certain conditions that lead the system to be balanced .we then assume that the total load of each consumer at some point is the sum of the external power delivered from the main grid.After setting up he then proposed two different approaches on is a static non-cooperative game where the utility set a cost function and the player play a scheduled game in order for him to minimize the respective cost here we have user with the pro of selling energy back to the utility grid known as reverse peak so the second game was introduced and it takes the utility grid as part of the game (participant) and then adjust the prices and schedule the trade a typical leader –follower structure which define the “Stackelberg game”. This proves that stackelberg equilibrium is equivalent to minimizing the peak-to average ratio.

Another study took a look at situation were the traditional power station wasn’t able to muddle through with the high demand at some point so it further buy the needed energy from energy consumers (electric vehicles, renewable energy farms and any participants involved with the central power station as individual).The researcher proposed a non-cooperative stackelberg game were we don’t deal with each individual alone but in instead operate a solution that serves the social benefit assuring that each participants is benefiting from involving in the energy trading, introducing a price model where the price might differ for different energy consumers. The author henceforth 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.

What's more a research covered the transition of energy among the MGs where the trade doesn’t happen directly with each other but instead try to trade the surplus energy with the market and request the deficiency as well from it . This multileader-multifollower stackelberg game proposed ,the sellers act as leaders and the buyers as followers in which the the surplus energy is proposed by the leaders to the followers propotional to the bids each buyer has placed .which leads us to know that the best solution for this scenario depends on the 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 increments. Simultaneously, the aggregate of the utility qualities for the dealers increments, since more costumers permit them to sell more.

Another stackelberg seller-buyer structure among MGs was taken into consideration like the ones before but here instead to make the model more expressive the author encompasses the known structure to the Bayesian game we mentioned in the GT properties where our knowledge is incomplete and we don’t have full awareness buy the game aspects and players states in means that each player is private about his information.In this case we take the players as normal or abnormal that is the emergency state in which the sellers are less profound to to sell energy and value the stored energy;from buyers point of view they tend to bid more to ensure the deliver of the requested energy .Adding to the last study another communication link is proposed 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 to two stage technique in stage one each Mg estimate the state based on the players given messages the second stage henceforth update the estimates based on information it gathered from the close neighbors in the structure ;looking into increasing the trust within the network showing the partial trused information in the end they debate that this will increase the power quality but was left for further work .

Unlike the late researches here the central unit doesn’t 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 the system the respective MG collects point that increases the contribution value.If this MG run into deficit of energy the high contribution value will gives it benefit of having the bigger chunk of energy given by the rest. That is set by the distributor 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 propotional 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 in order to minimize what is requested from the utility grid.in case not enough surplus energy to serve we have nash equilibrium property that even if particpants deviate from it the other don’t be impacted negatively .

The advertise/aggregator through which they are empowered to exchange excess energy. For security reasons, all correspondences are composed through thecentral unit

Seen from the point of view of 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 divide the MGs into merchants and purchasers, and plan a two phase Stackelberg game in which every one of these gatherings attempts to ﬁnd their best activities by methods or reinforcement learning algorthim. The same classification and giving principle based on proportionality is added. This implies there are two utility capacities, one for each group of buyers and sellers without the knowledge of the the other players it shows that the learning algorthim here converges to the best reply which is same as the solution to the optimization of the sellers and buyers respectively.in comaparsion the the iteration solutions earlier this take 100 time more to converge to nash equilibrium.

If we looked at the energy exchange proficiencies we discussed earlier in different games we find some focused on selling the energy back to the conventional grid while other took into account two type of participants the sellers and the buyers. The energy transition in most cases happens in indirect structure as it doesn’t happen between individuals but happens through an operator a third party that leads us to not fully decentralized scenario.

In another hand most of the utility functions taken by the games focused in the monetary function prespective from looking into the cost of storing the energy to cost of transmission of energy between different parties in other the utility function didn’t look into the price function but instead looked into the ratio between allocated energy and requested energy . Other approaches that acted upon and auction algorithm where the buyers and the sellers are balanced in the market.

Also in all scenarios the customers were referred to as sellers or buyers despite if they have surplus energy or deficit.In the above models there was a short in a model that combines a high quality demand analysis with the RE generation in the context of the 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 were short in the long term assessable suggestions opposing the merely one-day ahead analyses in energy trading.

we use reinforcement learning algorithm that works on solving the situation without prior information about the microgrid .As achieving Supply-Demand equilibrium is difficult when considering the non-formality of the RES’s and a lot of studies were proposed in market based energy trading among microgrids to fully utilize DERs across the network [ali2009electricity]

The idea of microgrids replacing convential powe grids in rural areas has been the subject of research, B. M. Sivapriya et. al[mothilal2018pv] worked with the problem of micogrd 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 Rawandan village, instlatoin of a microgrid with PV, battaries and a microhydro is a better financial alternative than extending the national power grid transmission to reach the village.

Applications of Reinforcement Learning in smat grids and microgrids is varying, A smart building energy management algorithm[kim2018reinforcement] that uses a markov decision process to model the smart building including interactions with the utility grid and internal RES, and the algorithm used Q-Leraning 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 namly SARSA and Q-learning in the context of predicting energy in a smart building, the algorithm can generalize a learnt behaviour 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 with finding the best combination of available power resources to provide required load with minimal cost. Their algorithm converged to the optimal solution as well as providing 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 and the algorithm can achieve minimal cost of drawing power from the main grid and achieve most grid independence. Finally Xiao et. al[xiao2018reinforcement] proposed an energy trading game between different microgrids with the aim of achieving 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§,